Title
Impacts of Cloud Microphysics on Extreme Precipitation and Lightning

Permalink
https://escholarship.org/uc/item/0sw3x7qx

Author
Charn, Alexander Benedict

Publication Date
2020

Peer reviewed|Thesis/dissertation

Impacts of Cloud Microphysics on Extreme Precipitation and Lightning

by

Alexander B. Charn

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Earth and Planetary Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor William D. Collins, Chair
Professor David M. Romps
Professor John C.H. Chiang

Summer 2020

Impacts of Cloud Microphysics on Extreme Precipitation and Lightning

Abstract

Impacts of Cloud Microphysics on Extreme Precipitation and Lightning

by

Alexander B. Charn

Doctor of Philosophy in Earth and Planetary Science

University of California, Berkeley

Professor William D. Collins, Chair

The microphysical processes involving water droplets and ice crystals in clouds are too small to be explicitly simulated by climate and weather models. Nevertheless, they play a critical role in the large-scale energy balance of the Earth and its atmosphere, as well as smaller-scale phenomena such as storms. This dissertation examines the impact of microphysics on the latter, specifically extreme precipitation and lightning. Climate change threatens to exacerbate such events, making the understanding of such extremes crucial.

We focus primarily on the effects of microphysical processes as they are simulated in a super-parameterized climate model, which is better suited to studying clouds and the associated extreme weather events than conventional models. We find statistically significant differences in extreme precipitation rates via two separate mechanisms when replacing one commonly used microphysics parameterization with another. We also find that the sign of changes in lightning flash rates with global warming depends on the microphysics representation used. Finally, we employ observations to address a longstanding question about the necessity of ice as a precursor of lightning. With the data available it is concluded that there is insufficient evidence to suggest that thunderstorm electrification can occur in the absence of ice.

For Cosmo

Contents

List of Figures

List of Tables

Acknowledgments

I first and foremost thank Bill Collins, who agreed to serve as my advisor five years ago. I will look back fondly on our weekly meetings, where his intelligence and humor frequently shone as he pushed me to become a better scientist. Perhaps Bill's greatest gift to me was giving me the confidence to carry out research independently, which I scarcely dreamed of possessing as an undergraduate.

I am also indebted to David Romps, whom I feel I will never be able to thank enough for getting me interested in atmospheric physics after taking me on as an undergraduate researcher all those years ago. In particular, I am thankful for David sparking my interest in lightning, which would come to encompass half my dissertation. Lastly, I am grateful for the numerous interactions, both within and without weekly group meetings, with Ben Fildier, Jake Edman, Jake Seeley, Nadir Jeevanjee, Wolfgang Langhans, Rusen Oktem, Suqin Duan, Jordan Mizerak, Dana Lapides, Ellyn Gray, Yi-Chuan Lu, Nathaniel Tarshish, Qindan Zhu, Katie Latimer, Leo Selker, and Alec Shelley.

I am grateful to have interacted with and gotten to know many of the brilliant EPS faculty, notably John Chiang, for serving on both my qualifying exam and my dissertation committee, and Bruce Buffett, for serving as my chair on the former. I am also lucky to have had many wonderful collaborators, including Hossein Parishani, Travis O'Brien, Mark Risser, Ben Timmermans, and Chris Paciorek.

I must thank numerous staff members for helping make my graduate program as smooth an experience as possible. Thanks to Margie Winn for serving as our Graduate Student Affairs Officer and for always remembering my Star Wars sweater that I wear at Santa Barbara's Day. Thanks to Rosie Davis and Kim Sopher for helping ensure I got paid on time and in general for being friendly faces whenever I shuttled up the hill to LBL. Thanks to Rachel Kowalik for assisting me with reimbursements and classroom reservations and for having Jolly Ranchers in her candy jar. Thanks to Bryan Taylor for helping me with my laptop issues! I would also like to thank the many individuals working behind the scenes at NERSC Consulting and Support, for helping me get SPCAM running amidst Cori software changes through countless support tickets.

It is a perhaps underappreciated fact that graduate students are oftentimes alone in their studies. Having a departmental community was invaluable in my having a positive experience within EPS. For me this started with our 1st-Year Fellowship, going to trivia nights and getting dumplings before EPS 375. Thanks to Tanis Leonhardi, Anna Clinger, Will Davis, and Rodrigo Chi for being great roommates in Albany.

I am truly blessed with friends and family who have made every moment in life worth living. I thank my regular hitting partners, including Teddy and R.J., as well as my tennis team captain John Andorf, for letting me get my mind off being a student in my favorite competitive way. I thank Natalie, Chaaru, Ross, Jason, Kyra, and Angelo for the escape rooms and virtual gaming marathons and for putting up with said competitiveness, during, for example, Overcooked! I am forever grateful to my parents Christina and Kenny, whom I can never repay for their love, enthusiasm, guidance, and unshakeable support for whatever

I choose to do. I am also forever grateful to my girlfriend Natalie for her love, our lunches and adventures around campus and California as a whole, and, honestly, for putting up with a lot.

Lastly, I thank Berkeley, both the city and the campus, for being my home away from home. I have been fortunate enough to be a student here for 9 years, meeting a bevy of fascinating, intelligent, and passionate people along the way. May it always strive to be a beacon of light for humanity. Fiat Lux.

This dissertation was typeset using the `ucbthesis` LaTeX template.

Chapter 1

Introduction

1.1 Microphysics and Extreme Precipitation

Microphysical processes encapsulate the formation of cloud droplets and ice crystals, which collide and grow into raindrops, snowflakes, graupel, and hail. The latent heat released during condensation, freezing, and deposition has implications for the energetics of the parent cloud, the broader storm within which the cloud is nested, and even the Earth system as a whole. The net radiative cooling of the atmosphere is balanced largely by turbulent fluxes of latent enthalpy, a result of evaporation from the surface. A balanced water budget on sufficiently long timescales implies that precipitation equals evaporation, thus completing the link between the former and Earth's energy balance. A consequence of this is that global-mean precipitation will increase with warming at a rate of 1-3%/K (Jeevanjee and Romps, 2018). Local extreme rain rates are expected to increase even faster, around 7%/K (Allen and Ingram, 2002), following the Clausius-Clapeyron relation, which describes how the saturation vapor pressure of water increases with temperature.

There exist a variety of ways in which microphysical processes are represented in climate and weather models. Since keeping track of all hydrometeors (the generic term for all particles derived from water) is infeasible, microphysical parameterizations generally divide them into discrete categories, typically cloud droplets, cloud ice, raindrops, snow, and graupel/hail (one or both may be included). Next, modelers must decide how to represent the size distribution of each individual species, which is important for both diffusion and collision-coalescence processes. Bin schemes contain prognostic equations for a number of discrete bins that divide up the particle size spectrum. Such an approach is often prohibitively computationally expensive. More commonly used are bulk schemes, which prescribe the functional size distribution a priori, typically as a gamma or exponential distribution. Here, prognostic equations are written for bulk, i.e., grid-box, quantities such as mixing ratio, which relates to mass conservation, and number concentration for each species.

Different bulk schemes prognose different numbers of quantities, or moments, of the hydrometeor distribution. The simplest and least computationally demanding parameteri-

zations have one predicted moment, usually mixing ratio, but some physical processes are not realistically represented with this approach. For example, accretion, or riming, of supercooled water droplets by ice particles leads to an increase in mass, but not number concentration, of the latter. Since the number concentration is diagnosed from the mixing ratio, the 1-moment scheme would predict an artificial increase in the amount of ice particles. Another example is raindrop breakup, which should affect (increase) the number concentration, but not the mixing ratio, of the raindrops.

The differences in the parameterizations described above can manifest in storm-level quantities, such as precipitation rate and propagation speed. Larger hydrometeors have faster terminal velocities, are less likely to evaporate, and thus might be expected to result in higher (extreme) rain rates over a particular area. The absorption of heat from the environment during re-evaporation is responsible for cold pool development, which then has ramifications for propagation speed, particularly for organized convection, such as mesoscale convective systems. This in turn affects how long a storm remains in one location. Thus, while we have some idea as to how precipitation will change between different climates, another important consideration is whether climate models can even represent the magnitude of extreme precipitation rates today. Chapters 2 and 3 are devoted to this question, with the former focusing on the contiguous United States and taking advantage of a rain-gauge dataset with high temporal resolution to compare model output with. The latter generalizes this analysis to the entire globe.

1.2 Microphysics and Lightning

The collisions between hydrometeors are thought to be responsible for charge separation within clouds, where sufficiently strong electric fields result in lightning. There are two main theories that attempt to elucidate the predominant form of such collisions. The inductive mechanism takes advantage of a preexisting electric field, such as the fair-weather field, which is oriented such as to drive positive charges downward toward the surface. Collisions between originally neutral, but polarized, hydrometeors result in positively charged particles aloft and negatively charged ones below. Since rebounding is critical in this process of charge separation, it is thought that ice must be involved since the coalescence efficiency of water-water collisions is generally high. However, the electrical conductivity of ice is generally not high enough to permit a complete transfer of charge in the brief time of interaction. Thus, collisions between water and ice particles are thought to be the predominant form of inductive charge transfer. There is, however, at least one limitation with the inductive theory: observational (e.g., Christian et al., 1980) and modeling (e.g., Kuettner et al., 1981) studies have concluded that it is not able to create a field strong enough to be responsible for the first lightning strike within a storm. For this, a non-inductive mechanism is believed to be dominant. Here, graupel and ice particles are thought to collide and rebound in a similar manner as described above, though the exact nature of the charge transfer is less agreed upon than in the inductive case.

Typical climate models, even those that are cloud-resolving, do not parameterize electric fields in any way. Any predictions of lightning are generally done via observed correlations with other atmospheric variables, such as convective cloud-top height, convective available potential energy, precipitation, or ice fluxes prognosed via a microphysics scheme. In this last case it is clear that the chosen microphysical representation can play a critical role in the simulation of lightning and its expected changes with global warming. Chapter 4 is devoted to addressing both the ability of such ice processes to capture the global flash rate distribution and any sensitivity to the choice of microphysics parameterization.

Finally, while most proposed theories of thunderstorm electrification rely in some way on ice particles, particularly precipitating ones, there have been purported eyewitness accounts of lightning occurring in clouds entirely below the local freezing level height. This would suggest that ice is in fact not essential, and some other factor, such as convective ascent velocity, is the true cause of electric field generation. Since there have been relatively few documented instances of such warm-cloud lightning, recent attempts via aircraft to confirm the phenomenon's existence ended, perhaps unsurprisingly, in failure. In Chapter 5 we take advantage of the recently launched Geostationary Operational Environmental Satellite-16, as well as three lightning datasets, to conduct a systematic search for warm-cloud lightning.

Chapter 2

Microphysical Sensitivity of Superparameterized Precipitation Extremes in the Contiguous United States

Superparameterized (SP) global climate models have been shown to better simulate various features of precipitation relative to conventional models, including its diurnal cycle as well as its extremes. While various studies have focused on the effect of differing microphysics parameterizations on precipitation within limited-area cloud-resolving models, we examine here the effect on contiguous-US (CONUS) extremes in a global SP model. We vary the number of predicted moments for hydrometeor distributions, the character of the rimed ice species, and the representation of raindrop self-collection and breakup. Using a likelihood ratio test and accounting for the effects of multiple-hypothesis testing, we find that there are some regional differences, particularly during spring and summer in the Southwest and the Midwest, in both the current climate and a warmer climate with uniformly increased sea-surface temperatures. These differences are most statistically significant and widespread when the number of moments is changed. To determine whether these results are due to (fast) local effects of the different microphysics or the (slower) ensuing feedback on the large-scale atmospheric circulation, we run a series of short, 5-day simulations initialized from reanalysis data. We find that the differences largely disappear in these runs and therefore infer that the different parameterizations impact precipitation extremes indirectly via the large-scale circulation. Finally, we compare the present-day results with hourly rain-gauge data and find that SP underestimates extremes relative to observations regardless of which microphysics scheme is used given a fixed model configuration and resolution.

The content of this chapter is published as Charn et al. (2020).

2.1   Introduction

Impacts from climate extremes such as heat waves, droughts, and heavy storms include geographic range shifts of plants and animals, food price increases, damage to infrastructure, and increased mortality (IPCC, 2014). In particular, rainfall extremes directly impact the frequency and the severity of floods and the runoff in rivers, which provide over half of the global potable water supply (Barnett et al., 2005). Many global climate models (GCMs) are able to successfully reproduce mean patterns on large scales, but they generally exhibit lower fidelity in simulating regional features and higher-order statistics (variance, extremes, etc.) (Flato et al., 2013). One reason is that critical convective processes that generate clouds and precipitation must be parameterized since the characteristic length scales of these processes are much smaller than a conventional GCM grid cell. Typical GCMs often diagnose convection by assuming a quasi-steady equilibrium (e.g., Zhang and McFarlane, 1995), an approximation that is notorious for underestimating higher-intensity updrafts and hence, extreme rainfall (Dai, 2006; Wilcox and Donner, 2007). While the employment of a global cloud-resolving model with horizontal resolution of O(1 km) provides an obvious solution, the disadvantage of this approach is the increased computational cost by a factor of a million.

One solution to this problem that balances resolution and computational cost is the creation and utilization of a superparameterized (SP) model such as the SP version of the Community Atmosphere Model (CAM), known as SPCAM (Khairoutdinov and Randall, 2001). Superparameterization replaces the convective and boundary-layer schemes in each GCM grid column with a cloud-resolving model (CRM). Depending on its resolution, the CRM can resolve deep-convective and mesoscale processes in response to large-scale GCM dynamics, and in return it provides subgrid convective heating and moistening tendencies to the larger grid. Hence, cloud-scale interactions between cloud dynamics, microphysics, radiation, and turbulence are more finely resolved. Various studies have documented improved correlation with observations, for example, of the Madden-Julian Oscillation (Benedict and Randall, 2009), African easterly waves (McCrary et al., 2014), and the diurnal precipitation cycle (Khairoutdinov et al., 2005). In the case of rainfall extremes, Li et al. (2012) studied the tails of the frequency distribution over the United States and noted that they are much better represented when using SPCAM as opposed to CAM. Globally, Kooperman et al. (2016) found that SPCAM has lower systematic biases relative to satellite-derived products since it simulates higher extreme rain rates, whereas CAM underestimates the heaviest rain rates at coarser ($\sim 2°$) resolutions. SPCAM also has a higher, more realistic amount mode, that is, the rain rate that delivers the most accumulated rainfall, at all resolutions. Conversely, Kooperman et al. (2018) showed that the amount mode in CAM is underestimated from 50°S to 50°N due to the component from the convective parameterization, which dominates precipitation amounts in low-to-mid-latitudes.

While the structural uncertainties associated with the representation of deep convection have been examined by comparing simulations from CAM to those from SPCAM, less attention has been paid to the parametric uncertainty that arises due to the representation of

cloud microphysics in global models. In limited-domain CRM simulations, which are typically run with non-periodic boundary conditions forced by reanalyses, differing microphysics parameterizations have been shown to impact precipitation structures and statistics. For example, Bryan and Morrison (2012) showed in their case-study simulation of a squall line that using hail instead of graupel as the rimed-ice species led to faster fall velocities and slower propagation speeds and hence to increased precipitation accumulation. Van Weverberg et al. (2012) also simulated a squall line and found that using hail led to a higher domain-maximum accumulated precipitation but also faster propagation speeds due to more intense cold pools. However, the largest factor contributing to the domain-maximum accumulation in their study is the combined treatment of raindrop self-collection and collisional breakup. Specifically, the representation involves a size threshold that, when exceeded by the number-weighted drop diameter, causes the collection efficiency to decrease exponentially from 1 in order to simulate the effects of increased breakup. We conduct a similar test involving this process, which is described in greater detail in Section 2.2.1.1. Finally, a radiative-convective equilibrium experiment (Singh and O'Gorman, 2014) using doubly periodic boundary conditions yielded instantaneous rainfall rates that varied by an order of magnitude as hydrometeor fall speeds fixed to constant values were varied by an order of magnitude.

To the best of our knowledge, one study to date has examined the effects of different parameterizations of microphysics within the SPCAM framework. Elliott et al. (2016) investigated mesoscale convective systems (MCSs) and found that sensitivities in MCS event counts and in precipitation rates were overshadowed by interannual variability. Recently, Parishani et al. (2017) introduced ultraparameterized-CAM (UPCAM), in which the embedded CRMs have fine enough resolution to explicitly capture boundary-layer turbulence. While not focused on precipitation, Parishani et al. (2018) compared SPCAM and UPCAM and found that microphysics, rather than the CRM grid resolution, was mostly responsible for differences in cloud feedback to +4 K sea surface temperature (SST) warming. They attributed this finding to differences in liquid and ice water path in the baseline climatology.

In this study we generalize our analysis (as compared to Elliott et al. (2016)) to precipitation rates. We apply extreme value theory to see whether disparate microphysics representations within SPCAM significantly alter the distribution tail, and if so, which representation best matches observations. This question has relevance for the spread in current and future projections of rainfall extremes. As a secondary question, we ask whether any significant differences are due to local effects or feedbacks on the large-scale circulation. The former would encapsulate processes connecting water vapor supersaturation to hydrometeor fallout, while the latter would include radiative processes, latent heat release from condensation and freezing, and evaporative cooling, which influences cold-pool development.

Section 2.2 describes the datasets and the model used, including the microphysics parameterizations analyzed. Section 2.3 describes how extremes are defined and fit with an extreme-value distribution. Results are presented in section 2.4 and conclusions in section 2.5.

2.2   Data and Methods

2.2.1   SPCAM

We have employed SPCAM, where CAM is the atmospheric component of the Community
Earth System Model (Hurrell et al., 2013), forced by prescribed monthly SSTs and sea ice
boundary conditions (a run-time configuration known as the `F_2000` compset). The em-
bedded CRM is the System for Atmospheric Modeling (SAM) (Khairoutdinov and Randall,
2003). In our experiments the CRMs have been configured in a 2D, east-west orientation
with periodic boundary conditions with grids of 32 columns at 2-km horizontal resolution
and 30 vertical levels corresponding to those in CAM. The large-scale, outer grid is imple-
mented using a finite-volume dynamical core with a $1.9 \times 2.5°$ resolution. Greenhouse gas
concentrations are held fixed at year-2000 levels, and (bulk) aerosols are prescribed. Precip-
itation rates have been output every 3 hours, and to be clear, average rates (equivalent to
accumulation over the 3-hour time period), as opposed to instantaneous rates every 3 hours,
were output.

2.2.1.1   Microphysics Parameterizations

SAM is equipped with two fundamentally different microphysics formulations, specifically
the original, 1-moment scheme (1MOM in Table 2.1) described in Khairoutdinov and Ran-
dall (2003) and the 2-moment one (M2005 in Table 2.1) from Morrison et al. (2005). The
number of moments refers to the number of degrees of freedom in the size distribution of a
precipitating hydrometeor, which is prescribed to be an inverse exponential (Marshall and
Palmer, 1948). The 1-moment scheme prognoses the mixing ratio and diagnoses the number
concentration while the 2-moment scheme prognoses both. We note that using 1-moment
microphysics was only able to be done by using SPCAM4, while the experiments with 2-
moment microphysics were only able to be done using SPCAM5. While CAM5 (Neale et al.,
2010b) uses different boundary layer and convection parameterizations from those of CAM4
(Neale et al., 2010a), as mentioned before, the embedded CRM overrides all of them, so
most of the difference between the two versions of SPCAM lies in the microphysics and the
aerosol processes. 1n 1MOM there is no explicit droplet activation or ice nucleation; rather,
the cloud condensate is diagnosed by assuming zero supersaturation at all times. Cloud
liquid and cloud ice are then partitioned using a linear relationship that depends only on the
temperature. In contrast, M2005 explicitly calculates droplet activation and ice nucleation.
One caveat is that CAM4 uses the CAMRT radiation scheme (Collins et al., 2002) while
CAM5 uses RRTMG (Iacono et al., 2008), a potentially confounding factor in this analysis.

Two additional experiments have been performed by modifying M2005. In the first,
graupel is replaced by hail, which has a higher bulk density as well as a faster fall speed
(2MH300). In the second (2MG600), the threshold diameter in the raindrop breakup/self-
collection process implemented following Verlinde and Cotton (1993) has been increased from
the default of 300 to 600 μm.

Table 2.1: Summary of Microphysics Experiments Performed

Name	SAM Scheme	Rimed-Ice Species	Size Threshold in Raindrop Breakup Process (μm)
1MG	1MOM	Graupel	n/a
2MG300	M2005	Graupel	300
2MH300	M2005	Hail	300
2MG600	M2005	Graupel	600

Table 1 compares the four representations. By allowing an extra free parameter in a hydrometeor's size distribution, a 2-moment scheme would be expected to generate more realistic cloud dynamics and properties, such as liquid and ice water contents and radiative fluxes, and indeed it has been shown to do so in several studies (Bryan and Morrison, 2012; Igel et al., 2015). Prognosing both the mixing ratio and the number concentration enables both quantities to evolve independently, thereby admitting size sorting to occur. For instance, a large mixing ratio and low number concentration implies larger particles, which fall faster and are less prone to evaporation than smaller ones. However, excessive sorting has been shown to occur when using 2-moment schemes, and there are various suggestions on how to ameliorate the problem, such as a 3-moment scheme in which the third moment is radar reflectivity (Milbrandt and Yau, 2005).

The 2MH300 and the 2MG600 representations might both be expected to increase rainfall extremes. By falling faster, hail might have less time to melt into raindrops susceptible to evaporation, thus increasing surface precipitation. This process is complicated, however, by the fact that differing melting and evaporation rates can impact cold pool development and propagation and thus affect the amount of time a storm spends over a particular location. Similarly, the larger raindrops allowed in the 2MG600 experiment should lead to faster fallout that is less impeded by breakup or re-evaporation.

The 2MG300 case is considered the baseline, and comparisons will be made between it and the other three in turn. All figures plotting a difference in return values between two microphysics schemes will have that of 2MG300 as the minuend (quantity being subtracted from) and that of the other case as the subtrahend (quantity being subtracted). We omit all maps/comparisons containing fewer than five grid cells showing statistically significant differences, as determined using the method described in Section 2.3.3 below.

2.2.2 Experiment Design

2.2.2.1 Climatological Runs

The 7-year climatological runs conducted for this analysis have start dates on 1 June 1991 and end dates on 30 November 1998. The first 6 months have been discarded to remove spinup effects. Two sets of experiments have been performed, one with present-day, annually cyclic, climatological SSTs, and the other with the same SST field uniformly increased by 4 K.

2.2.2.2 ILIAD

To determine whether the results seen in the (present-day) climatological runs are due to differing immediate, local effects or the microphysics' feedback on the large-scale circulation, we make use of the InitiaLized-ensemble, Analyze, and Develop (ILIAD) framework introduced by O'Brien et al. (2016). This software framework is designed to assess the effects of different parameterizations or model resolutions by conducting repeated short hindcast simulations initialized with reanalysis output and comparing the results against the meteorological conditions observed during the hindcast period. The experimental protocol, which we also use here, consists of one 5-day hindcast initialized at 00Z every day for 5 years. Precipitation output is taken from the fifth simulation day, a timeframe that allows the model to develop a somewhat distinctive dynamical state yet still be constrained to that observed on the verification date. The hindcasts are initialized by Climate Forecast System (CFS) (Saha et al., 2010) output taken from 1 December 2007 to 30 November 2010 and from 1 December 2011 to 30 November 2013, with the missing year reflecting our desire to conduct seasonal analyses amidst the fact that CFS version 2 (Saha et al., 2014) was made operational in March 2011. The CFS data was regridded to the large-scale SPCAM grid and include cloud water but no cloud ice or aerosol outputs. It should also be noted that the CFS data has zero land ice, in contrast to the monthly climatological data used in the free-running experiments. This is a potentially confounding factor but is partially mitigated by the fact that the precipitation probability density functions (PDFs) from the ILIAD runs lie within the interannual spread of those from the free-running ones (not shown).

2.2.3 Observations

The Climate Prediction Center (CPC) Hourly Precipitation Dataset (HPD; Higgins et al., 1996) is used to assess the fidelity of the model output. The HPD is derived from rain gauges and gridded using a modified Cressman scheme to cover a region spanning 20N-60N and 220E-297.5E. One third of the stations are first-order National Weather service stations, and the remaining two thirds consist of data from cooperative observers. The seven years of observations that are contemporaneous with the simulation period during 1991–1998 have been interpolated to the CAM grid.

2.3 Identifying and Comparing Extremes

2.3.1 Background on Extreme Value Distributions

To compare precipitation extremes, extreme value distributions were fitted to 3-hourly data for each season and in each grid cell within the contiguous United States (CONUS). There are two ways in which extremes from a time series are selected and modeled. The first is to divide the time series into regular intervals, say, months, then take the maximum of each period. Building on work done by Fisher and Tippett (1928), Von Mises (1954) and Jenkinson (1955)

independently showed that the generalized extreme value (GEV) distribution asymptotically
approximates the distribution of these block maxima. The primary drawback of this method
is that it ignores other extremes that might be present, for example, the second highest rain
rate in a block. An alternative approach to fitting extremes is the threshold exceedances
method, in which a distribution is fitted to all values over a chosen threshold. Pickands III
et al. (1975) showed that such a distribution can be asymptotically approximated by the
generalized Pareto distribution (GPD).

The GPD has a slight disadvantage in that the parameter estimates are a direct function
of the threshold u; for example, analytically finding the GPD representation of the standard
Fréchet model $F(x) = \exp(-1/x)$ yields a scale parameter $\sigma = u$. Thus, we employ a third
approach: the non-homogeneous Poisson point (NHPP) process (Pickands III, 1971), which
is also referred to as peaks-over-threshold. Like the threshold excess approach, the NHPP
method models all extreme values greater than a specified threshold but characterizes the
extreme value distribution using statistical parameters that correspond to the GEV approach.
An NHPP process is defined by an intensity *function* with the identical analytic expression
as that for the GEV PDF, namely

$$\lambda(t, y) = \frac{1}{\sigma}\left[1 + \xi\left(\frac{y - \mu}{\sigma}\right)\right]^{-1/\xi - 1}, \tag{2.1}$$

where μ is a location parameter, σ a scale parameter, and ξ a shape parameter. This is not
a coincidence, as it can be shown that both the GPD and the GEV are special cases of the
point process representation. In fact, return values for the NHPP process are calculated here
in the same manner as for the GEV distribution: the N-block return value is that which is
expected to be exceeded once every N blocks, where blocks here are years. However, λ is
not itself a PDF but rather the mean number of exceedances over a given rain rate in one
time block. This is analogous to a time rate r for a Poisson distribution, where $\lambda_{\text{Poisson}} = rt$
is the expected number of events in an interval t.

Thus, integrating $\int_0^T \int_u^\infty \lambda(t, y) dt dy$, where T is the total number of time blocks, yields the
intensity *measure*

$$\Lambda(T, u) = T\left[1 + \xi\left(\frac{u - \mu}{\sigma}\right)\right]^{-1/\xi}, \tag{2.2}$$

which is the expected number of exceedances of the rain rate over the threshold u over T
years.

Given a set of observed extremes $Y_1, ..., Y_N$ at times $t_1, ..., t_N$, the negative log-likelihood
of an NHPP process is

$$-\ln L = N \ln \sigma + \left(\frac{1}{\xi} + 1\right)\sum_{k=1}^{N} \ln\left[1 + \xi\left(\frac{Y_k - \mu}{\sigma}\right)\right] + \Lambda(T, u), \tag{2.3}$$

where L is the likelihood. We use optimization techniques to minimize Equation 2.3 to find
the parameters μ, σ, ξ.

2.3.2 Likelihood Ratio Test

To compare two microphysics schemes, we used the likelihood ratio test to determine whether the fitted NHPP processes are significantly different from each other. Such a test is generally framed with hypotheses

$$H_0 : \theta \in \Theta_0$$
$$H_a : \theta \in \Theta \backslash \Theta_0$$

with θ the distribution parameter(s) and Θ_0 a subset of the parameter space Θ.

Wilks (1938) proved that, given the observed data x and assuming H_0 is true, the test statistic

$$z = -2 \ln \frac{\sup\{L(\theta|x) : \theta \in \Theta_0\}}{\sup\{L(\theta|x) : \theta \in \Theta\}}$$

asymptotes to a chi-square distribution with k degrees of freedom, where k is the difference in dimensionality, or number of parameters, between Θ and Θ_0. Thus, a p-value can easily be computed.

2.3.3 Comparison Procedure

A null and an alternative hypothesis can be defined for this study as follows:

$$H_0 : \theta_i = \theta_j$$
$$H_a : \theta_i \neq \theta_j$$

where θ_i denotes the parameters for the NHPP fit of microphysics case i and θ_j those of case j. In other words, the null hypothesis states that the two cases being compared can be reasonably represented by one common distribution, while the alternative hypothesis says that the two individual fits are significantly different, that is, the extremes arise from distributions that exhibit meaningful differences. Our test statistic then can be written as

$$z = -2 \ln \frac{\sup\{L(\theta_{12}|x_1, x_2)\}}{\sup\{L(\theta_1|x_1)\} \sup\{L(\theta_2|x_2)\}}. \tag{2.4}$$

Because the test statistic requires that the parameter estimation in the numerator and the denominator be done using the same data, the former is found by analyzing the union of the extremes for each individual microphysics case, rather than extremes determined from the union of the whole datasets. This also necessitates a uniform threshold u. Therefore, when comparing two datasets in a given grid cell and a season, u was defined as the higher of the two individual 98th percentiles, and extremes for both cases were defined as values greater than u. The 98th percentile was chosen because it was found that higher thresholds, for example, the 99th percentile, led to too many failures of the fitting algorithm. The results also depend on the percentile chosen (Charn et al., 2020, Figures S1-S3), particularly when comparing

the 1-moment and 2-moment schemes. Thus, we chose the highest integer percentile possible that also mitigated the incidence of convergence failures. This also ensures that the results shown are the most robust. Lastly, to account for temporal clustering, for example, a storm with high rain rates over multiple 3-hour periods, we took only the maximum value in each series of consecutive exceedances.

Once the three sets of extremes were acquired, a p-value was calculated using the test statistic with $k = 3$ degrees of freedom since there are six free parameters in the alternative model and only three in the null, where the parameters are the same between the two cases being compared. However, because a p-value is the probability of obtaining a test statistic at least as extreme as that observed while assuming H_0 is true, carrying out N tests with a significance level α will, on average, result in αN rejections if the tests are independent , and $> \alpha N$ if there is spatial correlation, simply due to random sampling error. Thus, as noted by Wilks (2016), the effects of multiple-hypothesis testing (conducting individual tests at multiple grid points) must be taken into account. We applied the method pioneered by Benjamini and Hochberg (1995) known as controlling the false discovery rate (FDR), which is the expected fraction of rejected local null hypotheses that are actually true. The procedure involves choosing α_{FDR}, the level at which it is desired to limit the fraction of false negatives (Type I errors). In this paper we set $\alpha_{\mathrm{FDR}} = 0.05$. In practice, this means that there is then a value p_{FDR} such that an individual p-value p must satisfy $p \leq p_{\mathrm{FDR}} \leq \alpha_{\mathrm{FDR}}$ in order to be declared significant to ensure that overall the rate of false discoveries does not exceed the nominal level. This method is robust, and even conservative, when there is strong spatial correlation since the achieved FDR will be smaller than α_{FDR}. For more details, see Appendix A.

2.4 Results

2.4.1 Present Day

2.4.1.1 Climatological Runs

Figure 2.1 shows the map of 2-year return values for 2MG300. Generally, across all seasons, extremes are lowest near the Rocky Mountains and highest in and between the Midwest and the East Coast, with moderate values on the West Coast except in JJA. When comparing the 1- and 2-moment schemes in the present day (Figure 2.2), there are numerous significant differences over the full annual cycle in the Midwest and during the summer season in the Southwest and Southeast. In the case of the southwestern United States, the location and the timing might imply interactions with the North American monsoon. Previous work has found that aspects of microphysical processes, for example, a more maritime-like, drop-size distribution and greater liquid and ice mixing ratios, may be important in producing heavier rainfall in organized convection during the North American monsoon (Rowe et al., 2012). We caution, however, that much of this summertime signal may actually be noise from interannual variability (see Section 2.4.1.4). Note also that in this region the diverse

distributions have differences in return values that are quite small in magnitude. While this is likely in large part due to the lower rain rates in the region (as compared to the rest of CONUS), it should be noted that disparate distributions can yield identical return values since there are three parameters that describe an NHPP process.

When comparing the two rimed-ice species in the 2-moment scheme (Figure 2.3), the Southwest again displays significant differences, though to a lesser degree than in Figure 2.2. Graupel leads to higher extremes in this area in the spring, while the opposite is true in the summer. The fact that there are so few significant differences in Figure 2.3 is perhaps a surprising result given other studies (Bryan and Morrison, 2012; Van Weverberg et al., 2012) that have demonstrated impacts on cold-pool dynamics and subsequent accumulated precipitation. Finally, increasing the size threshold before increased raindrop breakup intensifies leads to higher 2-year return values in the Midwest spring (Figure 2.4). Of the present-day findings, this is the most robust when varying the percentile used to demarcate extremes (Charn et al., 2020, Figure S3).

The robust signal in the Midwest and the seasonality (occurring mostly in the spring and the summer but never in the winter) points to the choice of microphysics being more important in places and times where storms are dominated by convective processes as opposed to large-scale air motions. This is a similar conclusion reached by Li et al. (2012), who found that CAM and SPCAM yielded similar precipitation extremes in the western United States and during winter but vastly different results in the southeastern United States and during summer. The springtime signal in the Midwest, particularly in Figure 2.4, could suggest that SPCAM is impacting the transition from predominantly synoptic to convective precipitation, though we do not investigate this further.

2.4.1.2   5-Day Runs

The central and western United States display significant differences in annual extremes when comparing the 1- and 2-moment schemes in the ILIAD framework (Figure 2.5), though fewer than in the climatological case. These regional disparities are not collocated with those in Figure 2.2, implying that local microphysical effects can cause differences in SPCAM precipitation extremes when varying the number of moments prognosed. Nevertheless, this is the only case where there are significant differences; none occur in the seasonal comparisons with 1MG and 2MG300 nor in any of the comparisons between the 2-moment schemes. Given the general lack of stippling, a conclusion consistent with this evidence is that the microphysics parameterizations alone (i.e., operating under the same meteorological conditions) do not often lead to significantly different extremes in SPCAM simulations within CONUS. Instead, an alternate plausible hypothesis is that precipitation rates are indirectly influenced via feedbacks from the microphysics to the large-scale circulation that operate on longer timescales. These ILIAD runs are more similar to the aforementioned CRM studies forced by reanalyses at the boundaries (e.g., Bryan and Morrison, 2012; Van Weverberg et al., 2012), yet our findings, that is, that the local effect of microphysics do not appear to be significant, seem to differ with those in the literature.
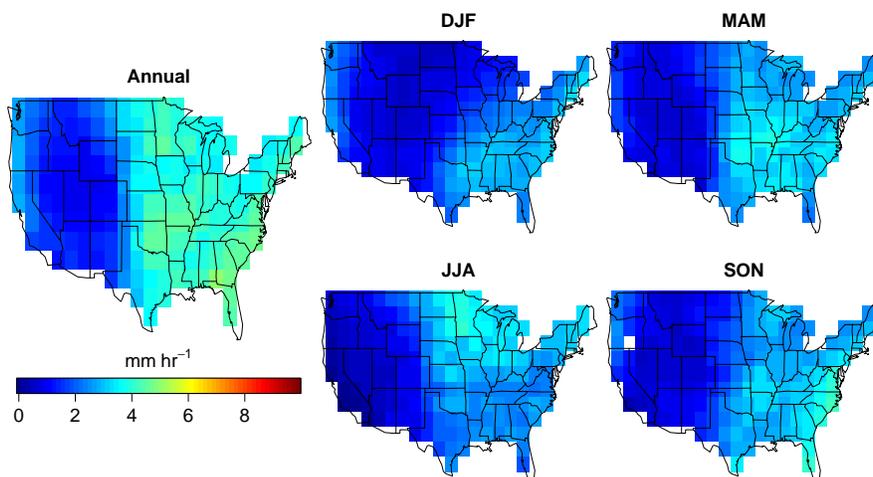
Figure 2.1: A 2-year return value for 2MG300 (Table 2.1) in the present-day climatological experiment. The legend axis matches that in Figure 2.10 for ease of comparison with observations.
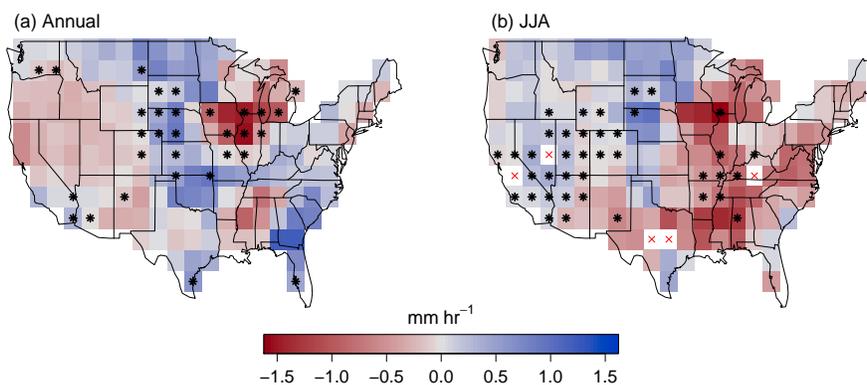
Figure 2.2: Climatological comparison of 1MG (with 2MG300) (Table 2.1). Following the convention defined in Section 2.2.1.1, the quantity plotted is the 2MG300 2-year return value minus that of 1MG. Asterisks denote grid cells with statistically significant differences between the two extreme-precipitation distributions. Grid cells where parameter estimation failed are denoted by a red X.

There are several reasons as to why the sensitivities are diminished in SPCAM. One is
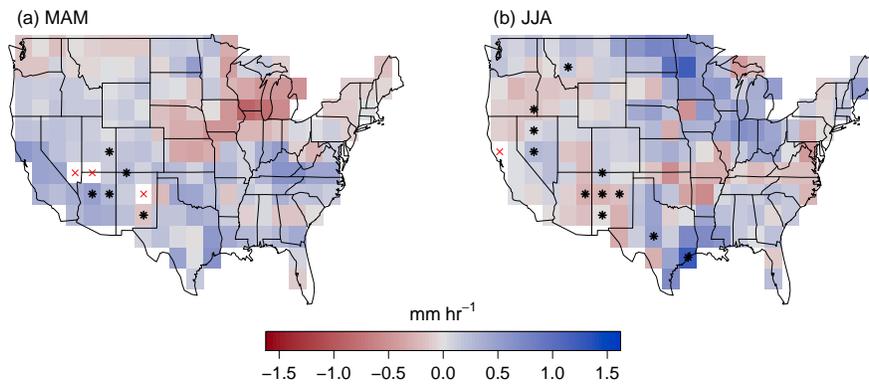
Figure 2.3: As in Figure 2.2 but with climatological comparison of 2MH300 and 2MG300 (Table 2.1).
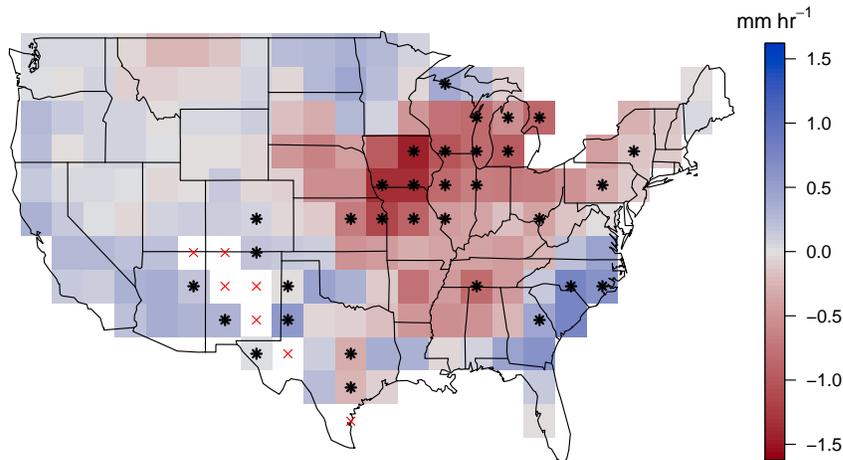
Figure 2.4: As in Figure 2.2 but with climatological comparison of 2MG600 and 2MG300 (Table 2.1) in MAM.

our use of a 2D domain in the embedded CRM. Various CRM studies (Wilhelmson, 1974; Phillips and Donner, 2006; Zeng et al., 2008) have noted that updrafts, especially the most intense ones associated with deep convection, have faster speeds in 3D simulations. Phillips and Donner (2006) and Zeng et al. (2008) found that this led to higher graupel mixing

ratios, and more graupel could potentially make the distinction between graupel and hail more important. Cloud dynamics would also affect the balance between warm-rain and cold-rain production because more moderate updrafts would favor increased coalescence of cloud water while faster vertical velocities lead to higher penetration, resulting in more ice (e.g., Redelsperger et al., 2000). A second possible cause for the reduced sensitivities is the fact that the data were averaged over the CRM domain within the GCM column, thus perhaps making a comparison to domain-mean (as opposed to domain-maximum) precipitation rates from CRM studies more appropriate. A third confounding factor is the limited vertical resolution (30 levels), which is restricted to match that of CAM. Khairoutdinov et al. (2009) found a dependence of cloud distribution and convection on vertical resolution, so it is possible that future SP models with an increased number of vertical levels could show more realism and/or sensitivity to experiments.

The findings here are robust regardless of the percentile used to denote extreme values (Charn et al., 2020, Figures S4-S6), lending support to the idea put forth by Khairoutdinov and Randall (2003), namely that nonlocal feedbacks can amplify microphysical sensitivities that would otherwise be inoperative or suppressed due to prescription of meteorological boundary conditions in limited-domain CRM simulations. It is also interesting to compare our finding with that of Li et al. (2012), who found significant disparities between CAM and SPCAM even when the model's large-scale fields were replaced with NCEP reanalysis data at each time step. Here, the model is simply initialized with reanalysis fields for short, 5-day runs, yet little difference between the sundry microphysics schemes is observed.
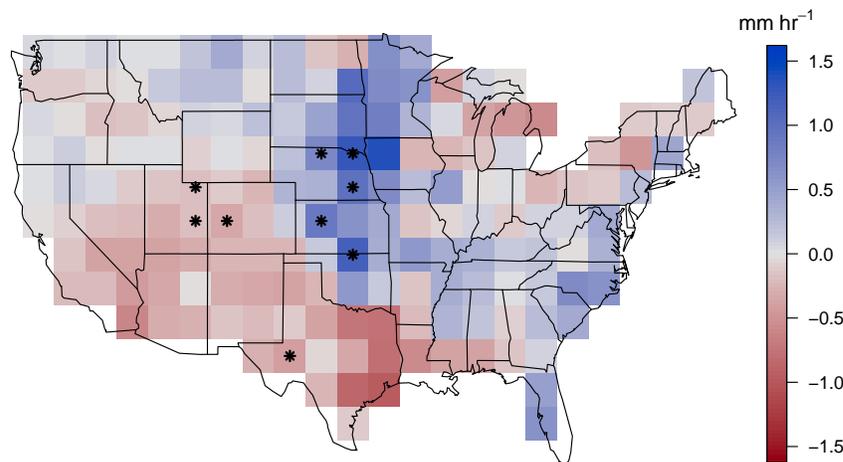
Figure 2.5: As in Figure 2.2 but with comparison of 1MG and 2MG300 (Table 2.1) annual extremes within the ILIAD framework.

2.4.1.3   Evidence of Large-Scale Circulation Changes

The results in Sections 2.4.1.1 and 2.4.1.2 support seeking evidence of large-scale circulation changes that could lead to the hypothesized feedbacks on extreme precipitation. Feng et al. (2018) compared two microphysical schemes within a convection-permitting model and found that one parameterization simulated more stratiform rainfall within MCSs, resulting in heating profiles that were top-heavier and had a sharper heating maximum. This in turn led to the strengthening of midlevel mesoscale convective vortices brought on by increased potential vorticity generation in this region, helping deepen the synoptic-scale trough within which the MCSs were embedded. Dry air intrusion, followed by increased evaporative cooling, allowed the systems to last longer and affect the large-scale circulation. This change in synoptic flow described by Feng et al. (2018) is a somewhat local phenomenon, with differences in heating profiles ultimately feeding back on the same storm, with the caveat that MCSs travel O(1000 km) over their lifetime. More importantly, however, MCSs have a lifetime of about one day, so we would expect to see effects from this process in the ILIAD runs. Indeed, it may partially explain the results in Figure 2.5, especially since most of the signal is in the Great Plains, where MCSs are frequent and produce much of their rainfall (Feng et al., 2018).

However, as stated above, effects that manifest in the climatological, but not ILIAD, runs imply nonlocal feedbacks affecting extreme precipitation. Figure 2.6a compares 1MG and 2MG300 heating rates at their respective midtropospheric (400-800 hPa) peaks in each grid column, and Figure 2.6b shows differences in the zonal mean. We see that the largest differences occur from about 20S-20N. While we do not attempt to further connect heating rates with the large-scale circulation, we note that this is consistent with the idea of a nonlocal feedback. We also looked at changes in convective available potential energy but found it larger nearly everywhere in CONUS when using 1-moment as opposed to 2-moment microphysics (Charn et al., 2020, Figure S7).

The top row of Figure 2.7 reprints some of the climatological results from Section 2.4.1.1, while the bottom row shows the respective differences in daily mean pressure velocity at the 500-hPa level (ω_{500}). In the cells/regions with statistically significant differences, the response in precipitation extremes generally follows that expected from the differences in ω_{500} (which can be on the order of 30%), particularly between panels 2.7a and 2.7d and between 2.7c and 2.7f. There is less similarity when comparing 1MG and 2MG300 in the summer (between panels 2.7b and 2.7e), especially in the southwestern United States. Low humidity, or relatively modest "extremes" (leading to small return values, as seen in Figure 2.1), could serve as an explanation. Note that an exact correspondence between the two rows of Figure 2.7 is not expected for multiple reasons. The differing output frequencies of precipitation and ω_{500} is one problem, which leads to another: it is not possible to pull ω_{500} values from the same time steps as the rainfall extremes. Thus, the bottom row of Figure 2.7 shows differences averaged over the entirety of the corresponding season and therefore might be more telling of shifts in mean precipitation. Similarities between the top and bottom rows, then, hint at shifts in the precipitation-rate distribution that move both the
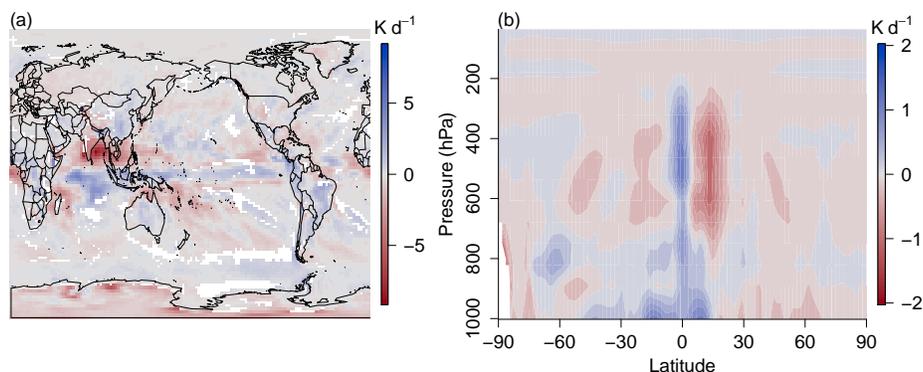
Figure 2.6: Climatological 2MG300 (Table 2.1) heating rates minus those of 1MG, averaged across times with columns with $\omega_{500\text{ hPa}} < 0$ hPa d^{-1}. (a) Difference in the midtropospheric (400-800 hPa) peaks in the climatological heating profile; only columns with peaks (between 400 and 800 hPa) less than 100 hPa apart are colored. (b) Zonal mean of difference in heating rates.

mean and extremes in the same direction, while disparities imply changes in the shape of the distribution. Lastly, as described further below, one would expect strong precipitation events to be proportional not only to the upward motion but to the moisture content (which was unfortunately not output) as well.

2.4.1.4 Estimating Natural Variability of SPCAM Return Values

It is possible that 7 years is too short an integration period, such that, despite the statistical lengths we have gone to try to ensure a robust signal, the differences we see are artificially significant, and attempts at interpretation are moot. To test this, we grouped the climatological output into 28 years (consisting of the four microphysics experiments of 7 years each). We then randomly reshuffled the 28 years and redid our analysis, with the idea that this gives a sense of the signal in a dataset where none is expected. We did this reordering 600 times, and the median number of CONUS cells flagged each season as statistically significant by our algorithm is listed in the last column of Table 2.2. Note that in calculating the median, for each reshuffling and season, the three values generated (from comparing the "four" microphysics schemes) are taken as individual data points (as opposed to being summed), so that a fair comparison can be made to the actual climatological results (the middle three columns in Table 2.2). We see that our statistical procedure generally reports no significant differences when reshuffling the data, except during JJA. This suggests that, at least for extreme precipitation, signal detection over interannual variability is more difficult in JJA, which is in line with the findings of Elliott et al. (2016) regarding summertime MCSs.
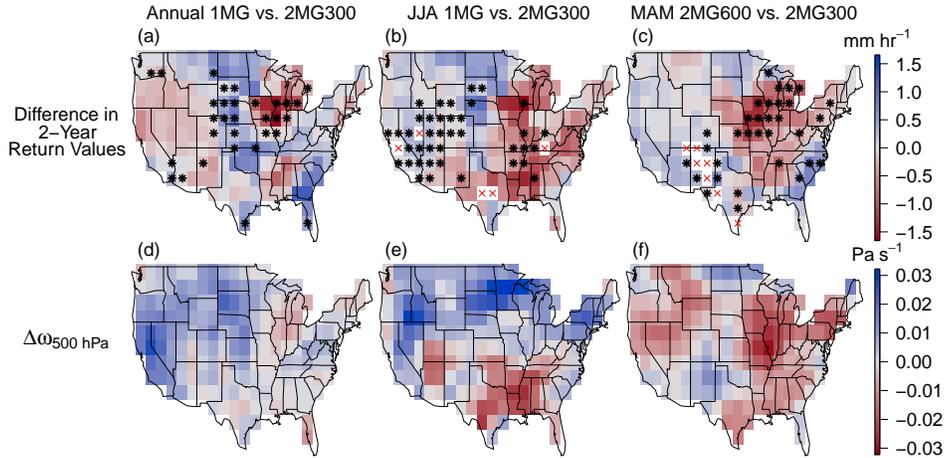
Figure 2.7: Climatological 2MG300 (Table 2.1) 2-year return value minus that of (a) 1MG in the annual, (b) 1MG in JJA, and (c) 2MG600 in MAM, with asterisks denoting grid cells with statistically significant differences. 2MG300 daily-mean ω_{500} subtracted from that of (d) 1MG in the annual, (e) 1MG in JJA, and (f) 2MG600 in MAM.

Table 2.2: Number of CONUS Cells with Statistically Significant Differences

Season	1MG[a]	2MH300[a]	2MG600[a]	Reshuffling Median
Annual	31	1	0	1
DJF	1	0	0	0
MAM	1	5	32	0
JJA	39	11	2	9
SON	0	3	0	0

[a]In climatological comparison with 2MG300.

2.4.1.5 Comparing with Observations

While there are significant differences in the climatological runs (Figures 2.2-2.4), the disparities in return values are small in comparison to those when measuring against observations. We show here only the comparison of the 1MG case to observations (Figure 2.8); the maps of the other three microphysics schemes (Charn et al., 2020, Figures S8-S10) look very similar. SPCAM systematically underestimates the magnitude of precipitation rates when compared to the CPC HPD data, particularly when looking at annual, summer, and fall extremes, with the closest (spatial) agreement in the winter. While this might seem to be at odds with the conclusions of Li et al. (2012), the conclusions are in fact quite similar; notably, the model underestimates in the southeastern United States while showing better agreement, or even overestimation, in the north, particularly in the 2MG600 case.

Figures 2.9 and 2.10 show the maps of return values for the 2MG300 ILIAD case and the CPC observations, respectively. One might expect that the ILIAD framework, with its short, 5-day runs initialized by reanalysis data, better captures the location, if not the magnitude, of precipitation events. Comparing with Figure 2.1, this appears to be the case at least in the summer, with the largest extremes placed further south in the Midwest, closer to Iowa. Thus, as in Li et al. (2012), it seems that winter precipitation events dominated by synoptic systems are better represented in terms of both magnitude and location, whereas summer events dominated by convective processes are potentially subject to greater errors, even in SPCAM.

Several studies (Rauscher et al., 2016; O'Brien et al., 2016; Fildier et al., 2018) have shown empirical evidence for the idea that, for strong events, precipitation rates are approximately equal to the cloud-base upward moisture flux, which is the product of vertical velocity (in pressure coordinates) and specific humidity. As mentioned above, 3D CRM simulations generally exhibit larger updraft speeds, particularly at the highest percentiles: Phillips and Donner (2006) found that the most vigorous 1% of updrafts were 20-40% faster, and Wilhelmson (1974) found that the updraft in a 3D thunderstorm simulation was twice as fast. Indeed, the latter study found that the maximum rainfall intensity in 3D was about three times that of the 2D thunderstorm. Thus, it is possible that embedding a 3D CRM domain (instead of a 2D one) within each GCM column would result in a closer match to observations. Increasing the horizontal resolution of the CRM might also help: Ooyama (2001) showed in 2D simulations of a single-cell cloud that 1-km resolution resulted in higher rainfall intensities than when using 2 km, which in turn yielded a peak precipitation rate twice that in a 4-km experiment. Nevertheless, this effect is not a surety, especially in a 3D setup: Bryan and Morrison (2012) found that domain accumulated precipitation–which, as noted above, might be the more relevant quantity in SPCAM–was relatively similar between a 4- and a 1-km simulation of a squall line. And in fact, precipitation *decreased* with 0.25-km resolution due to increased entrainment which weakened the system intensity.

2.4.2 Uniformly Increased SSTs

Finally, we analyzed identical free-running simulations but with SSTs uniformly increased by 4 K. Again, there are substantial regions with differences when comparing the 1- and the 2-moment schemes over the whole year as well as summertime (Figure 2.11). In this case, however, the stippling is not confined to the southwestern United States in the summer; parts of the Midwest show significant differences as well, with 1MG showing more extreme precipitation rates further east, 2MG300 further west. As when using present-day SSTs, there is no detectable signal when comparing 2MG300 and 2MH300. The springtime signal when comparing 2MG300 and 2MG600 (Figure 2.4) also disappears.

We also investigated changes in return values when going from the present climate to the +4 K SST climate. We would generally expect the most intense precipitation events to increase in magnitude with warming, but does the response depend on the microphysics scheme used within SPCAM? Figure 2.12 suggests the answer is no for annual extremes
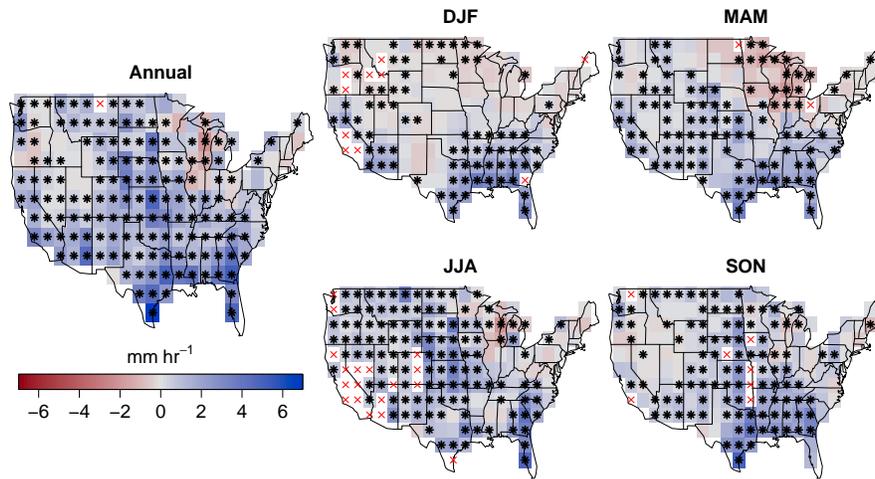
Figure 2.8: As in Figure 2.2, but the 1MG (Table 2.1) 2-year return value is subtracted from that of the CPC HPD observations.
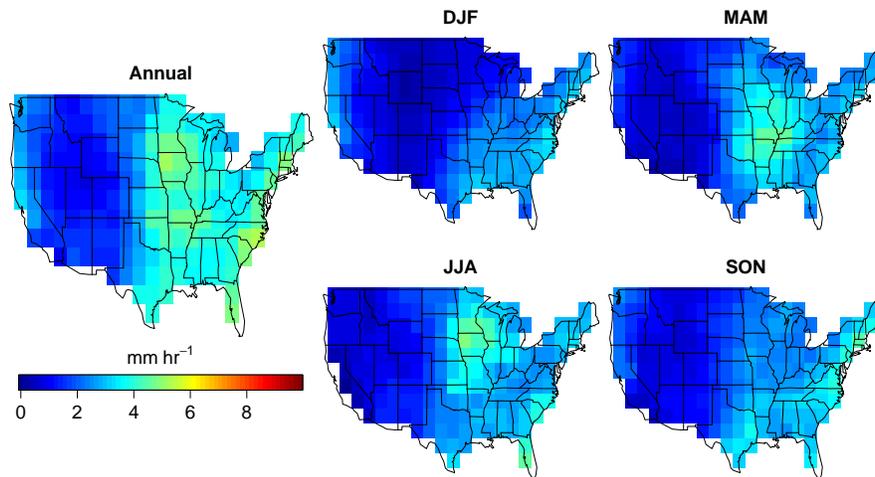
Figure 2.9: A 2-year return value for 2MG300 (Table 2.1) in the ILIAD framework. The legend axis matches that in Figure 2.10 for ease of comparison with observations.

within CONUS: the return values significantly increase almost everywhere, but there is not much distinction between the four cases. There is some slight variation in the spatial
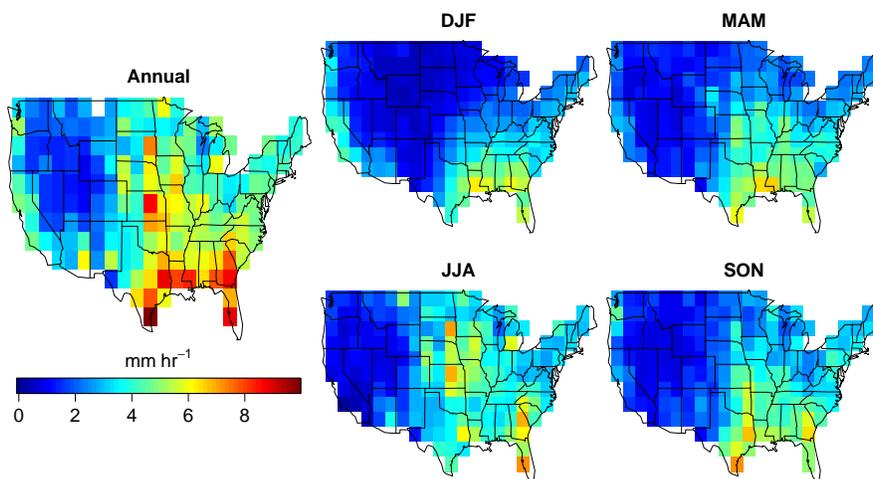
Figure 2.10: A 2-year return value for the CPC HPD observations.

structure, but generally the largest increases (in an absolute sense) are in the Midwest, with secondary increases on the West and East Coasts.

A potentially interesting and similar simulation would be a pseudo-global-warming experiment, as done by Pall et al. (2011), within the ILIAD framework. That is, first gather results from two (ensembles of) climatological (SP)CAM runs, one forced with observed historical time series of radiative and surface boundary conditions, the other with 4xCO2 (alternatively, the increased amount of CO2 necessary to produce an average of 4 K global warming) and the corresponding surface boundary conditions, as estimated by a fully coupled atmosphere-ocean model, for example, (SP)CESM. Take the difference in atmospheric variables between the future and present climates in, say, the month of the ILIAD simulation, and add it to the corresponding reanalysis fields used as initial conditions. This set of integrations would be like the "counterfactual" case in event-attribution studies, e.g, Pall et al. (2011, 2017), except that instead of being a scenario without already-occurred anthropogenic emissions, it represents a future world with continuing emissions. The local effect of changing microphysics schemes on return values in a warmer climate could then be evaluated in an ILIAD framework as in Section 2.4.1.2.

2.5   Conclusions

Earlier studies have found that superparameterization, which is a mechanism to introduce convection-permitting processes into GCMs, can increase the fidelity of the precipitation extremes simulated with these models. However, even as increasing numbers of climate
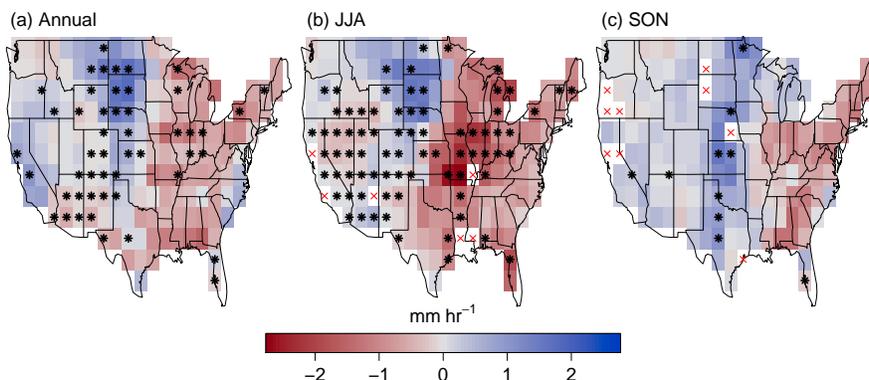
Figure 2.11: As in Figure 2.2 but with comparison of 1MG and 2MG300 (Table 2.1) in experiment with +4 K SSTs.

models transition to convection-permitting or -resolving atmospheric dynamics, microphysical processes will still require parameterization. Experiments with limited-domain CRMs have shown that rainfall statistics are sensitive to the choice of microphysical representations. However, it was previously unclear whether nonlocal feedbacks in a global model would amplify or diminish such a response.

In this study we compare precipitation extremes generated by SPCAM for various microphysical configurations. In the climatological simulations, we find significant differences in both present-day and warmer climates, particularly between the extreme statistics generated with the 1- and the 2-moment schemes during summer and the whole year. Short, reanalysis-initialized runs performed via the ILIAD framework suggest that this response is mostly due to effects on the large-scale circulation that then feed back into rainfall extremes. This conclusion is supported by comparing mean 500-hPa vertical velocity maps, whose differences we hypothesize stem from variations in vertical heating profiles caused by the disparate microphysics representations. Despite the disparities in rainfall statistics, we also find that none of the four analyzed microphysical parameterizations yield good agreement with rain-gauge data, as SPCAM consistently underestimates the observations. And in fact, if the intensity extremes are in turn underestimated in the observational data sets (both the one used here and others produced in the past), the discrepancies between SPCAM and the data would be further amplified. Recent work done by Risser et al. (2019) showed that current gridded precipitation products have diminished variability and extreme values due to the interpolation of a fractal field in precipitation. They argue that an improved product would first estimate the extreme-value distribution locally and only afterwards interpolate the climatology.

Our results are perhaps contrary to what might be expected based on the results of Elliott et al. (2016), namely that global feedbacks lead to internal variability of storms that dwarfs
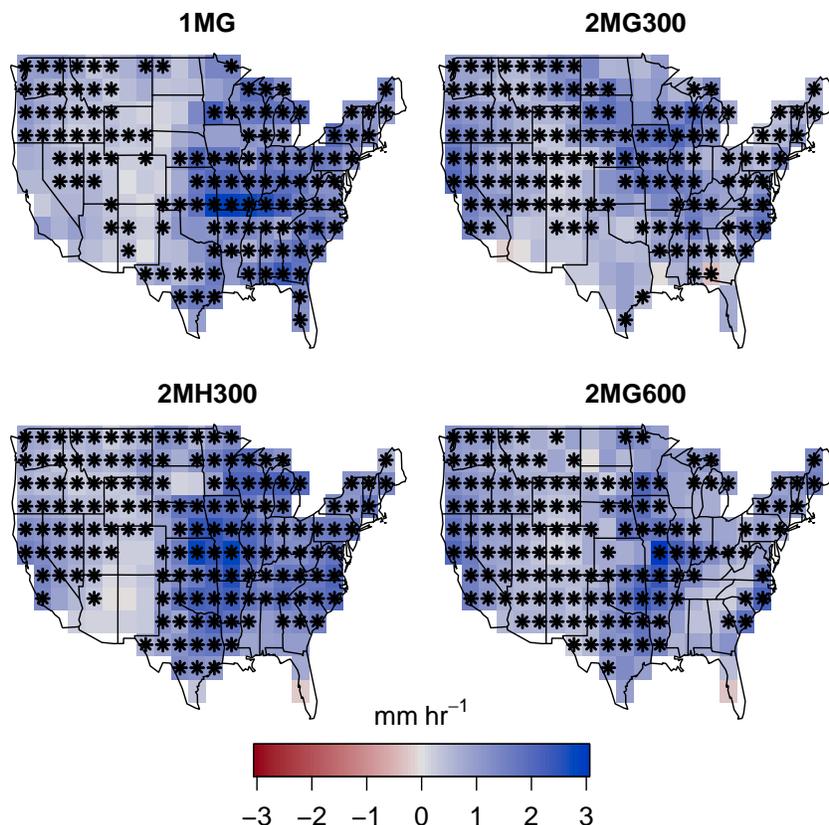
Figure 2.12: Present-day 2-year return value subtracted from that of the +4 K simulation. Asterisks denote grid cells with statistically significant differences between the two extreme-precipitation distributions.

sensitivity to microphysics configurations. Although, like Elliott et al. (2016), we found significant interannual noise in the summertime, we also found that nonlocal feedbacks magnify and highlight differences that might otherwise have gone undetected in shorter experiments. However, at least for the configuration and resolution used, none of the microphysical schemes tested are clearly preferable in terms of biases in return values of extreme precipitation. Helpful improvements in microphysics parameterizations might address both frozen and liquid processes. In the case of the former, better understanding of ice formation and evolution during aggregation and riming–which are poorly constrained processes and made even more difficult by the nonphysical partitioning of hydrometeors into discrete snow/ice/graupel/hail categories–is needed (Grabowski et al., 2019). For the latter, implementing stochasticity

into the collision-coalescence process, possibly using a Lagrangian, particle-based approach (Grabowski et al., 2019; Shima et al., 2009) might lead to a faster onset of rainfall due to some droplets growing faster than predicted in the usual continuous, deterministic formulation (Telford, 1955). In addition to broader work improving microphysics parameterizations, future work should involve seeing whether increasing the resolution, particularly that of the internal CRM, helps SPCAM align more closely with observations in its representation of rainfall extremes.

Chapter 3

Global Microphysical Sensitivity of Superparameterized Precipitation Extremes

A recent study found statistically significant differences in extreme precipitation distributions over the contiguous United States (CONUS) when changing the microphysics scheme in a superparameterized global climate model. Here, we repeat the analysis globally and similarly find that differences are widespread when varying the number of predicted moments in the microphysics parameterization but not when comparing variants of the double-moment scheme. However, contrary to the previous study in which differences largely disappeared over CONUS when 5-day simulations were conducted, we find that the signal in these shorter integrations remains within the tropics, implying a direct local effect of microphysics on precipitation extremes in these regions. The effect on precipitation is traced back to changes in vertical velocity profiles–changes that are then amplified in the climatological simulations compared to the 5-day ones.

The content of this chapter has been submitted to *Earth and Space Science.*

3.1 Introduction

Precipitation extremes are expected to increase with climate change (Held and Soden, 2006; Romps, 2011). Retrospective analyses of storms such as Hurricanes Harvey (Risser and Wehner, 2017), Katrina, Irma, and Maria (Patricola and Wehner, 2018) concluded that increased atmospheric temperatures likely led to more rainfall than would have been expected in a counterfactual world with no anthropogenic climate change. Despite their general skill at reproducing mean precipitation patterns on large spatial scales, global climate models (GCMs) still have difficulty representing extremes and even first-order phenomena such as the Intertropical Convergence Zone (ITCZ) (Flato et al., 2013). One reason is that the atmospheric convection that generates clouds and precipitation must be parameterized, as

opposed to explicitly simulated, since the characteristic length scales of these processes are much smaller than a conventional GCM grid cell. Typical GCMs often diagnose convection and its effects by assuming a quasi-steady equilibrium (e.g., Zhang and McFarlane, 1995), an approximation that is notorious for underestimating higher-intensity updrafts and hence, extreme rainfall (Dai, 2006; Wilcox and Donner, 2007). Convective parameterizations have also been heavily implicated in the so-called "double ITCZ" problem, where a spurious band of precipitation south of the Equator, comparable to its northern counterpart, appears in the annual mean (Hirota et al., 2011; Woelfle et al., 2018).

One alternative to the use of convective parameterizations that balances resolution and computational cost is the utilization of a superparameterized (SP) model such as the SP version of the Community Atmosphere Model (CAM), known as SPCAM (Khairoutdinov and Randall, 2001). Superparameterization replaces the moist physics parameterizations in each GCM grid column with a cloud-resolving model (CRM). Depending on its resolution, the CRM can resolve deep-convective and mesoscale processes in response to large-scale GCM dynamics, in return providing subgrid convective heating and moistening tendencies to the larger grid. Thus, cloud-scale interactions between cloud dynamics, microphysics, radiation, and turbulence are more finely resolved. Various studies have documented improved correlation with observations, for example, of the Madden-Julian Oscillation (Benedict and Randall, 2009), African easterly waves (McCrary et al., 2014), and global precipitation distributions (Kooperman et al., 2016). Li et al. (2012) also documented an improved representation of rainfall–but within the contiguous United States (CONUS) and with a focus on extremes– and with the caveat that superparameterized extremes were still unable to fully match the magnitude of those observed via rain gauges in the southern United States.

While the structural uncertainties associated with the representation of convection have been examined by comparing simulations from CAM to those from SPCAM, less attention has been paid to the parametric uncertainty that arises due to the representation of cloud microphysics in global models. See Charn et al. (2020) or Elliott et al. (2016) for a discussion of microphysical sensitivities within limited-domain CRMs. To the best of our knowledge, these two studies are also the only ones to have examined the effects of different parameterizations of microphysics within the SPCAM framework. Elliott et al. (2016) investigated summertime mesoscale convective systems (MCSs) within CONUS and found that sensitivities in MCS event counts and in precipitation rates were overshadowed by interannual variability. Charn et al. (2020) looked at extreme precipitation within CONUS and found significant differences, mostly when comparing 1- and 2-moment microphysics schemes, as a result of feedbacks onto the large-scale circulation.

In this study we generalize our analysis from that of Charn et al. (2020) to global precipitation rates. We again ask whether any significant differences are due to local effects or feedbacks on the large-scale circulation. While the latter was found to dominate the signal within CONUS, there is reason to expect the former's importance within the tropics. Fan et al. (2015) found vertical profiles of mass fluxes in their CRM to be insensitive to the choice of microphysics in two short-integration (24 hours) midlatitude case studies but strongly affected in a tropical case. They argued that the strong large-scale dynamic forcing

constrained the mass fluxes in the midlatitudes, but the upscale feedback of microphysics onto the convective organization in the tropical case was key in determining the flux profiles.

3.2 Data and Methods

3.2.1 SPCAM

The reader can refer to Charn et al. (2020) for details about the SPCAM configuration and the microphysics schemes within the embedded CRM, System for Atmospheric Modeling. The same four microphysics experiments were conducted (Table 2.1), and average precipitation rates were again output every 3 hours.

3.2.2 Experiment Design

3.2.2.1 Climatological Runs

The 7-year climatological runs used for this analysis are those conducted by Charn et al. (2020). The integrations have start dates on 1 June 1991 and end dates on 30 November 1998, with the first 6 months discarded to remove spinup effects. The annually cyclic, present-day climatological (1982-2001) SST and sea ice boundary conditions originally from Hurrell et al. (2008) are used.

3.2.2.2 ILIAD

To determine whether the results seen in the climatological runs are due to differing immediate, local effects or the microphysics' feedback on the large-scale circulation, we again make use of the InitiaLized-ensemble, Analyze, and Develop (ILIAD) framework (O'Brien et al., 2016). This software was created to assess the effects of different model parameterizations or resolutions by conducting repeated, short hindcast simulations initialized with reanalysis output and comparing the results against the meteorological conditions observed during the hindcast period. In this study, however, instead of using reanalysis outputs, each simulation is initialized with output from the 1MG and 2MG300 climatological runs. This will allow us to avoid the potentially confounding factor of Climate Forecast System reanalysis output (Saha et al., 2010) containing zero land ice, in contrast to the monthly climatological data used in the free-running experiments. It will also enable us to assess any sensitivity to the source of the initial conditions. The experimental protocol of O'Brien et al. (2016), which we also use here, consists of one 5-day integration initialized at 00Z every day for 5 years. Precipitation output is taken from the fifth simulation day, a timeframe that allows the model to develop a somewhat distinctive dynamical state, yet still be constrained to that of the initial condition.

3.2.3 Observations/Reanalysis

Because precipitation was saved in 3-hourly accumulations, the 20th Century Reanalysis, version 3 (NOAA-CIRES-DOE 20CRv3), which supplies this output as well, was used to provide a benchmark against the simulations (Compo et al., 2011; Slivinski et al., 2019). It is supported by the National Oceanic and Atmospheric Administration (NOAA), the Cooperative Institute for Research in Environmental Sciences (CIRES), and the U.S. Department of Energy (DOE). With data spanning back to 1804, all versions of 20CR to date have only assimilated surface pressure data. 20CRv3 uses the National Center for Environmental Prediction's Global Forecast System as its atmosphere-land model with a resolution of about 0.5°. Thus, it, like other models employing convective parameterizations, is possibly susceptible to underestimating heavy precipitation events (Li et al., 2012; Kooperman et al., 2016). For this reason we also compare our results from SPCAM with data from the Tropical Rainfall Measuring Mission (TRMM), specifically its Multi-satellite Precipitation Analysis product 3B42, version 7, hereafter TRMM 3B42 (Huffman et al., 2007). TRMM 3B42 precipitation primarily comes from two instruments, its TRMM Microwave Imager and its Precipitation Radar, and is available as a 0.25°, 3-hourly instantaneous product between 50°S and 50°N. Because of this latter point, a straightforward apples-to-apples comparison is not possible at the 3-hourly level, given that our SPCAM output is an accumulated quantity. Thus, when comparing against TRMM, we aggregate both our output and the observational data to daily resolution.

3.2.4 Comparison Procedure

To compare precipitation extremes, extreme value distributions (EVDs) were fitted to 3-hourly data for each grid cell. Specifically, we employed the non-homogeneous Poisson point (NHPP) process (Pickands III, 1971), also referred to as peaks-over-threshold, referring to the fact that it models all extreme values greater than a specified threshold, subject to temporal de-clustering (described below). Other possible EVDs that could have been used include the generalized extreme value and the generalized Pareto distributions; see Charn et al. (2020) for a discussion on how the NHPP was selected, as well as a mathematical description. We used the likelihood ratio test to determine whether the fitted NHPP processes are significantly different from each other. This involves first calculating the test statistic

$$z = -2 \ln \frac{\sup\{L(\theta_{12}|x_1, x_2)\}}{\sup\{L(\theta_1|x_1)\} \sup\{L(\theta_2|x_2)\}}, \tag{3.1}$$

where the numerator refers to the NHPP process fitted to the extremes from both microphysics cases "1" and "2", and the denominator refers to the processes fitted individually to the two cases. Wilks (1938) proved that a test statistic of this form asymptotes to a chi-square distribution with k degrees of freedom, where k is the difference in number of parameters between the null and alternative models. Here, $k = 3$ since there are three parameters in the null (corresponding to the three parameters in the NHPP process), i.e., both

distributions can be reasonably described by the same process, and six in the alternative, i.e., separate fits for the two distributions.

Because the test statistic requires that the parameter estimation in the numerator and the denominator be done using the same data, the former is found by analyzing the union of the extremes for each individual microphysics case, rather than extremes determined from the union of the whole datasets. This also necessitates a uniform threshold u. Therefore, when comparing two datasets in a given grid cell and a season, Charn et al. (2020) defined u as the higher of the two individual 98th percentiles, and extremes for both cases were defined as values greater than u. Here, when doing a global analysis, this led to some dry regions, mostly in northern Africa and the Arabian Peninsula, having a threshold of 0 mm/day. To remedy this, we define u as the higher 90th percentile of rain rates greater than 1 mm/day. Doing this leads to more failures in the optimization algorithm in the aforementioned regions (along with others, e.g., Antarctica), but it does not change the fundamental results of this paper (not shown). Lastly, to account for temporal clustering, for example, a storm with high rain rates over multiple 3-hour periods, we took only the maximum value in each series of consecutive exceedances.

Once the three sets of extremes are acquired, the test statistic and the corresponding p-value can be calculated. As in Charn et al. (2020), we control the false discovery rate (FDR) (Benjamini and Hochberg, 1995) with $\alpha_{\text{FDR}} = 0.05$ in order to account for the effects of multiple-hypothesis testing (conducting individual tests at multiple grid points) (Wilks, 2016).

3.3 Results

3.3.1 Climatological Runs

Figure 3.1 shows the map of 2-year return values for the climatological 2MG300 simulation. When comparing precipitation extremes (Figure 3.2), we see that the differences when looking at the number of predicted moments dwarf that when comparing variants of the 2-moment scheme. Altering the rimed ice species (Figure 3.2b) and the raindrop breakup parameter (Figure 3.2c) results in a handful of significant differences, mostly located between 30°S and 30°N and over the ocean. The limited amount of stippling culminates in a lack of much in terms of a spatially coherent signal. Thus, for the remainder of the paper, we do not delve into the 2MH300 and 2MG600 cases any further.

In contrast, a much larger fraction of grid cells are flagged when comparing 1MG and 2MG300 (Figure 3.2a). Again, much of the signal is between 30°S and 30°N and over the ocean. However, there is also an appreciable signal over land, including Europe, India, Africa, and South America, as well as at higher latitudes, for example, over the Southern Ocean. Broadly speaking, there is a narrow band of higher precipitation extremes in the 2MG300 experiment around the Equator, flanked by higher extremes in the 1MG run in the subtropical regions, though some exceptions to this trend occur over South America and
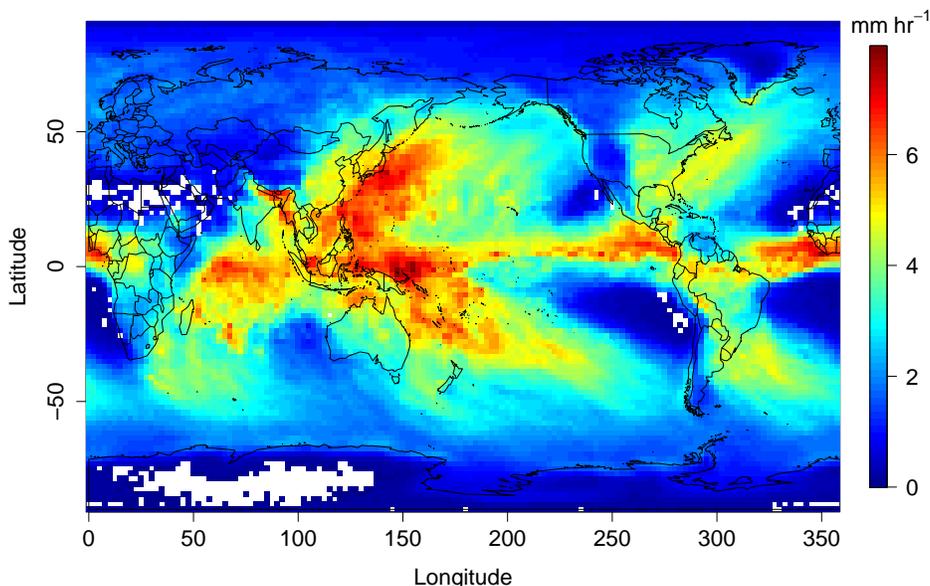
Figure 3.1: A 2-year climatological return value for 2MG300 (Table 2.1). Grid cells where parameter estimation failed are blank.

off the western coasts of Africa and Australia. This is suggestive of a strengthened Hadley circulation when using 2-moment microphysics, a topic we will revisit in the next section. There is even evidence of increased uplift over the Southern Ocean, at the southern edge of the Ferrell cell, though such a pattern is less obvious at the corresponding latitude range in the Northern Hemisphere.

Despite the vastly greater number of statistically significant differences when comparing 1MG and 2MG300, the differences in return values are generally comparable in magnitude within the three comparisons in Figure 3.2. This suggests that return values themselves are not a good indication of whether two extreme precipitation distributions are significantly different. As a comparison to a more traditional definition of extremes, Figure B.1a shows the map of annual 99.9th percentile 3-hourly precipitation rates for 2MG300. There is an extremely tight correlation between this map and that of the 2-year return values, with a correlation coefficient $r = 0.98$. Figure B.1b shows the 99.9th percentile of 2MG300 precipitation minus that of 1MG precipitation. Again, the spatial pattern is quite similar to that in Figure 3.2a, though the correlation with the difference in return values drops off to $r = 0.81$. This is because the threshold that demarcates an extreme, which ultimately influences the calculated return value, in each grid cell is defined as the minimum between that from 1MG and 2MG300, precluding a strict correlation with the straight 99.9th percentile.
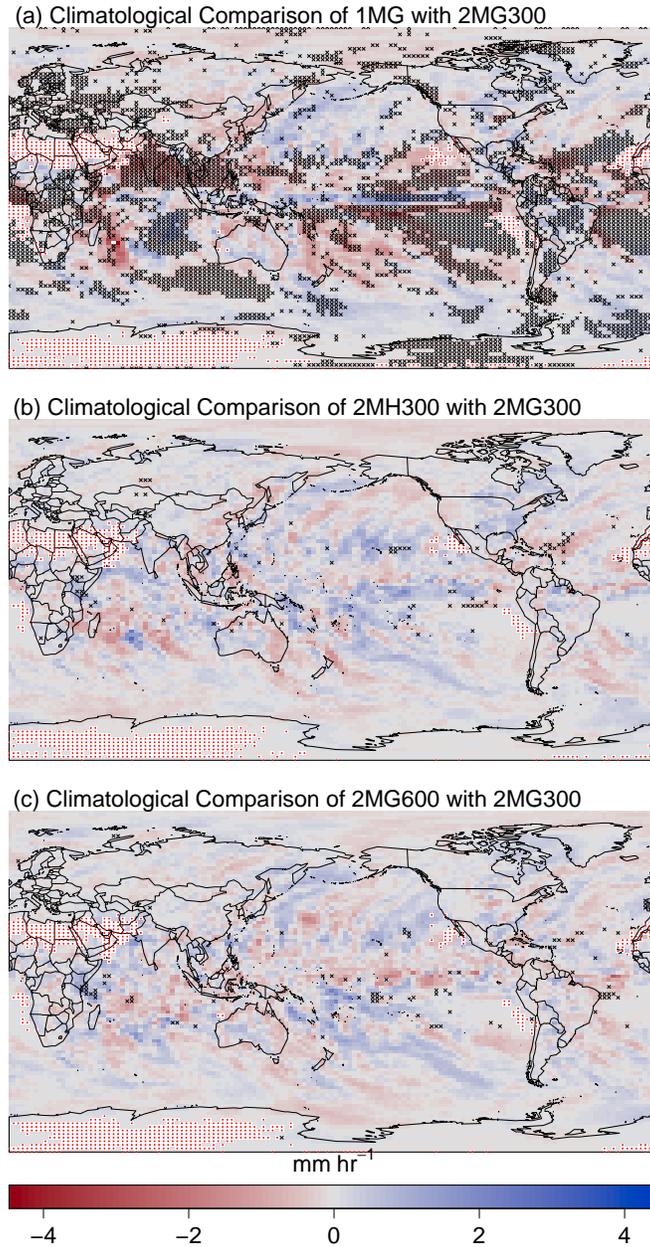
Figure 3.2: Climatological comparison of (a) 1MG, (b) 2MH300, and (c) 2MG600 with 2MG300 (Table 2.1). The quantity plotted is the 2MG300 2-year return value minus that from the other microphysics cases. Grid cells with statistically significant differences between the two extreme precipitation distributions are stippled in black. Grid cells where parameter estimation failed are stippled in red.

3.3.2 5-Day Runs

As discussed earlier, 5-day runs were conducted using the ILIAD framework to identify whether the differences in the climatological distributions of extreme precipitation (Figure 3.2) are due to local effects or feedbacks on the large-scale circulation. Because the differences between the 1- and 2-moment base cases vastly outnumbered those between the 2-moment variants, and due to limitations on computational resources, the ILIAD runs were only carried out with 1MG and 2MG300 microphysics. Four sets of simulations were performed: two–one using 1MG and the other using 2MG300–were branched off from the 1MG climatological run. Similarly, two other integrations were branched off from the 2MG300 climatological run. While the simulations using the same microphysics as the source (1MG→1MG, 2MG300→2MG300) are similar to the climatological runs themselves, they are not completely identical because the files used to initialize the 5-day runs contained a subset of the atmospheric fields used in the model's native restart files. Because the results do not depend on the source of the initial conditions, that is, which climatological run the ILIAD simulations were branched from (not shown), the remainder of the paper will be devoted to those derived from the 1MG large-scale meteorological conditions.

In the 5-day runs, unlike in Charn et al. (2020), a substantial signal with spatial coherence remains when comparing extreme precipitation between 1MG and 2MG300 (Figure 3.3). The signal is largely confined to latitudes between 30°S and 30°N. Much of the signal in the Southern Ocean as well as in the northern Pacific has disappeared, as has that in land regions such as Europe, the Indochinese Peninsula, northern India, Canada, and South America. Table 3.1 quantifies this: 24% of grid cells over land have significantly different precipitation extremes in the climatological run, while only 6% do in the 5-day runs. Also, the fraction of extratropical cells identified as significantly different drops to almost zero in the 5-day runs.

Table 3.1: Percentage of Cells with Statistically Significant Differences between 1MG and 2MG300

	Climatological			ILIAD		
	Land	Ocean	Total	Land	Ocean	Total
30°S-30°N	32%	47%	43%	7%	22%	18%
[a]$\|\phi\| >30°$	21%	17%	18%	5%	1%	2%
Total	24%	28%	26%	6%	8%	7%

[a]ϕ denotes latitude.

As mentioned in Section 3.3.1, the opposing signals in the ascending and descending regimes of the Hadley cell imply a strengthening of the circulation when using the 2-moment microphysics scheme, which could impact precipitation. Tao and Chern (2017) discussed a similar phenomenon within their superparameterized model (the Goddard Multiscale Modeling Framework) but due to changes in CRM domain size instead. Specifically, a smaller domain with coarser resolution in their embedded CRMs led to a strengthened Hadley cell,
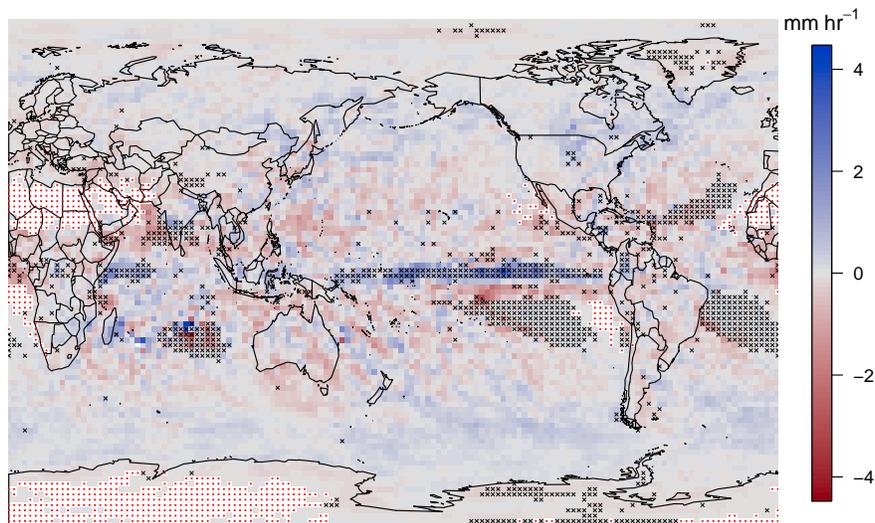
Figure 3.3: As in Figure 3.2 but with comparison of 1MG and 2MG300 (Table 2.1) extremes within the ILIAD framework.

which they argued enabled enhanced evaporation in the subtropics, in turn resulting in greater low-level water vapor flux convergence into the tropics. However, given the fact that the Hadley turnover time is on the order of a month, a time period of 5 days seems insufficient to allow either increased uplift in the deep tropics or increased subsidence in the subtropical regions to impact the other. Unfortunately, output from the first four days of each 5-day run was not saved to test this, so we do not explore this further.

3.3.3 Vertical Velocity

Figure 3.4 shows the zonal mean of 500-hPa vertical velocity ($\omega_{500\ hPa}$) from the 1MG case minus that from the 2MG300 case for both the climatological and ILIAD simulations. In the ILIAD runs, the difference in vertical velocity is most appreciable between 30°S and 30°N, with smaller magnitudes at higher latitudes. The 2MG300 microphysics generally leads to enhanced upward motion in the deep tropics around the Equator and enhanced downwelling in the subtropics. In the climatological simulations, where the large-scale circulation has the time to respond in the midlatitudes, differences in vertical velocity are intensified, particularly over the ocean (Figure B.2). 1MG microphysics leads to enhanced upward motion around 50°S in the southern parts of the Indian, Pacific, and Atlantic. Around 60°S, the reverse signal is amplified, with 2MG300 displaying greater upward motion throughout the Southern Ocean. The zonal-mean difference in $\omega_{500\ hPa}$ is weaker in the northern midlatitudes and displays more sign reversals within both the climatological and the ILIAD cases.

There is even a sign reversal between climatological and ILIAD at 50°N: slightly higher upward motion with 2MG300 in the latter is overtaken by higher updrafts in 1MG in Russia, Mongolia, and Europe. Figure 3.4 lends support to the above results–and those in Charn et al. (2020)–in that changes in vertical velocity and precipitation extremes are appreciable in the tropics in short, 5-day runs but only significantly appear in the midlatitudes in longer simulations. Lastly, we note that the tropical differences in vertical velocity are consistent with those in heating rates: Figure 6 from Charn et al. (2020) shows that, at least in climatological simulations, 2MOM results in greater mid-tropospheric equatorial heating, while 1MOM leads to greater heating around latitudes of 15° in either hemisphere, but particularly in the north.
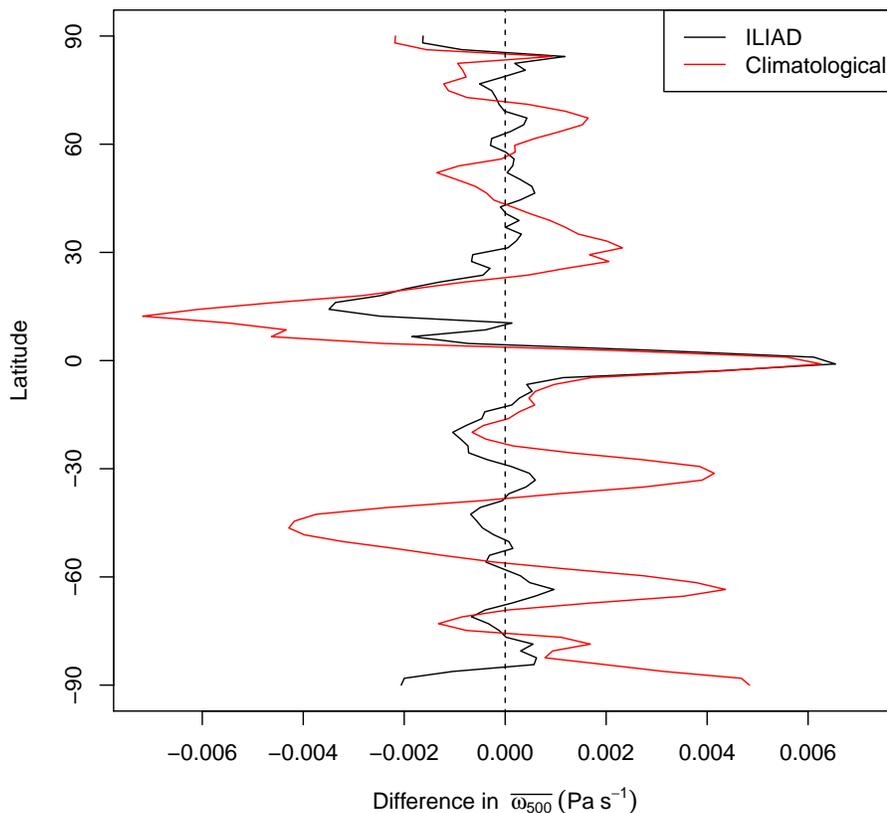
Figure 3.4: 1MG zonal mean of $\omega_{500\ \mathrm{hPa}}$ minus that of 2MG300.

To partition thermodynamic and dynamic contributions to extreme precipitation, scaling formulas involving vertical pressure velocity ω and specific humidity q_v (or its saturated

value q_v^*) have been previously developed. Both single-level (e.g., O'Brien et al., 2016), which involve products of predictor variables at specific heights or pressure levels, and full-tropospheric scalings (e.g., O'Gorman and Schneider, 2009), which involve integrals of water vapor flux, have been investigated in the literature. Here, we use a full-tropospheric approach, given in Equation 3.2, to investigate individual extreme precipitation events from the ILIAD integrations, complementing the results in Figure 3.4.

$$P = S(\omega, q_v^*) = \alpha \int_{1000 \text{ hPa}}^{200 \text{ hPa}} \omega \frac{\partial q_v^*}{\partial p} \frac{dp}{g}, \tag{3.2}$$

where P is precipitation, α is a precipitation efficiency, chosen so that $P = S(\omega, q_v^*)$ for each microphysics case, and g is gravitational acceleration.

Equation 3.2 can be rewritten, as in Fildier et al. (2017), as

$$P = S(\omega, q_v^*) = \alpha \underbrace{\langle\omega\rangle}_{M} \left\langle \underbrace{\frac{\omega}{\langle\omega\rangle}}_{\mu} \underbrace{\left(\frac{\partial q_v^*}{\partial p}\right)}_{\gamma} \right\rangle, \tag{3.3}$$

where the angle brackets denote a vertical integral over the troposphere: $\langle X \rangle = \int_{1000 \text{ hPa}}^{200 \text{ hPa}} X \frac{dp}{g}$. Thus, M is the column-integrated mass flux, μ is the normalized mass flux profile, and γ is the vertical gradient in q_v^*. Again as in Fildier et al. (2017), a fractional change in (extreme) precipitation can be written as follows:

$$\delta P = \underbrace{\delta\alpha}_{E} + \underbrace{\delta\text{M}}_{D1} + \underbrace{\frac{\langle\gamma\Delta\mu\rangle}{\langle\gamma\mu\rangle}}_{D2} + \underbrace{\frac{\langle\mu\Delta\gamma\rangle}{\langle\gamma\mu\rangle}}_{T} + N, \tag{3.4}$$

where ΔX represents a simple difference, $\delta X \equiv \Delta X / X$ a fractional difference, and N higher-order terms. Thus, $D1$ is a change in the column-integrated mass flux, $D2$ a change in the shape of the mass-flux vertical profile, and T any thermodynamic contribution to a difference in precipitation.

Here, we calculate E, $D1$, $D2$, and T for each precipitation extreme from the ILIAD simulations. More precisely, for each grid cell flagged as having a statistically significant extreme precipitation distribution, for each time in which both the 1MG and the 2MG300 rain rates are classified as extreme by our algorithm, the above four quantities are calculated. Simple and fractional differences are calculated with respect to the quantity from the microphysics case with the larger extreme; for example, if a 1MG precipitation rate is identified as the larger extreme, $D1 \equiv \delta M = (M_2 - M_1)/M_1$. We choose to only look at times at which both values are extreme in order to minimize the contribution of E, which can be relatively large when comparing an extreme with a non-extreme rain rate, reflecting the loss of accuracy of the scaling in the latter case. Relaxing this criterion and analyzing times at which precipitation is only extreme with one microphysics case is done in Figures B.3 and B.4 (described below).

Given that most of the statistically significant differences are between 30°S and 30°N, we focus our analysis within this latitude range. Figure 3.5a shows the zonal-median of E, $D1$, $D2$, and T for times when the 1MG extreme is larger, Figure 3.5b for times when the 2MG300 extreme is larger. Thus, each subplot allows us to infer which term is causing the rain rate from the other case (that is, 2MG300 in 3.5a and 1MG in 3.5b) to be smaller. Because we are not simulating climate change but rather conducting an experiment with two microphysics schemes in the present-day, we would not expect T to have a significant contribution. Indeed, this is confirmed in Figure 3.5, where T is generally less than 5% in magnitude. $D2$ is also generally associated with climate change: an upward shift of the mass flux profile is expected to accompany the upward shift in the radiative cooling profile (Singh and O'Gorman, 2012; Jeevanjee and Romps, 2018). Still, disparate microphysics schemes might be expected to cause changes not only to the column-integrated mass flux, i.e., $D1$, but also to the shape of the vertical profile, $D2$, via differences, for example, in riming and depositional growth, which in turn go on to affect latent heating (Van Weverberg et al., 2012). Nevertheless, the zonal profile of $D2$ is also small in Figure 3.5, with median values less than 5% in magnitude.

As expected, the contribution of E is relatively small, generally less than 10% in magnitude, though values can reach up to 15% at the outer edges of the tropics in Figure 3.5b. Figure B.3 shows the same plots as Figure 3.5 but is calculated for all extremes within grid cells flagged as having significantly different distributions, regardless of whether the rain rate from the other microphysics scheme is extreme. Figure B.4 is the same as Figure B.3 but for all grid cells in the tropics. As mentioned above, because it is possible that a rain rate is extreme in one microphysics case but not the other, the scaling in Equation 3.2 is not expected to hold for the non-extreme case in Figures B.3 and B.4. Thus, it is not surprising that E values are much higher in magnitude in these cases, reaching upwards of 80% in the subtropics in Figure B.3.

The largest contribution is that from $D1$; the median fractional decrease in mass flux is around 30% when the 1MG extreme is larger than the corresponding one from 2MG300, and around 15-20% when the 2MG300 extreme is larger (Figure 3.5). Figure 3.6 combines the information in Figure 3.5; here, the signs of all quantities used to compute Figure 3.5b are reversed, and the median across both cases is taken. Thus, positive values in Figure 3.6 denote latitudes where the contribution tends to yield higher extremes when using 2-moment microphysics. The total contribution, shown by the thick, black line, generally reflects the conclusions in Figure 3.3: 2MG300 results in higher extremes in a narrow band around the Equator, and the opposite is true in the subtropics. The total also follows the behavior of $D1$, confirming that changes in column-integrated mass fluxes are the primary driver of disparities in the tropical extreme precipitation distributions. This result is consistent with that of Fan et al. (2015), whose CRM domain-mean mass-flux vertical profile varied considerably with microphysics when simulating a mesoscale convective complex (MCC) during the Tropical Warm Pool International Cloud Experiment (TWP-ICE) (but not when modeling an MCC or a squall line during two case studies in the midlatitudes). We note, however, that Fan et al. (2015) argued that microphysical feedbacks to the convection's organization were responsible,
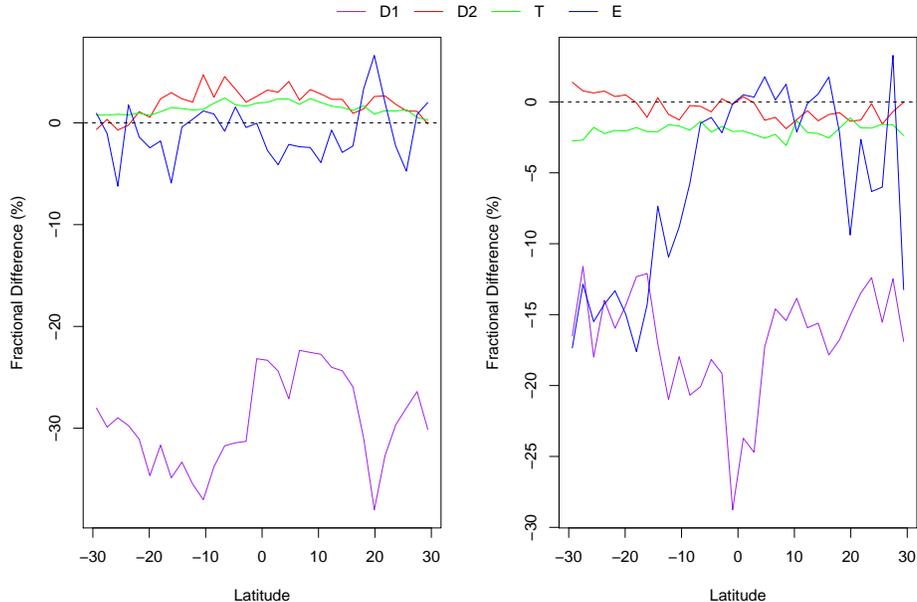
Figure 3.5: Zonal median of fractional changes of terms in Equation 3.2 for (left) times when the (ILIAD) 1MG precipitation extreme is greater than that from 2MG300 and (right) vice versa. Values are only calculated in grid cells with statistically significant differences in the (ILIAD) extreme precipitation distribution. Differences are computed such that negative values represent tendencies for lower column-integrated saturation vapor flux ($S(\omega, q_v^*)$ in Equation 3.2) in the microphysics case with the *smaller* extreme, i.e., 2MG300 in (a) and 1MG in (b).

whereas we do not investigate the frequency of such convective organization, though the work of Tao and Chern (2017) suggests convection might generally be more scattered in a CRM domain as small as ours (64 km).

3.3.4 Estimating Natural Variability of SPCAM Return Values

It is natural to wonder about the impact of interannual variability on our results, particularly in regions such as the subtropics with low annual rainfall, given our relatively short integration period of 5 years for the ILIAD simulations. To test this, we take the same approach as in Charn et al. (2020) by grouping the output into 10 years (consisting of the two microphysics experiments of 5 years each). The 10 years were then randomly reshuffled and the analysis redone. This process was performed 400 times, and the median number of grid cells by latitude marked as statistically significant is plotted in Figure 3.7. While there is indeed evidence of increased noise in the subtropics, interannual variability is still
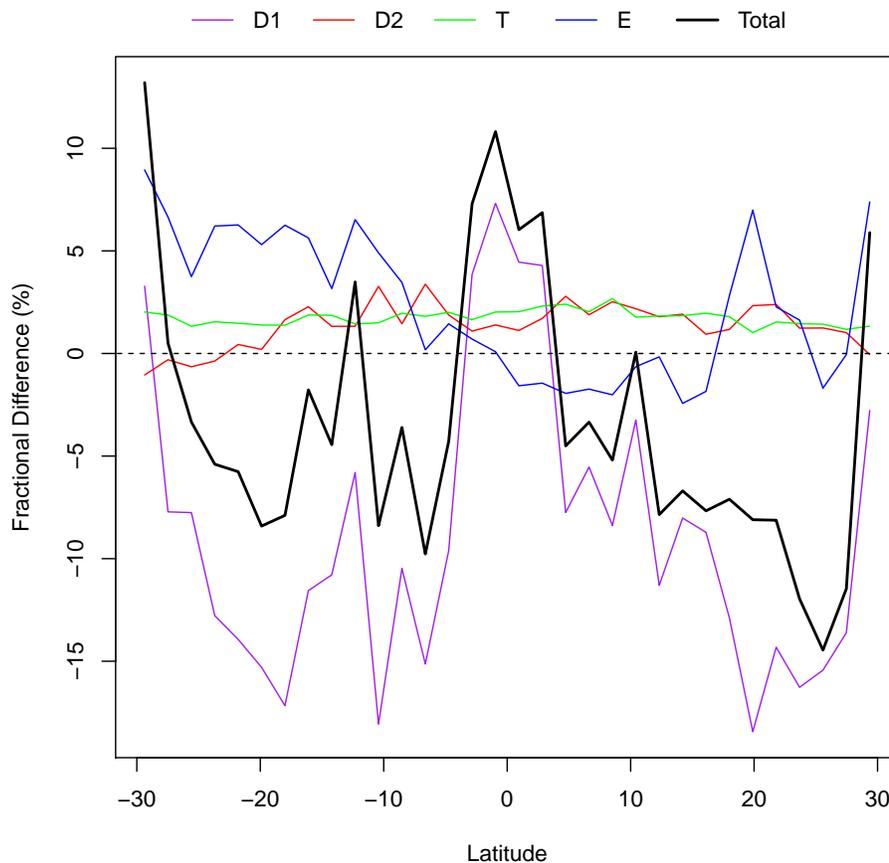
Figure 3.6: Zonal median of fractional changes of terms in Equation 3.2 for times when both
the 1MG and the 2MG300 precipitation rates are classified as extreme. Positive values repre-
sent tendencies for higher column-integrated saturation vapor flux ($S(\omega, q_v^*)$ in Equation 3.2)
for 2MG300. Values are only calculated in grid cells with statistically significant differences
in the (ILIAD) extreme precipitation distribution.

overwhelmed by the actual signal between 30°S and 30°N.

3.3.5   Comparing with Observations

Figure 3.8a (B.5) compares the climatological 1MG (2MG300) output with NOAA-CIRES-
DOE 20CRv3 and TRMM 3B42. As mentioned in Section 3.2.3, the comparison with TRMM
was carried out at daily resolution. With a few exceptions, 20CRv3 shows lower extremes
compared to SPCAM essentially everywhere. This is not surprising given that it is a re-
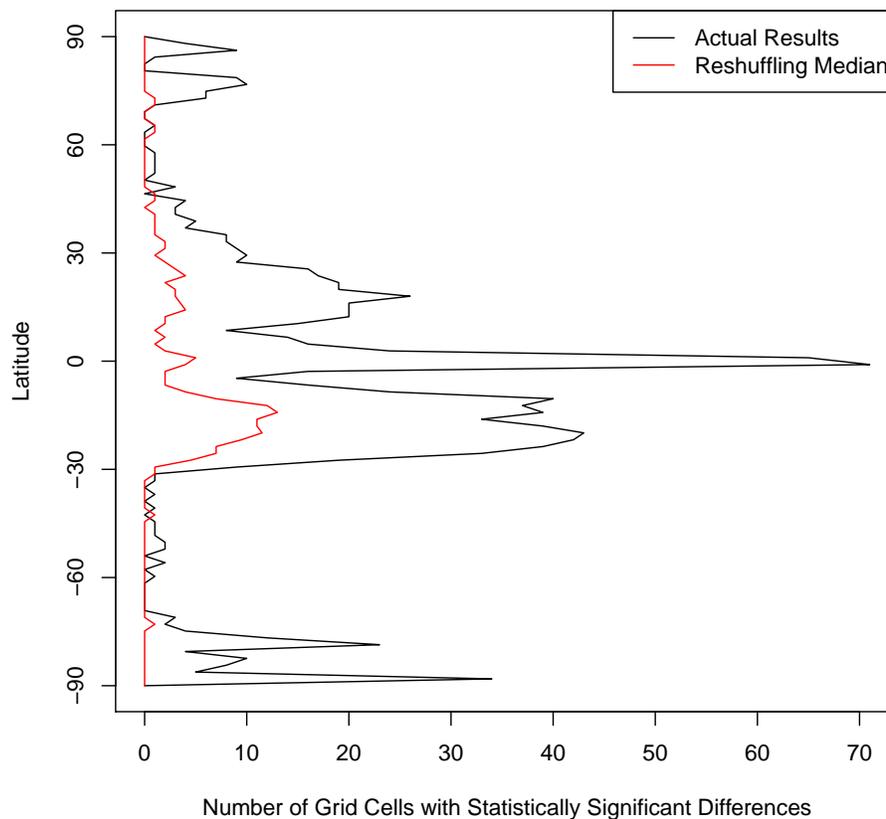
Figure 3.7: Number of grid cells at each latitude showing statistically significant differences
between 1MG and 2MG300 (Table 2.1) in the ILIAD simulations.

analysis product that relies on an atmospheric model with a cumulus parameterization to
generate its precipitation. On the other hand, TRMM 3B42's observed extremes generally
surpass those of SPCAM's. In particular, our findings of SPCAM's underestimation over
eastern CONUS, South America, the eastern equatorial Pacific, and the midlatitude Atlantic
Ocean match those of Kooperman et al. (2016) but have a robust, statistical footing here.

3.4   Conclusions

Earlier studies have found that superparameterization, which is a mechanism to introduce
convection-permitting processes into GCMs, can increase the fidelity of the precipitation
extremes simulated with these models. However, even as increasing numbers of climate
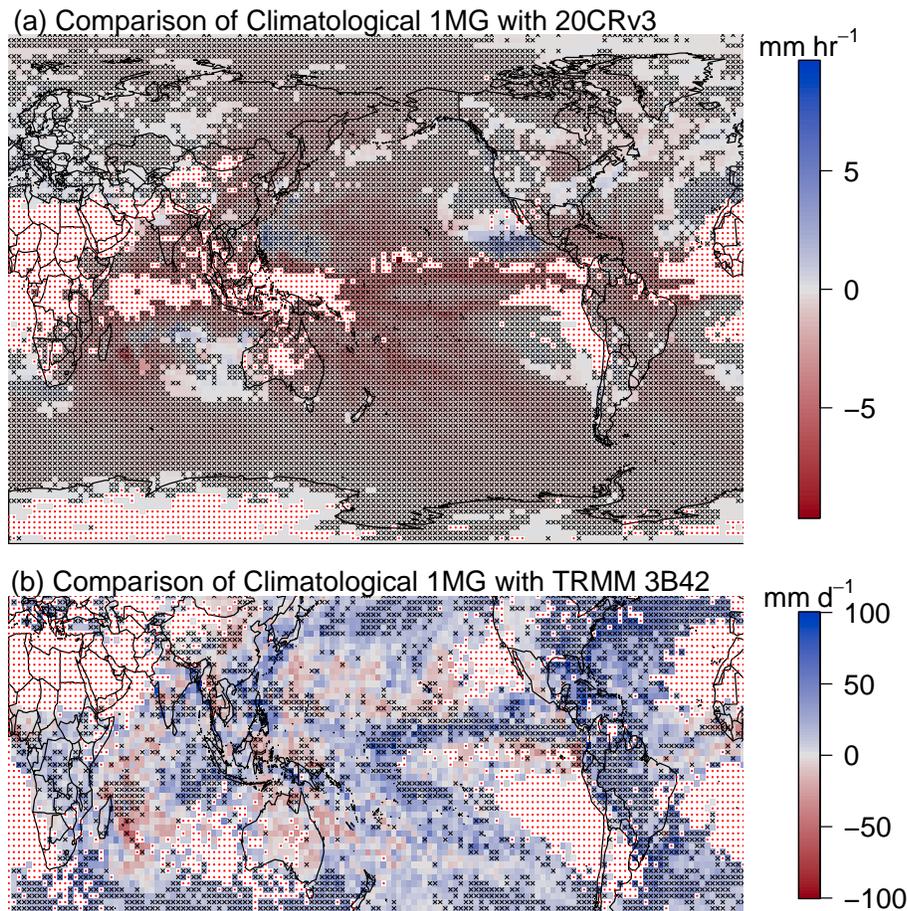
Figure 3.8: As in Figure 3.2, but the climatological 1MG (Table 2.1) 2-year return value
is subtracted from that of (a) 20CRv3 and (b) TRMM 3B42. Comparison with 20CRv3 is
done using 3-hourly precipitation while that with TRMM is done using daily data.

models transition to convection-permitting or -resolving atmospheric dynamics, microphys-
ical processes will still require parameterization. Two previous studies have investigated
the sensitivity of microphysics schemes within superparameterized models, though with a
regional focus on CONUS. While Elliott et al. (2016) found no discernible signal in sum-
mertime MCSs, Charn et al. (2020) found significant differences in extreme precipitation
distributions.

In this paper we generalize the methodology of Charn et al. (2020) to investigate extreme
precipitation globally. Similar to the aforementioned study, the largest differences occurred
when comparing simulations with different numbers of predicted moments. In a depar-
ture from previous findings, disparities persisted even during shorter integrations, though

primarily only in the tropics (30°S-30°N). Changes in vertical velocities, particularly the column-integrated profiles, were found to be responsible. 2-moment microphysics leads to greater upward motion in the deep tropics and greater subsidence in the subtropics. When looking back to the climatological simulations, changes in vertical velocities became readily apparent at higher latitudes as well.

As elaborated on in Charn et al. (2020), future improvements in microphysics parameterizations, particularly ice processes, would be helpful in terms of simulating clouds and precipitation. In addition, when employing superparameterization, finer resolution and larger domain sizes for the embedded cloud-resolving models may be necessary to properly represent organized convection, for example, mesoscale convective systems (Tao and Chern, 2017). While useful in lowering mean biases in precipitation, given that MCSs supply more than half of rainfall in most tropical regions (Nesbitt et al., 2006) and much of warm-season rainfall in midlatitudes (Schumacher and Johnson, 2006; Zhang et al., 2019), it is reasonable to assume that extreme rain rates would be more faithfully simulated as well.

Chapter 4

Predictive Proxies of Present and Future Lightning in a Superparameterized Model

A superparameterized climate model is used to assess the global performance of several previously proposed proxies for lightning. In particular, predictors incorporating hydrometeor (ice, graupel) profiles and convective vertical velocities are compared to observations, then used to estimate changes in flash rates with global warming. The choice of microphysics parameterization is also investigated, with four out of five predictors showing lower root-mean-square errors (up to 10%) when using a 2-moment scheme compared to a 1-moment representation. All proxies generally agree in their response to warming over tropical land, with notable decreases in Africa, the Middle East, and northern South America but disagree over oceans and the midlatitudes. The product of convective available potential energy and precipitation predicts increases over these areas, as do the 2-moment ice-based proxies, while those of the 1-moment model tend to show decreases, highlighting the importance of cloud microphysics when using climate models to simulate lightning.

The content of this chapter has been submitted to *Geophysical Research Letters*.

4.1   Introduction

Lightning simulation within weather and climate models remains an outstanding challenge, particularly when it comes to predicting changes with global warming. A variety of predictors have been proposed in the literature, ranging from large-scale variables, including cloud-top height (Price and Rind, 1992) and convective available potential energy and precipitation (Romps et al., 2014), to storm-scale updraft velocities and precipitating ice volumes (Barthe et al., 2010), to explicit simulation of the storm's electric field (Barthe et al., 2012; Fierro et al., 2013). Today, even machine learning is being brought to bear on lightning forecasting, though typically with an eye towards imminent warnings up to 30 minutes in advance

(Mostajabi et al., 2019).

When it comes to trying to predict the lightning response to global warming, most proxies have been evaluated in models with horizontal resolution on the order of 100 km (e.g., Price and Rind, 1994; Michalon et al., 1999; Romps et al., 2014; Finney et al., 2018), which is incapable of directly resolving the storm updrafts thought to be responsible for lightning generation. However, with increasing computing power comes the possibility of instead using cloud-resolving models (CRMs), which have the resolution to explicitly simulate convective processes. Indeed, Romps (2019) recently used the regional simulations of Rasmussen and Liu (2017), as well as his own limited-domain CRM, to investigate changes in lightning over the contiguous United States (CONUS) and the tropical oceans, respectively.

Despite advances in computing power, global cloud-resolving models with kilometer-scale resolution, as is typically used by its limited-domain counterparts, remain impractical for climate-length simulations. A compromise is superparameterization, in which 2D CRMs are embedded in each grid column of a conventional global climate model (GCM), replacing the convective and boundary layer schemes. At each GCM time step, the large-scale grid forces the CRMs, which then provide subgrid convective heating and moistening tendencies. Superparameterization has been shown to improve simulation of the diurnal cycle of convection (Khairoutdinov et al., 2005), the Madden-Julian Oscillation (Benedict and Randall, 2009), and extreme precipitation (Li et al., 2012). Here, we implement lightning proxies into a superparameterized (SP) model, compare them to present-day observations, and investigate changes in a climate subject to a uniform sea-surface temperature (SST) increase as a proxy for changes in response to global warming. Because lightning generation, and thus, many of the predictors developed to model it, is believed to be heavily dependent on cloud microphysics (Reynolds et al., 1957; Takahashi, 1978; Barthe et al., 2010), we also test the sensitivity of our results to the microphysics scheme within the embedded CRMs.

4.2 Simulation Setup

We analyze the output of two simulations from Charn et al. (2020), conducted using SPCAM, where CAM is the atmospheric component of the Community Earth System Model (Hurrell et al., 2013). The following text in this section is derived from Charn et al. (2020) with minor modifications. In the control run SPCAM was forced by climatological monthly SSTs and sea ice boundary conditions (Hurrell et al., 2008), and in the warmer-climate simulation, SSTs were uniformly increased by 4 K. The integrations were carried out for 5.5 years, with the first 6 months discarded as spinup. The embedded CRM is the System for Atmospheric Modeling (Khairoutdinov and Randall, 2003). In the experiments the CRMs were configured in a 2D, east-west orientation with 32 columns at 2-km horizontal resolution and 30 vertical levels corresponding to those in CAM. The large-scale, outer grid was implemented using a finite-volume dynamical core with a 1.9° latitude × 2.5° longitude resolution. Greenhouse gas concentrations were held fixed at year-2000 levels, and (bulk) aerosols were prescribed.

We investigate the impacts of using two different microphysics formulations: the 1-moment scheme described in Khairoutdinov and Randall (2003), and the 2-moment one from Morrison et al. (2005), hereafter 1MOM and 2MOM, respectively. The number of moments refers to the number of degrees of freedom in the size distribution of a precipitating hydrometeor, which is prescribed to be an inverse exponential (Marshall and Palmer, 1948). 1MOM contains prognostic equations for hydrometeor mixing ratios, from which the number concentrations are diagnosed, while 2MOM prognoses both. With its extra free parameter and set of prognostic equations, a 2-moment scheme would be expected to generate more realistic cloud dynamics and properties, such as liquid and ice water contents and radiative fluxes, and indeed it has been shown to do so in several studies (Lee and Donner, 2011; Bryan and Morrison, 2012; Igel et al., 2015).

4.3   Proxies

We evaluate five proxies, the product of convective available potential energy (CAPE) and precipitation (hereafter CAPE × P) calculated two different ways, and three proxies that directly involve ice-related quantities.

4.3.1   CAPE × P

CAPE × P has been mostly evaluated over CONUS (Romps et al., 2014, 2018; Romps, 2019; Tippett et al., 2019), though Romps et al. (2018) also documented an excellent global correspondence with lightning products over land. While, like many other proxies, it does not capture the long-noted land-ocean contrast in flash rates (Brooks, 1925; Orville and Henderson, 1986), CAPE × P does demonstrate a very good correlation over tropical oceans themselves (Romps, 2019).

CAM calculates CAPE using a dilute-plume model, in which entrainment of environmental air is incorporated (Neale et al., 2008). In addition to the model's output (which we will call $CAPE_d$), we also compute CAPE pseudoadiabatically ($CAPE_{ud}$). Here, a parcel does not undergo entrainment but does experience complete condensate fallout, both of which will tend to result in higher values. An adiabatic approach is more typically used, for example, in previous calculations of the CAPE × P proxy (Romps et al., 2014; Romps, 2019) and in studies aiming to understand why CAPE increases with global warming (Seeley and Romps, 2015; Romps, 2016). We will compare the performances of both methods when calculating CAPE × P as well as investigate whether predicted changes with warming differ in sign and/or magnitude. In both cases CAPE is calculated instantaneously, and precipitation is averaged over the following 3 hours, with CAPE × P as the product.

4.3.2   IFLUX

Finney et al. (2014) proposed the flux of non-precipitating ice at the 440-hPa level in an

effort to develop a more physically motivated lightning proxy related to the non-inductive charging mechanism, which is thought to be primarily responsible for charge separation. Cloud ice is also available in the ERA-Interim (and now ERA5) product, making testing possible with reanalysis datasets. Romps (2019) evaluated the proxy within cloud-resolving simulations over CONUS with one alteration: instead of looking at the flux at the 440-hPa isobar, the 260-K isotherm was used, which improved the performance over the western United States as well as further bolstered the proxy's physical basis. Here, we adopt the same modification, denoted as IFLUX, but with the -15°C (258.15-K) isotherm to be consistent with the graupel-flux proxy described below.

4.3.3   GFLUX

Since the non-inductive charging mechanism involves the collisions of ascending cloud ice with falling graupel, other studies have investigated the correlation between lightning and graupel volumes and fluxes (e.g., Carey and Rutledge, 1996; Carey et al., 2014). McCaul Jr et al. (2009) implemented a proxy into the Weather Research and Forecasting model that consisted of two parts, one of which is the flux of graupel at the -15°C isotherm. Here, we denote this as GFLUX, slightly modified to incorporate air density, which gives the proxy mass units, unlike its original formulation.

4.3.4   I × G

The rate of collisions per unit volume between ascending ice crystals and descending graupel is proportional to the product of the two number concentrations. While a more detailed proxy would take into account collision and subsequent separation efficiencies and relative size distributions and fall speeds, we adopt the approach of Romps (2019) and simply vertically integrate the product of the densities of ice and graupel, referring to this quantity as I × G. We note the caveat that while this may be a reasonable approach when using 1MOM, 2MOM allows hydrometeor mixing ratios and number concentrations to evolve independently. Unfortunately, number concentrations were not saved during the 2MOM simulations of Charn et al. (2020), and due to computational limitations, we did not carry out those runs again.

4.3.5   Summary of Proxies

All five predictors are output on the large-scale GCM grid. To do this for the three ice-based proxies, the IFLUX, GFLUX, and I × G quantities are calculated for each of the 32 CRM columns, then summed to arrive at the final GCM-grid value. Table 4.1 provides a summary.

Table 4.1: Summary of Lightning Proxies and Present-Day Land-Ocean Ratio

Proxy	Land-Ocean Ratio[a]	
	1MOM	2MOM
"IFLUX"$= \sum\limits_{j=1}^{32} \rho_i(j)w(j)\mathcal{H}[w(j) - 1 \text{ m/s}]$, where $T(j) = 258.15$ K	1.1	2.0
"GFLUX"$= \sum\limits_{j=1}^{32} \rho_g(j)w(j)\mathcal{H}[w(j) - 1 \text{ m/s}]$, where $T(j) = 258.15$ K	1.1	1.5
"I $\times$ G"$= \sum\limits_{j=1}^{32} \int\limits_{0}^{\infty} \rho_i(j)\rho_g(j)dz$	1.0	2.1
"CAPE$_\text{d}$ $\times$ P"$=$ CAPE$_\text{d}$ $\times$ P	0.6	0.8
"CAPE$_\text{ud}$ $\times$ P"$=$ CAPE$_\text{ud}$ $\times$ P	0.6	0.8

Note. $\mathcal{H}$ denotes the Heaviside step function, w the (CRM) vertical velocity, and ρ_i and ρ_g the mass of ice and graupel per volume of air, respectively. The three ice-based proxies (IFLUX, GFLUX, and I×G) are calculated for each CRM column $j \in 1, 2, ...32$, then summed across the 32 columns to yield values on the GCM grid, where the CAPE $\times$ P proxies are evaluated.

[a]As a comparison, the LIS/OTD land-ocean ratio is 8.3.

4.4 Proxy Validation

To validate the proxies, we make use of a merged climatology using data from the Lightning Imaging Sensor and the Optical Transient Detector (LIS/OTD; Cecil et al., 2014). None of the proxies reproduces the land-ocean contrast in lightning particularly well, though 2MOM does a better job in all cases (Table 4.1). We note that while the LIS/OTD land-ocean ratio is 8.3, oceanic flash rates are higher in the World Wide Lightning Location Network observations (Field et al., 2018), leading to a ratio of around 2.6 (Romps et al., 2018).

Figure 4.1 shows the annual mean of the LIS/OTD observations, as well as the difference when comparing to the five proxies, which are hereafter normalized to reproduce global-average flash rates. Different normalization constants are used over land and ocean in order for each proxy to match the respective global means of LIS/OTD. This is clearly not a panacea for a proper prediction of the land-sea contrast. All five predictors overestimate flash rates over the Intertropical Convergence Zone (ITCZ) and underestimate just off coastlines, notably those on the eastern side of continents in the midlatitudes, where convection would be expected to have a more continental character. Over land, we see that all of the ice-based proxies generally overestimate lightning in the midlatitudes, though this is significantly ameliorated with 2MOM over Canada, Europe, Russia, and the western parts of the United States and China. The ice proxies mostly underestimate in the tropics, especially the peak in central Africa, though again 2MOM does better in this regard. Given that most lightning occurs in the tropics (Figure 4.1a), the ice proxies arguably spread the latitudinal lightning distribution too far poleward. But CAPE $\times$ P is too concentrated in the tropics (Figure

C.1), with predicted flash rates generally underestimated over land in the midlatitudes and overestimated in the tropics, though even CAPE × P has trouble capturing the peak in central Africa. Perhaps unsurprisingly, the main differences between 1MOM and 2MOM CAPE × P are in the tropics, with 2MOM performing better over Africa and eastern South America but worse over western South America. The root-mean-square errors (RMSEs) in Figure 4.1b show that GFLUX has the best match to observations of all five proxies, followed by IFLUX. I × G and $CAPE_d$ × P are roughly similar, and $CAPE_{ud}$ × P shows the largest RMSEs.

4.5 SST + 4 K Warming Projections

Figure 4.2 depicts changes in flash rates when going from the control climate to one in which SSTs are uniformly increased by 4 K using 1MOM, while Figure 4.3 does the same for 2MOM. Some general observations follow.

4.5.1 CAPE × P Projections

The CAPE × P response is fairly similar regardless of the microphysics scheme, predicting decreases across most of Africa, a substantial part of northern South America, southern Europe, and the Middle East while simulating increases in Russia, the United States, Canada, and most oceans except parts of the North Atlantic.

While not obvious from Figures 4.2 and 4.3, Figure C.2 shows that, in the tropics, $CAPE_{ud}$ × P tends to predict a ~7%/K greater increase than its dilute counterpart, suggestive of a Clausius-Clapeyron (CC) scaling. Indeed, $CAPE_d$ changes in magnitude less than ~1%/K in the tropics (not shown), which might be surprising given previous findings that CAPE undergoes a CC scaling with warming in radiative-convective equilibrium (RCE) simulations, which the tropics are generally considered to be in (Singh and O'Gorman, 2013; Seeley and Romps, 2015; Romps, 2016). However, as stated in Section 4.3.1, the CAM formulation of CAPE includes entrainment (at a fixed rate of 100% km^{-1}). Singh and O'Gorman (2013) observed that tropical cloud buoyancies, both in observations and CRM simulations, are small (typically < 0.5 K), indicating that the tropical lapse rate is close to that set by entraining clouds, and that this persists in simulations with increased SST. Undilute buoyancy thus increases with warming due to the greater difference in moist static energy, specifically the moisture component, which undergoes a CC scaling and can then be converted to a change in temperature between a dilute plume and an undilute plume. Thus, it is perhaps not surprising here that tropical $CAPE_{ud}$ increases at ~7%/K while tropical $CAPE_d$ stays roughly constant. $CAPE_d$ is more likely to capture the magnitude of the vertical velocity of small, entraining clouds, such as those that set the lapse rate of the tropical free troposphere. However, we speculate that clouds that achieve the high vertical velocities thought to be needed for lightning are more likely to be larger and thus insulated from the effects of entrainment. While CRM simulations show typical entrainment rates of O(100% km^{-1})
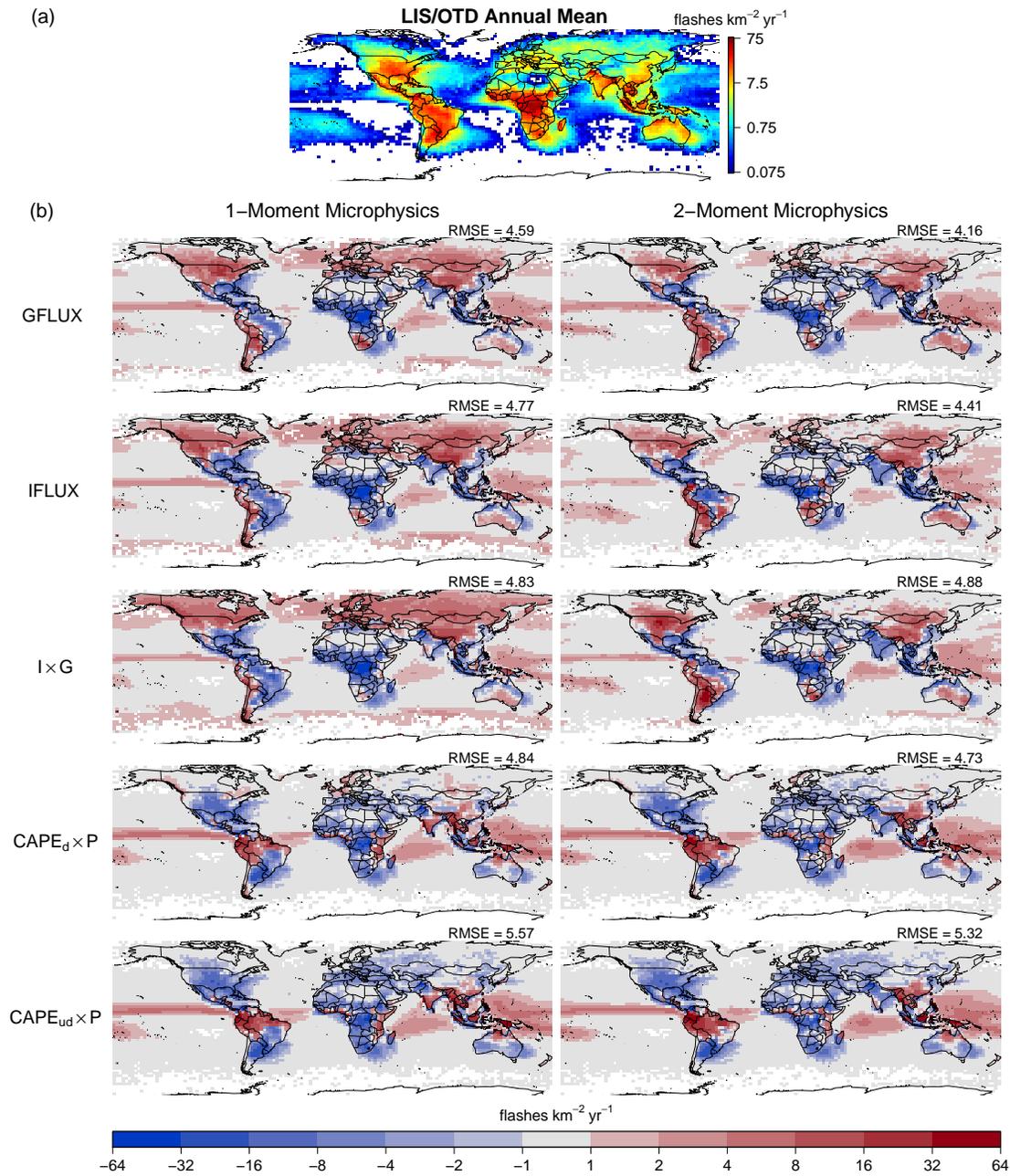
Figure 4.1: (a) Annual mean flash rate of LIS/OTD observations. (b) Model results minus observation for the five lightning proxies simulated by SPCAM. The proxies were interpolated from the large-scale SPCAM resolution (1.9° latitude × 2.5° longitude) to the LIS/OTD grid (2.5° × 2.5°). White regions correspond to grid cells with no observations.

(Romps, 2010), the same value as in CAM's dilute-plume model, deep convective updrafts that make it to the tropopause have effective entrainment rates on the order of 5% km^{-1} (Romps and Kuang, 2010). Thus, a best guess for the projected change in flash rates using CAPE × P is somewhere in between the two cases analyzed here.

4.5.2 Ice-Proxy Projections

Projections of future lightning diverge when considering the two microphysics schemes and the ice-based proxies. Figure 4.2 shows that 1MOM predicts decreases over most of the globe, a result even more extreme than that of Finney et al. (2018), whose decreases were mostly confined to the tropics. Here, only a few areas, such as southern South America, the southeastern United States, northern Russia, and northeastern China, show the reverse trend. Interestingly, the ITCZ also shows signs of increases over the ocean. On the other hand, Figure 4.3 shows that 2MOM predicts more in the way of increases, particularly over oceans, though also in the United States, Canada, China, and Russia.

4.5.3 Multi-Proxy Mean Changes

To get a sense of where the proxies agree on future changes in flash rates, the upper right sub-panels of Figures 4.2 and 4.3 depict the multi-proxy mean changes using their respective microphysics schemes. Regions where at least four out of the five predictors agree in sign are *un*stippled. Regardless of the scheme, Africa, the Middle East, and northern South America tend to show large areas of decreased lightning, while southern South America, the southeastern United States, and northern Russia tend to show increased lightning. The 2MOM proxies tend to be in agreement in sign over a greater fraction of the world, and Figures C.2 and C.3 show that the spread between the ice-based predictors and the CAPE × P ones tends to be smaller with 2MOM, particularly over the ocean. Thus, 2-moment microphysics leads to a possible reconciliation of the opposing responses between CAPE × P and ice-based proxies in the 1-moment, tropical maritime RCE simulations of Romps (2019).

4.6 Discussion and Conclusions

While the disagreement between 1MOM and 2MOM is arguably more pronounced in the midlatitudes, we investigate the disparities generated from the two microphysics schemes by plotting profiles of ice quantities for the tropics (30°S-30°N) in Figure 4.4 to compare with the results of Romps (2019). There, in tropical RCE simulations, the water vapor mass flux was found to be an invariant function of isotherm, a not unexpected result given previous findings that many other atmospheric quantities, for example, relative humidity (Romps, 2014), radiative cooling (Jeevanjee and Romps, 2018), and the tropopause (Seeley et al., 2019) are also invariant functions of isotherm. Since hydrometeor mass fluxes were also roughly invariant with temperature in the simulations of Romps (2019), two implications
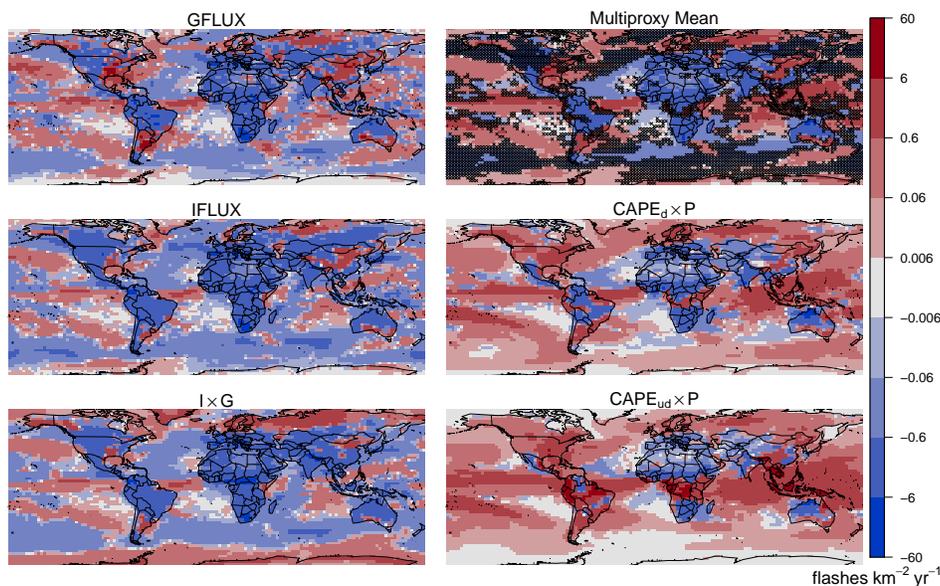
Figure 4.2: Change in lightning flash rate predicted by the five proxies using 1MOM, as well as the multi-proxy mean change. Stippling in this subpanel denotes grid cells where *fewer* than four out of the five proxies agree on the sign of the change.

followed: 1) IFLUX showed small changes with warming, and 2) the domain-mean product of ice and graupel densities (and thus, I × G) decreased with warming, given that vertical velocities increased. Here, we come to a similar conclusion using 1MOM microphysics, but we also find that the 2MOM microphysics can break that invariance. While the water vapor mass flux appears to still be an invariant function of isotherm, this is less the case with the ice and graupel fluxes, which show appreciable fractional increases in the mixed-phase region between 240 K and 273 K. In addition, the domain-mean product of ice and graupel densities actually *increases* with warming, in stark contrast to the findings of Romps (2019). As noted by Romps (2019), however, there is not an obvious reason to expect individual hydrometeor mass fluxes to be invariant with temperature even if the water vapor flux is, since the formation of condensates could be sensitive to the microphysics formulation. Indeed, Figure C.4 shows that the simulated total hydrometeor flux (other than the rain component, which was not output for 2MOM) displays the expected temperature invariance, at least above 273 K, below which rain would be a significant component. The graupel flux increases in the warmer climate at the expense of the cloud-water flux. The ice flux increases as well, but it is a relatively small fraction of the total hydrometeor flux in 2MOM. As to the result in Figure 4.4d, the decrease in the stock of ice (and I × G) in Romps (2019) was the result of the combination of condensate mass-flux invariance and an increase in vertical velocities. Here,
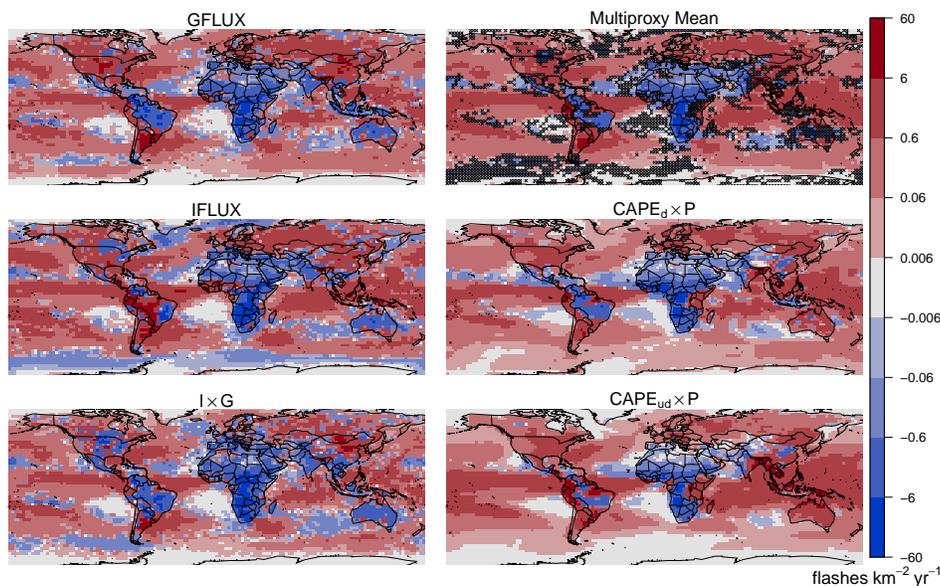
Figure 4.3: As in Figure 4.2 but using 2MOM.

it appears that increases in stock (and thus, I × G) and in vertical velocity work together
in 2MOM to drive higher ice and graupel mass fluxes in the mixed-phase region (and thus,
IFLUX and GFLUX, respectively) in the tropics.

In summary, using superparameterization, we have found that 2-moment microphysics
is generally better able to predict current spatial distributions of lightning, based on the
lower root-mean-square errors in four out of the five tested proxies when comparing against
LIS/OTD observations. There is also a broad agreement between 2MOM ice-based predictors
and CAPE × P regarding changes in a warmer climate–an agreement conspicuously lacking
in the 1MOM experiments. None of the proxies capture the observed land-ocean contrast in
lightning, though 2MOM does better quantitatively, with two out of three ice proxies showing
a twofold amplification over land. While all ice-based predictors generally overestimate flash
rates in the midlatitudes and the tropical oceans and underestimate over tropical land, there
is a general improvement with 2MOM, and the graupel flux at the -15°C isotherm shows the
smallest RMSE. CAPE × P shows the opposite trend, underestimating in the midlatitudes
and overestimating in the tropics except for the peak in central Africa.

While the uniform +4 K SST experiment does allow for circulation changes, unlike the
pseudo global warming approach, effects due to patterned SST increases and the direct effect
of CO_2 are not accounted for. Since the direct effect reduces the increase in precipitation with
global warming by roughly 25% (Romps, 2020), a first estimate of the impact of increased
CO_2 would be to ameliorate increases in flash rates using CAPE × P. Further superparame-
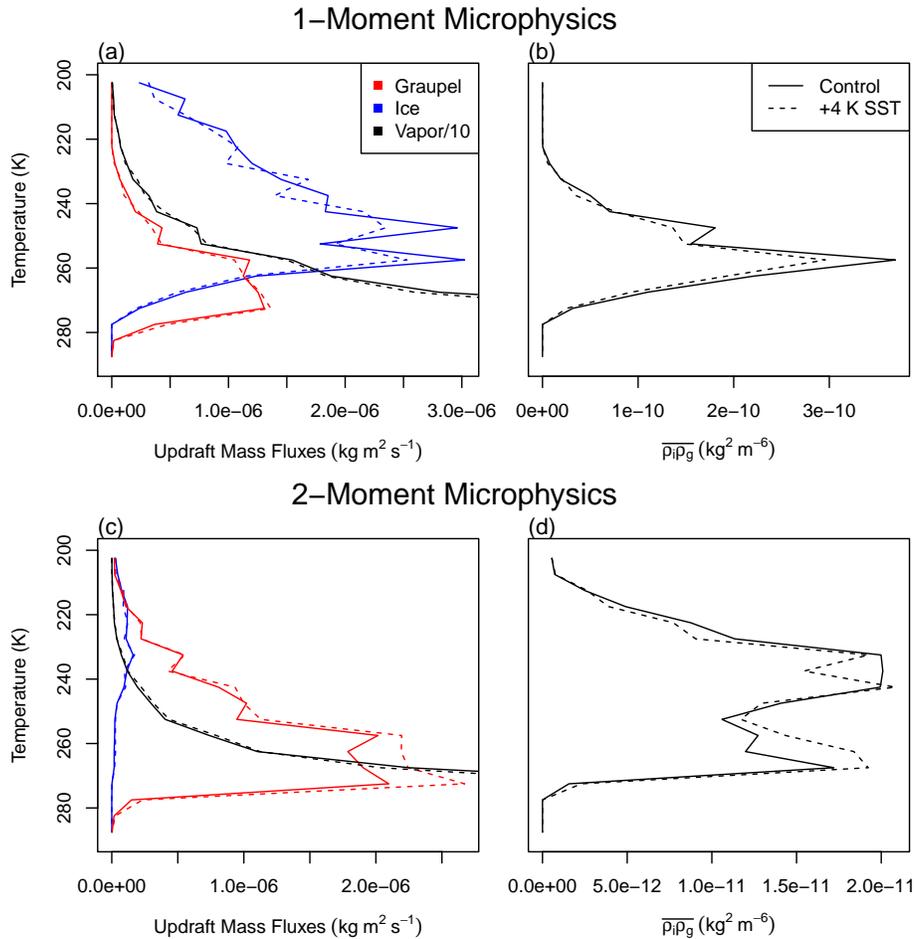
Figure 4.4: (a,c) Updraft mass fluxes, plotted as tropics-wide (30°S-30°N) means of $\rho_x w \mathcal{H}(w - 1 \, \text{m/s})$ in bins of width 5 K, where x denotes graupel, ice, or vapor. Note that vapor fluxes are divided by 10 to fit the x-axis scale, that is, actual values are 10 times larger. (b,d) Tropics-wide mean product of ice and graupel densities in bins of width 5 K.

terized (or global cloud-resolving) simulations that are coupled with an active ocean model, or atmosphere-only models such as SPCAM that are spun up from coupled conventional GCMs simulating climate change, would be useful for future work.

Chapter 5

No Definitive Evidence for Warm-Cloud Lightning using GOES-16 Observations of Cloud-Top Temperatures

Ice is often invoked as a necessary ingredient for thunderstorm electrification, yet there have been reported visual accounts of lightning occurring in clouds below the local freezing-level height. Subsequent attempts via aircraft to confirm the existence of such warm-cloud lightning have been unsuccessful. However, given the seeming rarity of this phenomenon, a more systematic search is desirable. The launch of satellites within the last few decades that are able to measure cloud-top temperatures (CTTs) provides such an opportunity. Here, CTTs and lightning detected by the Geostationary Operational Environmental Satellite-16 are combined with the National Lightning Detection Network and the Lightning Imaging Sensor onboard the International Space Station to search for warm-cloud lightning from July 27 to August 31, 2019. While warm-cloud flashes are found within each individual lightning dataset, efforts to validate flashes across datasets resulted in zero matches. Thus, we conclude that there is no definitive evidence of warm-cloud lightning using the listed datasets over this time period.

5.1   Introduction

Cloud electrification is generally theorized to necessitate ice, whether via an inductive process (i.e., within an already-present electric field) or a non-inductive one (Kuettner et al., 1981; Shi et al., 2016). Notably, precipitating graupel, or soft hail, is usually invoked as colliding with ascending ice crystals or, possibly in the case of the inductive process, supercooled water droplets. However, there have been purported eyewitness accounts of warm-cloud lightning: lightning occurring in clouds completely below the local freezing-level height (Foster, 1950;

Pietrowski, 1960; Moore et al., 1960; Rossby, 1966; Lane-Smith, 1971). A lack of ice, or even merely of precipitating ice, in clouds producing lightning, would demand an electrification mechanism separate from those mentioned above (Moore et al., 1960).

Given that all of these visual observations were made from a distance of tens or even hundreds of kilometers away, some have cast doubt on the accuracy of these accounts, e.g., Appleman (1957). Indeed, even the aforementioned eyewitnesses expressed difficulty in obtaining repeated observations (Moore et al., 1960; Lane-Smith, 1971). Moore et al. (1960) suggested several reasons for the exceedingly low frequency of observed warm-cloud lightning. First, all accounts describe the phenomenon as comprised of intracloud, as opposed to cloud-to-ground, flashes; such lightning is only readily visible at twilight or at night. Second, all previous observations were made at sea, with the exception of that of Lane-Smith (1971), and even this instance was documented on the coast of Sierra Leone. If indeed warm-cloud lightning occurs primarily over the ocean, it follows that there would be a smaller chance of human detection. Third, Moore et al. (1960) suggests warm-cloud lightning is a transitory phenomenon, with the parent clouds dissipating after a few strokes or continuing their ascent past the freezing level, though this is perhaps somewhat at odds with Pietrowski (1960), who observed the "interval between flashes...to be three or four seconds...for a period of five minutes".

Considering that warm-cloud lightning likely occurs rarely, if at all, it is perhaps not surprising that more recent, intentional efforts to document this phenomenon via aircraft studies have been unsuccessful (e.g., Petersen et al., 1996; Carey and Rutledge, 2000). An alternate method, using satellite observations of cloud-top temperatures and lightning, was suggested in the 1980s (Davis et al., 1983), and, interestingly, Vaughan Jr and Boeck (1998) claimed to have identified an instance of warm-cloud lightning using video footage taken during the STS-70 mission of the Space Shuttle *Discovery*. A corresponding GOES-8 image was used to estimate the cloud-top temperature at 271 K, suggesting "little if any evidence of frozen or melting precipitation." Here, a more systematic approach is proposed, in which observations of cloud-top temperature via the recently launched Geostationary Operational Environmental Satellite (GOES)-16 are used in conjunction with various lightning datasets, with the goal of robustly identifying warm-cloud flashes, which will be defined more precisely below.

5.2 Observational Datasets

5.2.1 Cloud-Top Temperature

GOES-16 Advanced Baseline Imager

The Advanced Baseline Imager (ABI) is the primary Earth-observing sensor on GOES-16, which sits in the GOES-East position at 75°W, providing a centered view of the Americas. Of its sixteen spectral channels, three in the infrared range (11.2-13.3 μm) are used in the ABI Cloud Height Algorithm (ACHA), which calculates the cloud-top height, pressure,

and temperature in all pixels deemed cloudy by the ABI Cloud Mask. The ACHA has been validated against data from the Moderate Resolution Imaging Spectroradiometer and the Cloud-Aerosol Lidar with Orthogonal Polarization and compares favorably in terms of accuracy (mean bias), though less so in precision (standard deviation of bias) (Heidinger, 2013). The ABI cloud-top temperature (CTT) product has a temporal resolution of 10 minutes and a nadir horizontal resolution of 2 km.

5.2.2 Lightning

5.2.2.1 GOES-16 Geostationary Lightning Mapper

The Geostationary Lightning Mapper (GLM) detects the optical signature of lightning by sensing the characteristic emission at 777.4 nm of excited oxygen atoms decaying to their original state. The GLM employs a variety of filters to improve its performance. One example that increases the signal-to-noise ratio is the use of 56 Real-Time Event Processors (RTEPs), which independently analyze the background scene within their subarrays and set the threshold above which a signal must be in order to be classified as a lightning event (Goodman et al., 2012). Another is a blooming filter, which helps decrease the number of false positives (Koshak et al., 2018). Incorporated as a software update on July 25, 2019, the blooming filter decreases the number of false events due to solar glint (specular reflection off of cloud boundaries or bodies of water) and solar intrusion (when the Sun shines directly into the GLM). GLM data was taken from July 27 to August 31, 2019 for this analysis.

The GLM detects what are called *events*: individual pixels that exceed a background threshold during one frame. Events that occur in adjacent pixels in the same 2-ms integration period are combined into *groups*. Two or more groups within 16.5 km and 330 ms of each other, using a weighted Euclidean distance (WED), are classified as belonging to the same *flash* if WED < 1.0.

$$\text{WED}^2 = (X/16.5 \text{ km})^2 + (Y/16.5 \text{ km})^2 + (T/330 \text{ ms})^2, \tag{5.1}$$

where X is the east-west distance and Y the north-south distance between the two group centroids, and T is the time between the two groups.

5.2.2.2 Lightning Imaging Sensor

After the decommissioning of the Tropical Rainfall Measuring Mission satellite, which housed the original Lightning Imaging Sensor (LIS), in 2015, a spare LIS was made operational on the International Space Station (ISS) in 2017. The LIS is essentially the same instrument as the GLM (Goodman et al., 2013) but with lesser sensitivity due to its decreased telemetry bandwidth (even after accounting for the fact that it has fewer pixels). Its clustering algorithm is also nearly identical and uses Equation 5.1 to classify flashes but with a maximum distance between groups of 5.5 km instead of 16.5 km. We note that "non-quality controlled" data are used here (Blakeslee, 2019), as quality controlled data are not yet available at the

time of writing. As with the GLM, LIS data was taken only from July 27 to August 31, 2019. We note that while LIS has nearly global coverage, the search for warm LIS lightning will be constrained to the GOES-East domain.

5.2.2.3 National Lightning Detection Network

The National Lightning Detection Network (NLDN) is comprised of over 100 ground-based sensors around the contiguous United States (CONUS), as well as two in the Bahamas (Nag et al., 2014). The sensors employ a combination of magnetic direction finding and time-of-arrival techniques to identify the time and location of lightning strokes. Following a first stroke, subsequent strokes no more than 10 km and 1 s away are grouped into the same flash. Only flash (as opposed to stroke) data was available for this analysis, again for the time period between July 27 to August 31, 2019. The NLDN data here are unfiltered so as to keep flashes that are likely to be intracloud, which as mentioned before, comprise all of the previous observational accounts of warm-cloud lightning.

5.3 Identifying Warm-Cloud Lightning

In this section we describe the procedure used to identify warm-cloud lightning using the observations listed above. Given that our goal is to confirm the existence of even a single instance of this phenomenon, we adopt a conservative approach.

5.3.1 Identifying Warm-Cloud Candidates

First, flashes from each of the three lightning datasets are evaluated. For GLM and LIS, every event comprising a flash must be warm in order for the flash to be considered a warm-cloud candidate. To be considered warm, all ABI pixels within 10 km and 30 minutes of a GLM event or an NLDN flash must be either clear-sky or have a CTT greater than 273.15 K. A 6-km radius is used for LIS events.

1. The 10(6)-km criterion corresponds to the approximate distance from the center of the coarsest GLM(LIS) pixel, 14(8) km, to the pixel's corner, i.e, $7\sqrt{2}(4\sqrt{2})$ km. As mentioned above, 10 km also is the radius within which NLDN strokes are assigned to the same flash.

2. Since the ABI's CTT product has a temporal resolution of 10 minutes, the 30-minute criterion effectively means that all pixels within the 6 or 10-km radius must be clear or warm for eight time steps, four on either side of the event or flash. It was deemed necessary for the surrounding region to be warm not only before the flash (in order to let any previous storms in the area dissipate), but afterwards as well, since the ABI could be seeing radiation primarily from lower altitudes if a cold updraft is small compared to the pixel's size. 30 minutes was deemed a sufficient time for such an updraft to

grow and/or detrain (possibly forming an anvil cloud) to the point where it dominates enough of a pixel to make itself visible as a cold cloud. Since this could conceivably eliminate lightning-producing warm clouds that simply continue their convective ascent above the freezing level (Moore et al., 1960; Rossby, 1966), we will test the sensitivity of our results to the 30-minute criterion.

The locations of the resulting warm-flash candidates are plotted in Figure 5.1. Looking at the GLM-candidate locations in Figure 5.1a, the potential for artifacts is apparent, most notably via the horizontal streaks between 5°N and 30°N. Such features have been noted (e.g., Rudlosky et al., 2019) and ascribed to bright clouds over the boundaries between the independently tuned RTEPs, causing the threshold-to-noise ratio to decrease and, thus, leading to false events. While multiple software updates have attempted to mitigate this phenomenon (Bateman and Mach, 2020), Figure 5.1a suggests that there is additional work to be done in this regard. Furthermore, recent work suggests that GLM artifacts are more common in the daytime (Bateman and Mach, 2020); Peterson (2020) estimates that 40% of recorded events between 10:00 and 12:00 local time in 2018 were solar-related artifacts. While, as mentioned above, a blooming filter was implemented into the GLM software in 2019, the significant fraction of GLM candidates that occurred during the daytime (Figure 5.2) should be met with skepticism.

Figure 5.1b displays the location of the LIS candidates. While, as noted before, the LIS is essentially the same instrument as the GLM, there is not as conspicuous a signal of artifacts as there was with GLM. Indeed, Blakeslee et al. (2020) estimated that LIS has had a false alarm rate (FAR) of less than 5% over its first three years of operation–as compared to the GLM FAR, estimated to be in the range of 20-40% in 2018 (Koshak et al., 2018). Aside from being less prone to eclipses during the equinoxes (leading to solar glint), LIS would be expected to have a lower FAR because it only has one RTEP, thus having no subarray boundaries over which false events are likely to occur, as with GLM (K. Virts 2020, personal communication).

Figure 5.1c shows that the NLDN candidates are primarily outside CONUS: in Mexico, the Caribbean, and the Atlantic Ocean off the U.S. East Coast. While many of the reported instances of warm-cloud lightning occurred around the Caribbean (Pietrowski, 1960; Moore et al., 1960; Rossby, 1966), we note that the NLDN location accuracy suffers dramatically outside CONUS, where estimated errors are on the order of 10 km (Nag et al., 2014).

5.3.2 Cross-Validating Warm-Cloud Candidates

The number of candidates compared to the total number of flashes is extremely low in all three datasets (Table 1), especially considering the estimated GLM/LIS false event rates. Combined with the potential for significant location errors using NLDN outside CONUS, it is clear than an additional step is needed to robustly identify warm-cloud lightning. Here, we look to cross-validate flashes by looking for candidates that are within a certain spatiotemporal distance of each other. The fact that only NLDN flash data was available for
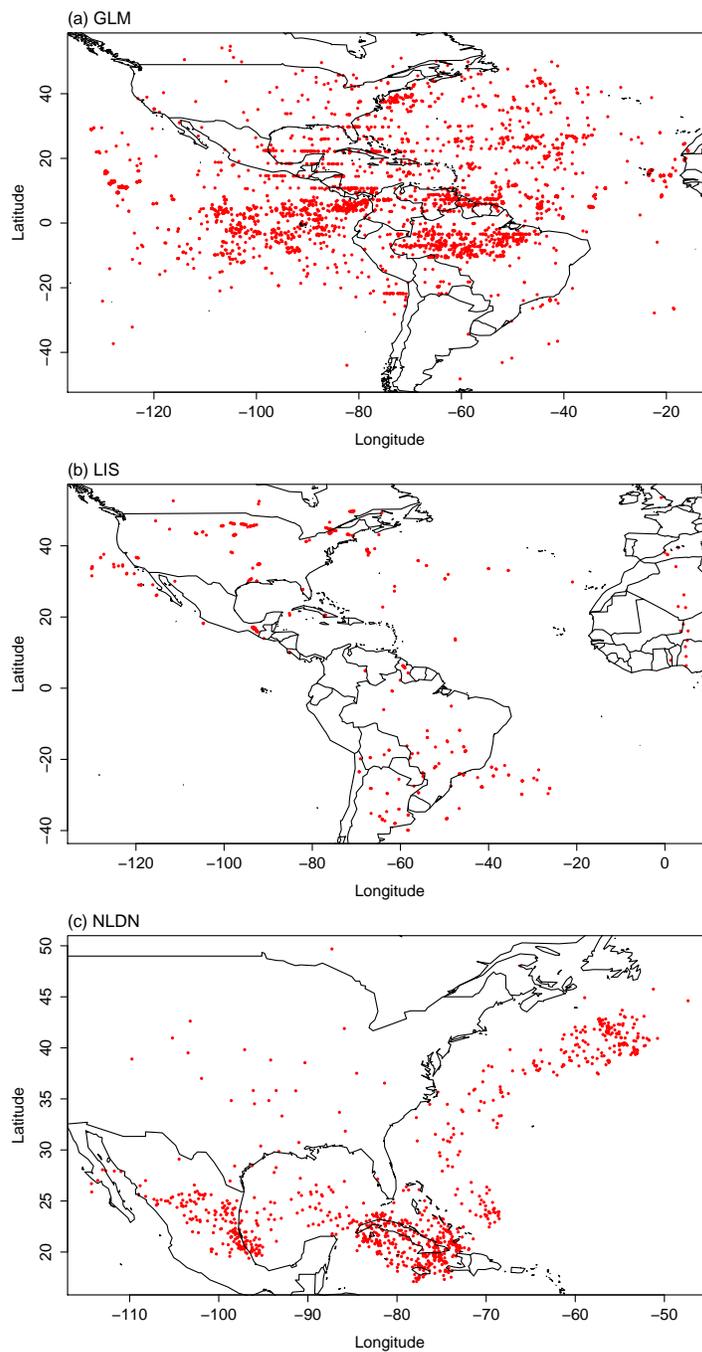
Figure 5.1: Warm-cloud flash-candidate locations.

this analysis prevented us from matching NLDN strokes and GLM/LIS groups, which are more comparable due to the differing flash classification methods. The criteria we use are 50
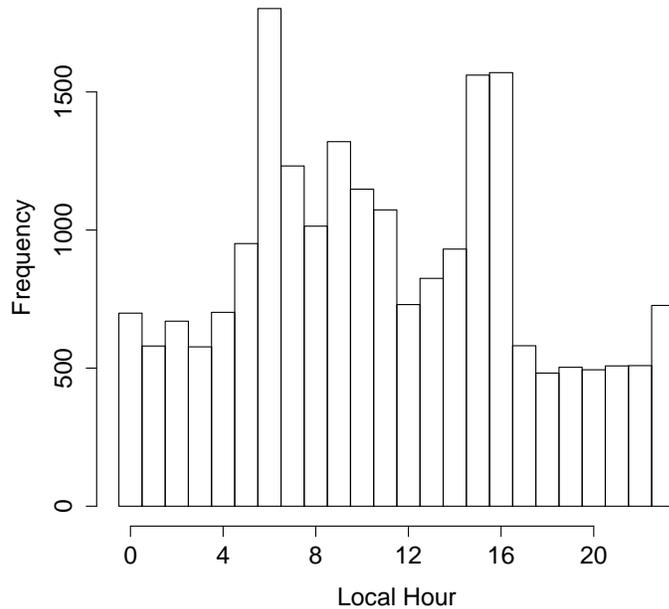
Figure 5.2: Histogram of the local hour of GLM warm-cloud flash candidates, where the local hour is estimated as the recorded UTC hour plus the largest integer smaller than $\frac{\theta_{\text{flash}}}{15°}$, where $\theta_{\text{flash}} \in (-180°, 180°)$ is the longitude of a flash.

km and 3 seconds. While this may seem to break with our conservative approach, (Bateman and Mach, 2020), it turns out that there are zero matching flashes comparing any two out of the three sets of candidates. Furthermore, this finding holds even if we only require one ABI time step before and after an NLDN flash, or each GLM/LIS event comprising a flash, to be completely clear/warm. It is on the basis of this that we proclaim no definitive evidence for warm-cloud lightning.

Table 5.1: Total number of flashes analyzed within GOES-EAST domain from July 27-August 31, 2019 and fraction of warm-cloud flash candidates.

Dataset	GLM	LIS	NLDN
Total number of flashes	46,798,182	68,827	6,044,156
Warm-cloud flash-candidate fraction	4.5×10^{-4}	5.4×10^{-3}	1.5×10^{-4}

5.4 Discussion and Conclusions

We note two comments regarding the cross-validation procedure.

1. Not all of the GLM candidates were able to be cross-validated. Given that NLDN's domain is effectively restricted to CONUS, and LIS is not in a geostationary orbit, while all NLDN and LIS candidates–which were constrained to be in the GOES-East domain–were able to compared with the GLM candidates, the converse is not true. GLM candidate flashes outside CONUS would have been unable to be validated if LIS was not overhead.

2. Attempting to validate GLM candidates with the entire dataset of NLDN flashes, that is, not just the warm-cloud candidates, resulted in some matches using the above spatial and temporal criteria. Visual inspection of the CTT maps confirmed that the matched NLDN flashes were located within cold clouds, lending support to the idea that the radiation associated with GLM events in a cold thunderstorm sometimes does not come out the (cold) cloud top but rather reflects off of lower clouds nearby (Peterson, 2019).

If warm-cloud lightning is indeed a real phenomenon, it is plausible that this analysis repeated over a longer time period would reveal instances of it. It is also possible that some of the flash candidates identified were in fact warm-cloud, and, for reasons unrelated to those listed previously (false alarms, location errors), they were simply not detected by the other instruments. For example, satellite-based sensors like GLM and LIS may be more likely to see small intracloud flashes, but they may be less likely to see strokes at lower levels in a cloud, when the associated light simply does not penetrate the cloud top (R. Holzworth 2019, personal communication). Further work could be conducted using other lightning datasets, for example, Earth Network's Total Lightning Network (ENTLN) (Heckman and Liu, 2010), the Worldwide Lightning Location Network (WWLLN) (Dowden et al., 2002; Lay et al., 2004), the Canadian Lightning Detection Network (CLDN) (Burrows et al., 2002), and Vaisala's Global Lightning Dataset (GLD360) (Demetriades et al., 2010).

Chapter 6

Conclusions

Microphysical processes, as their name implies, are often too small to observe with the naked eye, yet their effects manifest in macro-scale phenomena, such as clouds, storms, and the global atmospheric energy balance. Climate and weather modelers must capture these effects, which stem from myriad processes, including nucleation on aerosols, diffusional growth, accretion of supercooled water, collision and coalescence, sedimentation, drop breakup, and ice enhancement. Given their first-order impact on extreme precipitation and lightning, it is important to assess the degree to which our ability to simulate such weather phenomena depends on the chosen microphysical parameterization.

In this dissertation we have shown how microphysical processes shape the macro-scale phenomena around them, both locally and non-locally. In Chapter 2 we showed that changing the microphysics parameterization within a superparameterized climate model has only a small local effect on precipitation extremes. However, differences become apparent when allowing the microphysics to act on longer time scales, impacting the large-scale circulation, which then feeds back onto rain rates. In Chapter 3 we showed that this was a common theme in the midlatitudes, but, in the tropics, microphysics can have a large, immediate impact on the local vertical mass flux, which in turn affects local precipitation extremes.

In Chapter 4 we implemented lightning proxies into a superparameterized model for the first time and compared the performance of ice-based predictors against that of a large-scale one in CAPE × P. In general the former matches better with observations than does the latter. In addition, the choice of microphysics scheme has a first-order impact on the sign of change in flash rate with global warming in the midlatitudes and over oceans, with a 2-moment scheme predicting much more in the way of increases, consistent with CAPE × P, and the 1-moment scheme suggesting decreases. Finally, in Chapter 5 we conducted a systematic search for warm-cloud lightning using satellite observations of cloud-top temperature and three lightning datasets. While each individual dataset suggested a number of potential candidates, efforts to spatiotemporally match candidates across datasets were unsuccessful. Thus, in support of the idea that ice is necessary for thunderstorm electrification, we found no conclusive evidence for warm-cloud lightning in this analysis.

Bibliography

Allen, M. R. and Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. *Nature*, 419(6903):228–232.

Appleman, H. (1957). Comments on "initial electrification processes in thunderstorms". *Journal of Meteorology*, 14(1):89–89.

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066):303.

Barthe, C., Chong, M., Pinty, J.-P., Bovalo, C., and Escobar, J. (2012). Cells v1. 0: updated and parallelized version of an electrical scheme to simulate multiple electrified clouds and flashes over large domains. *Geoscientific Model Development*, 5(1):167–184.

Barthe, C., Deierling, W., and Barth, M. C. (2010). Estimation of total lightning from various storm parameters: A cloud-resolving model study. *Journal of Geophysical Research: Atmospheres*, 115(D24).

Bateman, M. and Mach, D. (2020). Preliminary detection efficiency and false alarm rate assessment of the Geostationary Lightning Mapper on the GOES-16 satellite. *Journal of Applied Remote Sensing*, 14(3):032406.

Benedict, J. J. and Randall, D. A. (2009). Structure of the Madden–Julian oscillation in the superparameterized CAM. *Journal of the Atmospheric Sciences*, 66(11):3277–3296.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 289–300.

Blakeslee, R. J. (2019). Non-quality controlled Lightning Imaging Sensor (LIS) on International Space Station (ISS) [July 27 to August 31, 2019].

Blakeslee, R. J., Lang, T. J., Koshak, W. J., Buechler, D., Gatlin, P., Mach, D. M., Stano, G. T., Virts, K. S., Walker, T. D., Cecil, D. J., et al. (2020). Three years of the Lightning Imaging Sensor onboard the International Space Station: Expanded global coverage and enhanced applications. *Journal of Geophysical Research: Atmospheres*.

Brooks, C. E. P. (1925). The distribution of thunderstorms over the globe. *Geophys. Mem. London*, 3(24):147–164.

Bryan, G. H. and Morrison, H. (2012). Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Monthly Weather Review*, 140(1):202–225.

Burrows, W. R., King, P., Lewis, P. J., Kochtubajda, B., Snyder, B., and Turcotte, V. (2002). Lightning occurrence patterns over Canada and adjacent United States from lightning detection network observations. *Atmosphere-ocean*, 40(1):59–80.

Carey, L. and Rutledge, S. (1996). A multiparameter radar case study of the microphysical and kinematic evolution of a lightning producing storm. *Meteorology and Atmospheric Physics*, 59(1-2):33–64.

Carey, L. D., Bain, A. L., and Matthee, R. (2014). Kinematic and microphysical control of lightning in multicell convection over Alabama during DC3. In *5th International Conference on Lightning meteorology*, pages 20–21.

Carey, L. D. and Rutledge, S. A. (2000). The relationship between precipitation and lightning in tropical island convection: A C-band polarimetric radar study. *Monthly Weather Review*, 128(8):2687–2710.

Cecil, D. J., Buechler, D. E., and Blakeslee, R. J. (2014). Gridded lightning climatology from TRMM-LIS and OTD: Dataset description. *Atmospheric Research*, 135:404–414.

Charn, A. B., Collins, W. D., Parishani, H., Risser, M. D., and O'Brien, T. A. (2020). Microphysical sensitivity of superparameterized precipitation extremes in the contiguous United States due to feedbacks on large-scale circulation. *Earth and Space Science*, 7(7).

Christian, H., Holmes, C., Bullock, J., Gaskell, W., Illingworth, A., and Latham, J. (1980). Airborne and ground-based studies of thunderstorms in the vicinity of Langmuir Laboratory. *Quarterly Journal of the Royal Meteorological Society*, 106(447):159–174.

Collins, W. D., Hackney, J. K., and Edwards, D. P. (2002). An updated parameterization for infrared emission and absorption by water vapor in the National Center for Atmospheric Research Community Atmosphere Model. *Journal of Geophysical Research: Atmospheres*, 107(D22):ACL–17.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., et al. (2011). The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, 137(654):1–28.

Dai, A. (2006). Precipitation characteristics in eighteen coupled climate models. *Journal of Climate*, 19(18):4605–4630.

Davis, M., Brook, M., Christian, H., Heikes, B. G., Orville, R. E., Park, C. G., Roble, R. G., and Vonnegut, B. (1983). Some scientific objectives of a satellite-borne lightning mapper. *Bulletin of the American Meteorological Society*, 64(2):114–119.

Demetriades, N. W., Murphy, M. J., and Cramer, J. A. (2010). Validation of Vaisala's global lightning dataset (GLD360) over the continental United States. In *Preprints, 29th Conf. on Hurricanes and Tropical Meteorology, Tucson, AZ, Amer. Meteor. Soc. D*, volume 16.

Dowden, R. L., Brundell, J. B., and Rodger, C. J. (2002). VLF lightning location by time of group arrival (TOGA) at multiple sites. *Journal of Atmospheric and Solar-Terrestrial Physics*, 64(7):817–830.

Elliott, E. J., Yu, S., Kooperman, G. J., Morrison, H., Wang, M., and Pritchard, M. S. (2016). Sensitivity of summer ensembles of fledgling superparameterized US mesoscale convective systems to cloud resolving model microphysics and grid configuration. *Journal of Advances in Modeling Earth Systems*, 8(2):634–649.

Fan, J., Liu, Y.-C., Xu, K.-M., North, K., Collis, S., Dong, X., Zhang, G. J., Chen, Q., Kollias, P., and Ghan, S. J. (2015). Improving representation of convective transport for scale-aware parameterization: 1. convection and cloud properties simulated with spectral bin and bulk microphysics. *Journal of Geophysical Research: Atmospheres*, 120(8):3485–3509.

Feng, Z., Leung, L. R., Houze Jr, R. A., Hagos, S., Hardin, J., Yang, Q., Han, B., and Fan, J. (2018). Structure and evolution of mesoscale convective systems: Sensitivity to cloud microphysics in convection-permitting simulations over the United States. *Journal of Advances in Modeling Earth Systems*, 10(7):1470–1494.

Field, P., Roberts, M., and Wilkinson, J. (2018). Simulated lightning in a convection permitting global model. *Journal of Geophysical Research: Atmospheres*, 123(17):9370–9377.

Fierro, A. O., Mansell, E. R., MacGorman, D. R., and Ziegler, C. L. (2013). The implementation of an explicit charging and discharge lightning scheme within the WRF-ARW model: Benchmark simulations of a continental squall line, a tropical cyclone, and a winter storm. *Monthly Weather Review*, 141(7):2390–2415.

Fildier, B., Parishani, H., and Collins, W. (2017). Simultaneous characterization of mesoscale and convective-scale tropical rainfall extremes and their dynamical and thermodynamic modes of change. *Journal of Advances in Modeling Earth Systems*, 9(5):2103–2119.

Fildier, B., Parishani, H., and Collins, W. (2018). Prognostic power of extreme rainfall scaling formulas across space and time scales. *Journal of Advances in Modeling Earth Systems*, 10(12):3252–3267.

Finney, D. L., Doherty, R., Wild, O., Huntrieser, H., Pumphrey, H. C., and Blyth, A. (2014). Using cloud ice flux to parametrise large-scale lightning. *Atmospheric Chemistry and Physics*, 14:12665–12682.

Finney, D. L., Doherty, R. M., Wild, O., Stevenson, D. S., MacKenzie, I. A., and Blyth, A. M. (2018). A projected decrease in lightning under climate change. *Nature Climate Change*, 8(3):210–213.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M. (2013). In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter Evaluation of Climate Models. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Foster, H. (1950). An unusual observation of lightning. *Bulletin of the American Meteorological Society*, 31(4):140–141.

Goodman, S., Mach, D., Koshak, W., and Blakeslee, R. (2012). GLM lightning cluster-filter algorithm. *Algorithm Theoretical Basis Document, Ver*, 3.

Goodman, S. J., Blakeslee, R. J., Koshak, W. J., Mach, D., Bailey, J., Buechler, D., Carey, L., Schultz, C., Bateman, M., McCaul Jr, E., et al. (2013). The GOES-R geostationary lightning mapper (GLM). *Atmospheric Research*, 125:34–49.

Grabowski, W. W., Morrison, H., Shima, S.-I., Abade, G. C., Dziekan, P., and Pawlowska, H. (2019). Modeling of cloud microphysics: Can we do better? *Bulletin of the American Meteorological Society*, 100(4):655–672.

Heckman, S. and Liu, C. (2010). The application of total lightning detection and cell tracking for severe weather prediction. In *Proc. Int. Conf. on Ground. and Earth & 4th Int. Conf. on Light. Phys. and Eff.(GROUND'2010 & 4th LPE)*, pages 234–240.

Heidinger, A. (2013). ABI cloud height algorithm theoretical basis document (ATBD). NOAA NESDIS Center for Satellite Applications and Research (STAR), ver. 3.0. Technical report, released 06/11/2013. Retrieved from http://www. star. nesdis. noaa. gov . . . .

Held, I. M. and Soden, B. J. (2006). Robust responses of the hydrological cycle to global warming. *Journal of Climate*, 19(21):5686–5699.

Higgins, R. W., Janowiak, J. E., and Yao, Y.-P. (1996). *A gridded hourly precipitation data base for the United States (1963-1993)*. US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service.

Hirota, N., Takayabu, Y. N., Watanabe, M., and Kimoto, M. (2011). Precipitation reproducibility over tropical oceans and its relationship to the double ITCZ problem in CMIP3 and MIROC5 climate models. *Journal of Climate*, 24(18):4859–4873.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of hydrometeorology*, 8(1):38–55.

Hurrell, J. W., Hack, J. J., Shea, D., Caron, J. M., and Rosinski, J. (2008). A new sea surface temperature and sea ice boundary dataset for the Community Atmosphere Model. *Journal of Climate*, 21(19):5145–5153.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J.-F., Large, W. G., Lawrence, D., Lindsay, K., et al. (2013). The Community Earth System Model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, 94(9):1339–1360.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, 113(D13).

Igel, A. L., Igel, M. R., and van den Heever, S. C. (2015). Make it a double? Sobering results from simulations using single-moment microphysics schemes. *Journal of the Atmospheric Sciences*, 72(2):910–925.

IPCC (2014). In *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, page 1132. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. Edited by Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White.

Jeevanjee, N. and Romps, D. M. (2018). Mean precipitation change from a deepening troposphere. *Proceedings of the National Academy of Sciences*, 115(45):11465–11470.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171.

Khairoutdinov, M., Randall, D., and 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.

Khairoutdinov, M. F., Krueger, S. K., Moeng, C.-H., Bogenschutz, P. A., and Randall, D. A. (2009). Large-eddy simulation of maritime deep tropical convection. *Journal of Advances in Modeling Earth Systems*, 1(4).

Khairoutdinov, M. F. and Randall, D. A. (2001). A cloud resolving model as a cloud parameterization in the NCAR Community Climate System Model: Preliminary results. *Geophysical Research Letters*, 28(18):3617–3620.

Khairoutdinov, M. F. and 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.

Kooperman, G. J., Pritchard, M. S., Burt, M. A., Branson, M. D., and Randall, D. A. (2016). 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(1):140–165.

Kooperman, G. J., Pritchard, M. S., O'Brien, T. A., and Timmermans, B. W. (2018). Rainfall from resolved rather than parameterized processes better represents the present-day and climate change response of moderate rates in the Community Atmosphere Model. *Journal of Advances in Modeling Earth Systems*, 10(4):971–988.

Koshak, W., Mach, D., Bateman, M., Armstrong, P., and Virts, K. (2018). Product performance guide for data users.

Kuettner, J. P., Sartor, J. D., and Levin, Z. (1981). Thunderstorm electrification—inductive or non-inductive? *Journal of the Atmospheric Sciences*, 38(11):2470–2484.

Lane-Smith, D. (1971). A warm thunderstorm. *Quarterly Journal of the Royal Meteorological Society*, 97(414):577–578.

Lay, E. H., Holzworth, R. H., Rodger, C. J., Thomas, J. N., Pinto Jr, O., and Dowden, R. L. (2004). WWLL global lightning detection system: Regional validation study in Brazil. *Geophysical Research Letters*, 31(3).

Lee, S. S. and Donner, L. J. (2011). Effects of cloud parameterization on radiation and precipitation: A comparison between single-moment microphysics and double-moment microphysics. *Terrestrial, Atmospheric & Oceanic Sciences*, 22(4).

Li, F., Rosa, D., Collins, W. D., and Wehner, M. F. (2012). "Super-parameterization": A better way to simulate regional extreme precipitation? *Journal of Advances in Modeling Earth Systems*, 4(2).

Marshall, J. S. and Palmer, W. M. K. (1948). The distribution of raindrops with size. *Journal of Meteorology*, 5(4):165–166.

McCaul Jr, E. W., Goodman, S. J., LaCasse, K. M., and Cecil, D. J. (2009). Forecasting lightning threat using cloud-resolving model simulations. *Weather and Forecasting*, 24(3):709–729.

McCrary, R. R., Randall, D. A., and Stan, C. (2014). Simulations of the West African monsoon with a superparameterized climate model. Part II: African easterly waves. *Journal of Climate*, 27(22):8323–8341.

Michalon, N., Nassif, A., Saouri, T., Royer, J., and Pontikis, C. (1999). Contribution to the climatological study of lightning. *Geophysical Research Letters*, 26(20):3097–3100.

Milbrandt, J. and Yau, M. (2005). A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *Journal of the Atmospheric Sciences*, 62(9):3051–3064.

Moore, C., Vonnegut, B., Stein, B., and Survilas, H. (1960). Observations of electrification and lightning in warm clouds. *Journal of Geophysical Research*, 65(7):1907–1910.

Morrison, H., Curry, J., and Khvorostyanov, V. (2005). A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description. *Journal of the Atmospheric Sciences*, 62(6):1665–1677.

Mostajabi, A., Finney, D. L., Rubinstein, M., and Rachidi, F. (2019). Nowcasting lightning occurrence from commonly available meteorological parameters using machine learning techniques. *npj Climate and Atmospheric Science*, 2(1):1–15.

Nag, A., Murphy, M. J., Cummins, K. L., Pifer, A. E., and Cramer, J. A. (2014). Recent evolution of the U.S. National Lightning Detection Network. In *23rd International Lightning Detection Conference & 5th International Lightning Meteorology Conference.*

Neale, R., Chen, C., and Gettelman, A. (2010a). Description of the NCAR Community Atmosphere Model (CAM 4.0) NCAR. *TN-485+ STR http://www. cesm. ucar. edu/models/ccsm4. 0/cam/docs/description/cam4_desc. pdf.*

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Williamson, D. L., Conley, A. J., Garcia, R., Kinnison, D., Lamarque, J.-F., et al. (2010b). Description of the NCAR Community Atmosphere Model (CAM 5.0). *NCAR Tech. Note NCAR/TN-486+ STR*, 1(1):1–12.

Neale, R. B., Richter, J. H., and Jochum, M. (2008). The impact of convection on ENSO: From a delayed oscillator to a series of events. *Journal of Climate*, 21(22):5904–5924.

Nesbitt, S. W., Cifelli, R., and Rutledge, S. A. (2006). Storm morphology and rainfall characteristics of TRMM precipitation features. *Monthly Weather Review*, 134(10):2702–2721.

O'Brien, T. A., Collins, W. D., Kashinath, K., Rübel, O., Byna, S., Gu, J., Krishnan, H., and Ullrich, P. A. (2016). Resolution dependence of precipitation statistical fidelity in hindcast simulations. *Journal of Advances in Modeling Earth Systems*, 8(2):976–990.

O'Gorman, P. A. and Schneider, T. (2009). Scaling of precipitation extremes over a wide range of climates simulated with an idealized GCM. *Journal of Climate*, 22(21):5676–5685.

Ooyama, K. V. (2001). A dynamic and thermodynamic foundation for modeling the moist atmosphere with parameterized microphysics. *Journal of the Atmospheric Sciences*, 58(15):2073–2102.

Orville, R. E. and Henderson, R. W. (1986). Global distribution of midnight lightning: September 1977 to August 1978. *Monthly Weather Review*, 114(12):2640–2653.

Pall, P., Aina, T., Stone, D. A., Stott, P. A., Nozawa, T., Hilberts, A. G., Lohmann, D., and Allen, M. R. (2011). Anthropogenic greenhouse gas contribution to flood risk in England and Wales in Autumn 2000. *Nature*, 470(7334):382.

Pall, P., Patricola, C. M., Wehner, M. F., Stone, D. A., Paciorek, C. J., and Collins, W. D. (2017). Diagnosing conditional anthropogenic contributions to heavy Colorado rainfall in September 2013. *Weather and Climate Extremes*, 17:1–6.

Parishani, H., Pritchard, M. S., Bretherton, C. S., Terai, C. R., Wyant, M. C., Khairoutdinov, M., and Singh, B. (2018). Insensitivity of the cloud response to surface warming under radical changes to boundary layer turbulence and cloud microphysics: Results from the ultraparameterized CAM. *Journal of Advances in Modeling Earth Systems*, 10(12):3139–3158.

Parishani, H., Pritchard, M. S., Bretherton, C. S., Wyant, M. C., and 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.

Patricola, C. M. and Wehner, M. F. (2018). Anthropogenic influences on major tropical cyclone events. *Nature*, 563(7731):339–346.

Petersen, W. A., Rutledge, S. A., and Orville, R. E. (1996). Cloud-to-ground lightning observations from TOGA COARE: Selected results and lightning location algorithms. *Monthly Weather Review*, 124(4):602–620.

Peterson, M. (2019). Research applications for the Geostationary Lightning Mapper operational lightning flash data product. *Journal of Geophysical Research: Atmospheres*, 124(17-18):10205–10231.

Peterson, M. (2020). Removing solar artifacts from Geostationary Lightning Mapper data to document lightning extremes. *Journal of Applied Remote Sensing*, 14(3):032402.

Phillips, V. T. and Donner, L. J. (2006). Cloud microphysics, radiation and vertical velocities in two-and three-dimensional simulations of deep convection. *Quarterly Journal of the Royal Meteorological Society*, 132(621C):3011–3033.

Pickands III, J. (1971). The two-dimensional Poisson process and extremal processes. *Journal of applied Probability*, 8(4):745–756.

Pickands III, J. et al. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, 3(1):119–131.

Pietrowski, E. (1960). An observation of lightning in warm clouds. *Journal of Meteorology*, 17(5):562–563.

Price, C. and Rind, D. (1992). A simple lightning parameterization for calculating global lightning distributions. *Journal of Geophysical Research: Atmospheres*, 97(D9):9919–9933.

Price, C. and Rind, D. (1994). Possible implications of global climate change on global lightning distributions and frequencies. *Journal of Geophysical Research: Atmospheres*, 99(D5):10823–10831.

Rasmussen, R. and Liu, C. (2017). High resolution WRF simulations of the current and future climate of North America, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory.

Rauscher, S. A., O'Brien, T. A., Piani, C., Coppola, E., Giorgi, F., Collins, W. D., and Lawston, P. M. (2016). A multimodel intercomparison of resolution effects on precipitation: Simulations and theory. *Climate Dynamics*, 47(7-8):2205–2218.

Redelsperger, J.-L., Brown, P., Guichard, F., How, C., Kawasima, M., Lang, S., Montmerle, T., Nakamura, K., Saito, K., Seman, C., et al. (2000). A GCSS model intercomparison for a tropical squall line observed during TOGA-COARE. I: Cloud-resolving models. *Quarterly Journal of the Royal Meteorological Society*, 126(564):823–863.

Reynolds, S., Brook, M., and Gourley, M. F. (1957). Thunderstorm charge separation. *Journal of Meteorology*, 14(5):426–436.

Risser, M. D., Paciorek, C. J., Wehner, M. F., O'Brien, T. A., and Collins, W. D. (2019). A probabilistic gridded product for daily precipitation extremes over the United States. *Climate Dynamics*, 53(5-6):2517–2538.

Risser, M. D. and Wehner, M. F. (2017). Attributable human-induced changes in the likelihood and magnitude of the observed extreme precipitation during Hurricane Harvey. *Geophysical Research Letters*, 44(24):12457–12464.

Romps, D. M. (2010). A direct measure of entrainment. *Journal of the Atmospheric Sciences*, 67(6):1908–1927.

Romps, D. M. (2011). Response of tropical precipitation to global warming. *Journal of the Atmospheric Sciences*, 68(1):123–138.

Romps, D. M. (2014). An analytical model for tropical relative humidity. *Journal of Climate*, 27(19):7432–7449.

Romps, D. M. (2016). Clausius–Clapeyron scaling of CAPE from analytical solutions to RCE. *Journal of the Atmospheric Sciences*, 73(9):3719–3737.

Romps, D. M. (2019). Evaluating the future of lightning in cloud-resolving models. *Geophysical Research Letters*, 46(24):14863–14871.

Romps, D. M. (2020). Climate sensitivity and the direct effect of carbon dioxide in a limited-area cloud-resolving model. *Journal of Climate*, 33(9):3413–3429.

Romps, D. M., Charn, A. B., Holzworth, R. H., Lawrence, W. E., Molinari, J., and Vollaro, D. (2018). CAPE times P explains lightning over land but not the land-ocean contrast. *Geophysical Research Letters*, 45(22):12623–12630.

Romps, D. M. and Kuang, Z. (2010). Do undiluted convective plumes exist in the upper tropical troposphere? *Journal of the Atmospheric Sciences*, 67(2):468–484.

Romps, D. M., Seeley, J. T., Vollaro, D., and Molinari, J. (2014). Projected increase in lightning strikes in the United States due to global warming. *Science*, 346(6211):851–854.

Rossby, S. A. (1966). Sferics from lightning within a warm cloud. *Journal of Geophysical Research*, 71(16):3807–3809.

Rowe, A. K., Rutledge, S. A., and Lang, T. J. (2012). Investigation of microphysical processes occurring in organized convection during NAME. *Monthly Weather Review*, 140(7):2168–2187.

Rudlosky, S. D., Goodman, S. J., Virts, K. S., and Bruning, E. C. (2019). Initial geostationary lightning mapper observations. *Geophysical Research Letters*, 46(2):1097–1104.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8):1015–1058.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H.-y., Iredell, M., et al. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, 27(6):2185–2208.

Schumacher, R. S. and Johnson, R. H. (2006). Characteristics of US extreme rain events during 1999–2003. *Weather and Forecasting*, 21(1):69–85.

Seeley, J. T., Jeevanjee, N., and Romps, D. M. (2019). FAT or FiTT: Are anvil clouds or the tropopause temperature invariant? *Geophysical Research Letters*, 46(3):1842–1850.

Seeley, J. T. and Romps, D. M. (2015). Why does tropical convective available potential energy (CAPE) increase with warming? *Geophysical Research Letters*, 42(23):10429–10437.

Shi, Z., Hui-Qiang, T., and Yong-Bo, T. (2016). Effects of the inductive charging on the electrification and lightning discharges in thunderstorms. *TAO: Terrestrial, Atmospheric and Oceanic Sciences*, 27(2):2.

Shima, S.-i., Kusano, K., Kawano, A., Sugiyama, T., and Kawahara, S. (2009). The super-droplet method for the numerical simulation of clouds and precipitation: A particle-based and probabilistic microphysics model coupled with a non-hydrostatic model. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 135(642):1307–1320.

Singh, M. S. and O'Gorman, P. A. (2013). Influence of entrainment on the thermal stratification in simulations of radiative-convective equilibrium. *Geophysical Research Letters*, 40(16):4398–4403.

Singh, M. S. and O'Gorman, P. A. (2014). Influence of microphysics on the scaling of precipitation extremes with temperature. *Geophysical Research Letters*, 41(16):6037–6044.

Singh, M. S. and O'Gorman, P. A. (2012). Upward shift of the atmospheric general circulation under global warming: Theory and simulations. *Journal of Climate*, 25(23):8259–8276.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., et al. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, 145(724):2876–2908.

Takahashi, T. (1978). Riming electrification as a charge generation mechanism in thunderstorms. *Journal of the Atmospheric Sciences*, 35(8):1536–1548.

Tao, W.-K. and Chern, J.-D. (2017). The impact of simulated mesoscale convective systems on global precipitation: A multiscale modeling study. *Journal of Advances in Modeling Earth Systems*, 9(2):790–809.

Telford, J. (1955). A new aspect of coalescence theory. *Journal of Meteorology*, 12(5):436–444.

Tippett, M. K., Lepore, C., Koshak, W. J., Chronis, T., and Vant-Hull, B. (2019). Performance of a simple reanalysis proxy for US cloud-to-ground lightning. *International Journal of Climatology*, 39(10):3932–3946.

Van Weverberg, K., Vogelmann, A. M., Morrison, H., and Milbrandt, J. A. (2012). Sensitivity of idealized squall-line simulations to the level of complexity used in two-moment bulk microphysics schemes. *Monthly Weather Review*, 140(6):1883–1907.

Vaughan Jr, O. H. and Boeck, W. L. (1998). Space shuttle video images: An example of warm cloud lightning.

Verlinde, J. and Cotton, W. R. (1993). Fitting microphysical observations of nonsteady convective clouds to a numerical model: An application of the adjoint technique of data assimilation to a kinematic model. *Monthly Weather Review*, 121(10):2776–2793.

Von Mises, R. (1954). La distribution de la plus grande de n valeurs. in (ed.), selected papers (vol. ii, pp. 271-294). providence, ri. *American Mathematical Society*.

Wilcox, E. M. and Donner, L. J. (2007). The frequency of extreme rain events in satellite rain-rate estimates and an atmospheric general circulation model. *Journal of Climate*, 20(1):53–69.

Wilhelmson, R. (1974). The life cycle of a thunderstorm in three dimensions. *Journal of the Atmospheric Sciences*, 31(6):1629–1651.

Wilks, D. S. (2016). "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it. *Bulletin of the American Meteorological Society*, 97(12):2263–2273.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9(1):60–62.

Woelfle, M., Yu, S., Bretherton, C., and Pritchard, M. (2018). Sensitivity of coupled tropical pacific model biases to convective parameterization in CESM1. *Journal of Advances in Modeling Earth Systems*, 10(1):126–144.

Zeng, X., Tao, W.-K., Lang, S., Hou, A. Y., Zhang, M., and Simpson, J. (2008). On the sensitivity of atmospheric ensembles to cloud microphysics in long-term cloud-resolving model simulations. *Journal of the Meteorological Society of Japan. Ser. II*, 86:45–65.

Zhang, G. J. and McFarlane, N. A. (1995). Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model. *Atmosphere-ocean*, 33(3):407–446.

Zhang, L., Min, J., Zhuang, X., and Schumacher, R. S. (2019). General features of extreme rainfall events produced by MCSs over east China during 2016–17. *Monthly Weather Review*, 147(7):2693–2714.

Appendix A

Appendix to Chapters 2 and 3

A.1   Controlling the False Discovery Rate

Once α_{FDR}, the level at which it is desired to limit the false discovery rate (Section 2.3.3) is chosen, the procedure for rejecting null hypotheses is as follows. For each comparison between two datasets in a given season, and given N finite p-values p_i with $i = 1, ..., N$, sort the p_i into ascending order. Note that N is bounded above in Chapter 2(3) by $N \leq 175(13,824)$, the number of GCM grid cells covering CONUS (the world) given the resolution used for the GCM spatial grid. Using standard statistical notation, these p-values are now denoted with parenthetical subscripts, such that $p_{(1)} \leq p_{(2)} \leq ... \leq p_{(N)}$. Local null hypotheses are now rejected if their respective p-values are no larger than p_{FDR}:

$$p_{\mathrm{FDR}} = \max_{i=1,...,N} \left[p_{(i)} : p_{(i)} \leq (i/N)\alpha_{\mathrm{FDR}} \right] \tag{A.1}$$

Thus, it can be seen that this method, known as the Benjamini-Hochberg procedure, requires even smaller p-values to reject a local null hypothesis than would be needed when examining individual grid points in isolation.

Appendix B

Supplementary Figures for Chapter 3
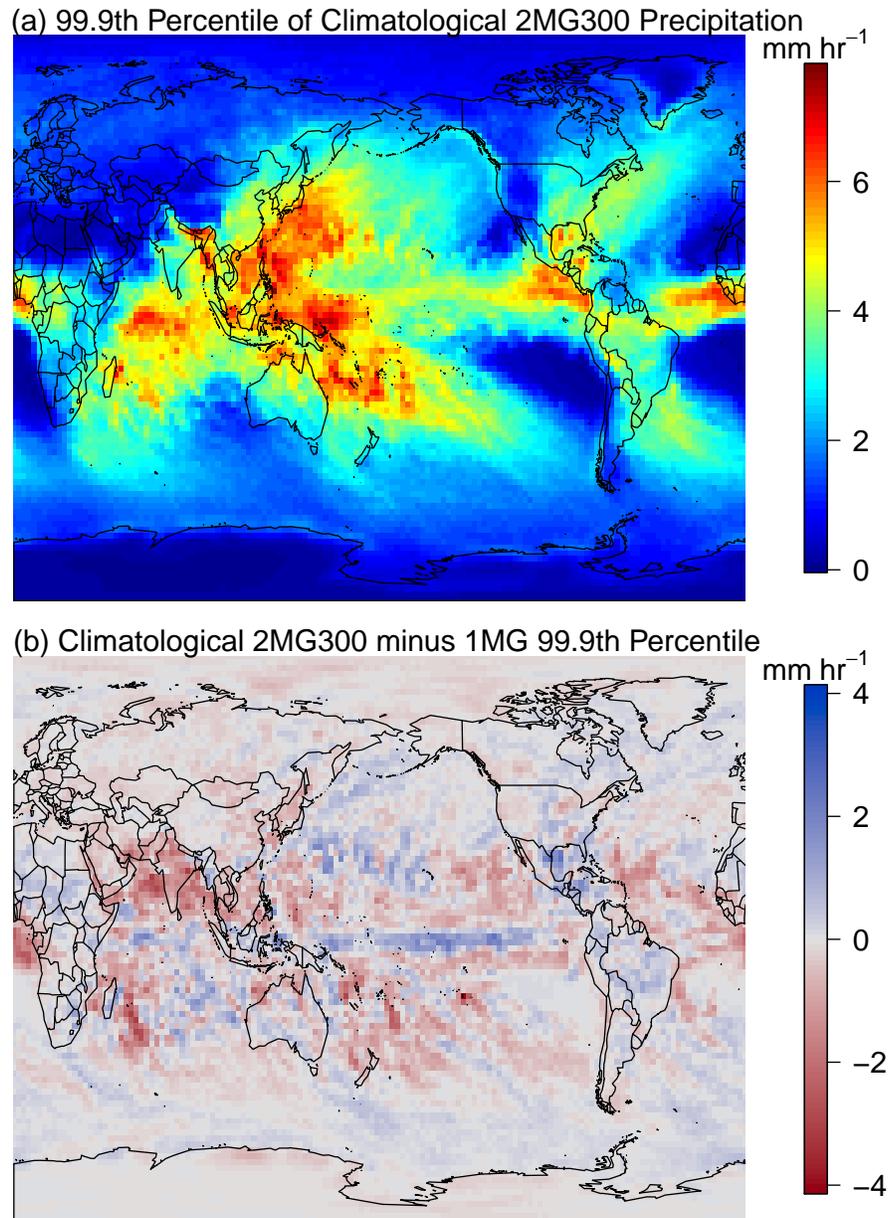
Figure B.1: (a) Annual 99.9th percentile of climatological 2MG300 precipitation. (b) Annual 99.9th percentile of climatological 1MG precipitation subtracted from that of 2MG300.

(a) ILIAD $\Delta \overline{\omega_{500}}$

(b) Climatological $\Delta \overline{\omega_{500}}$
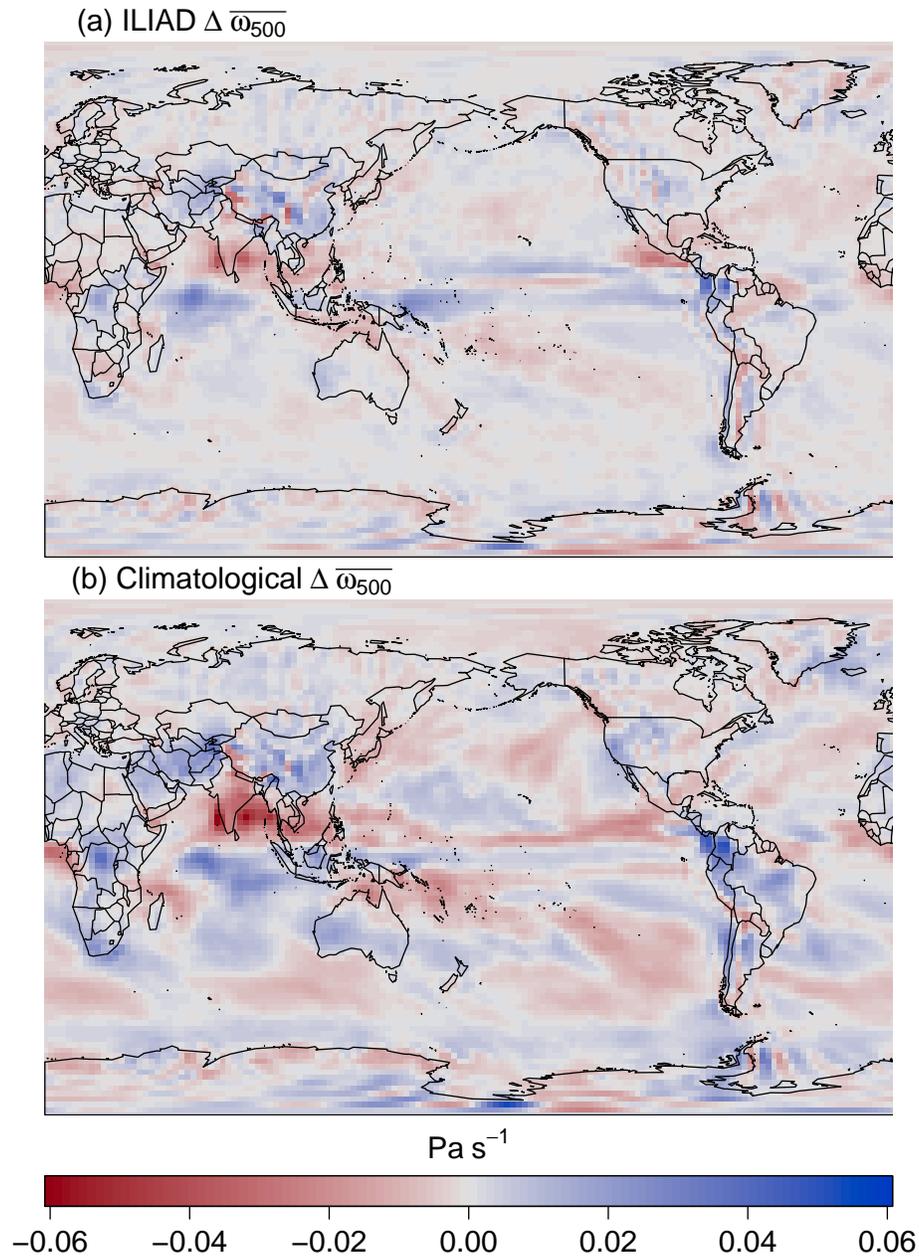
Figure B.2: Mean (a) ILIAD and (b) climatological 2MG300 500-hPa vertical velocity subtracted from that of 1MG.
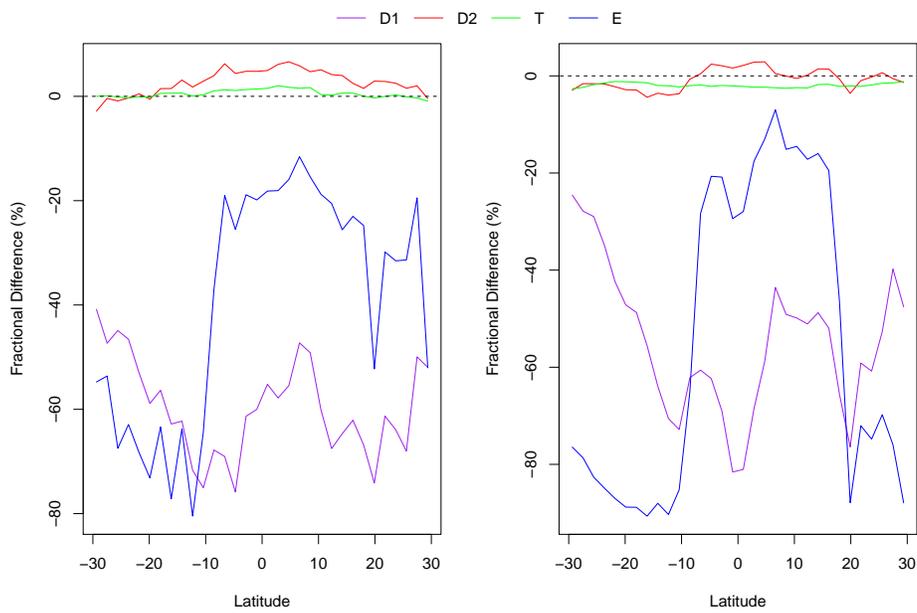
Figure B.3: As in Figure 3.5, but with medians calculated for all extremes in grid cells with statistically significantly different (ILIAD) precipitation distributions, that is, regardless of whether the extreme from the other microphysics case is denoted as extreme.
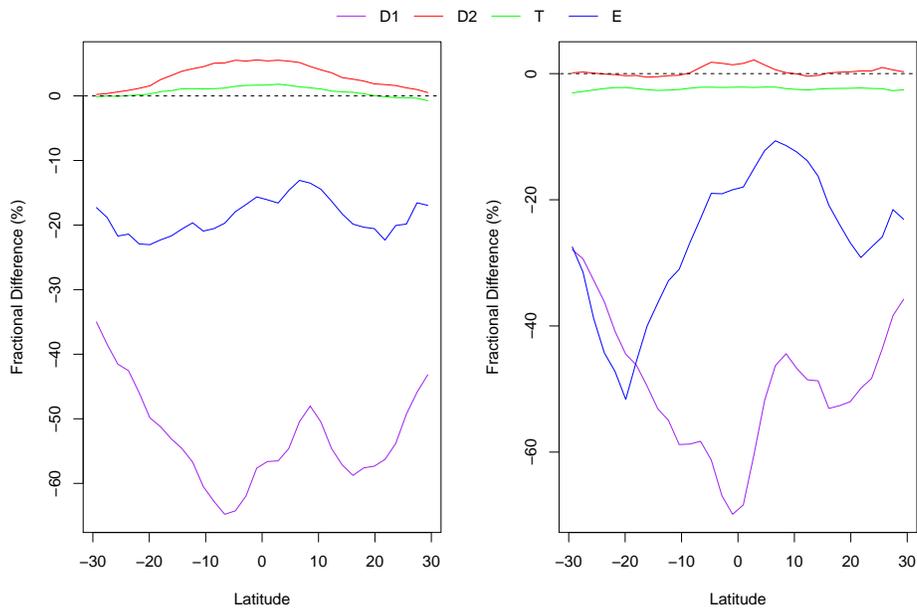
Figure B.4: As in Figure B.3, but for all grid cells between 30°S and 30°N, that is, even when the ILIAD extremes are not statistically significantly different between 1MG and 2MG300.
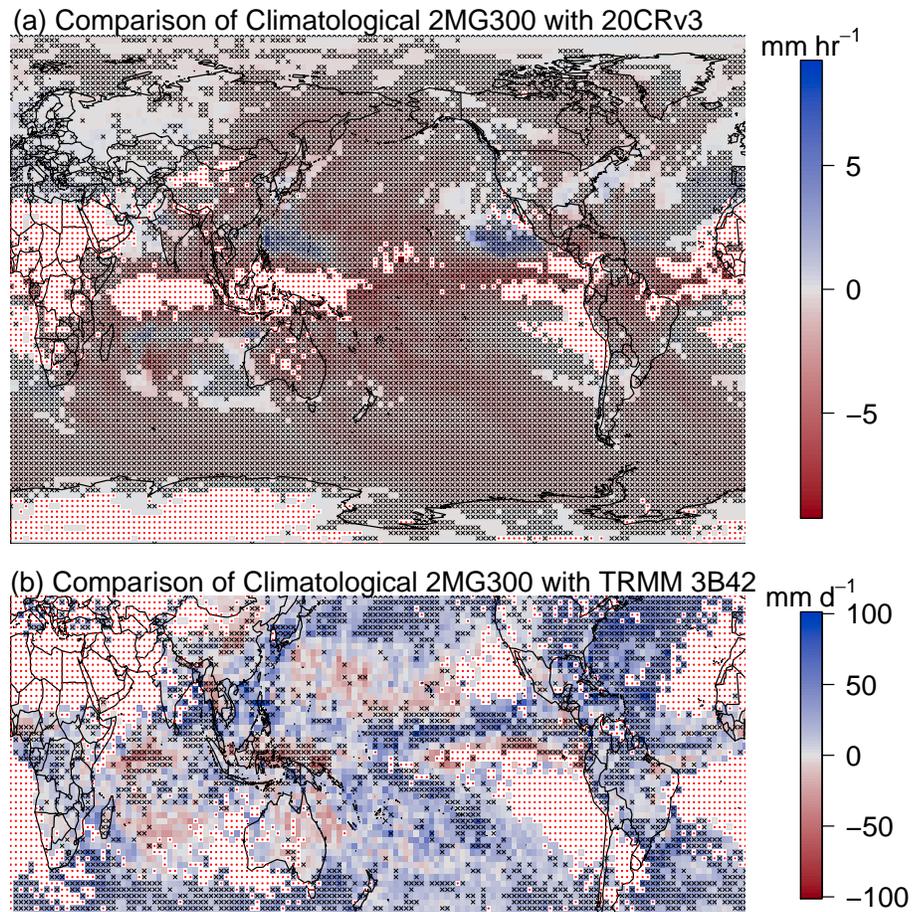
Figure B.5: As in Figure 3.8, but observations are compared to extremes from the climatological 2MG300 (Table 1) experiment.

Appendix C

Supplementary Figures for Chapter 4
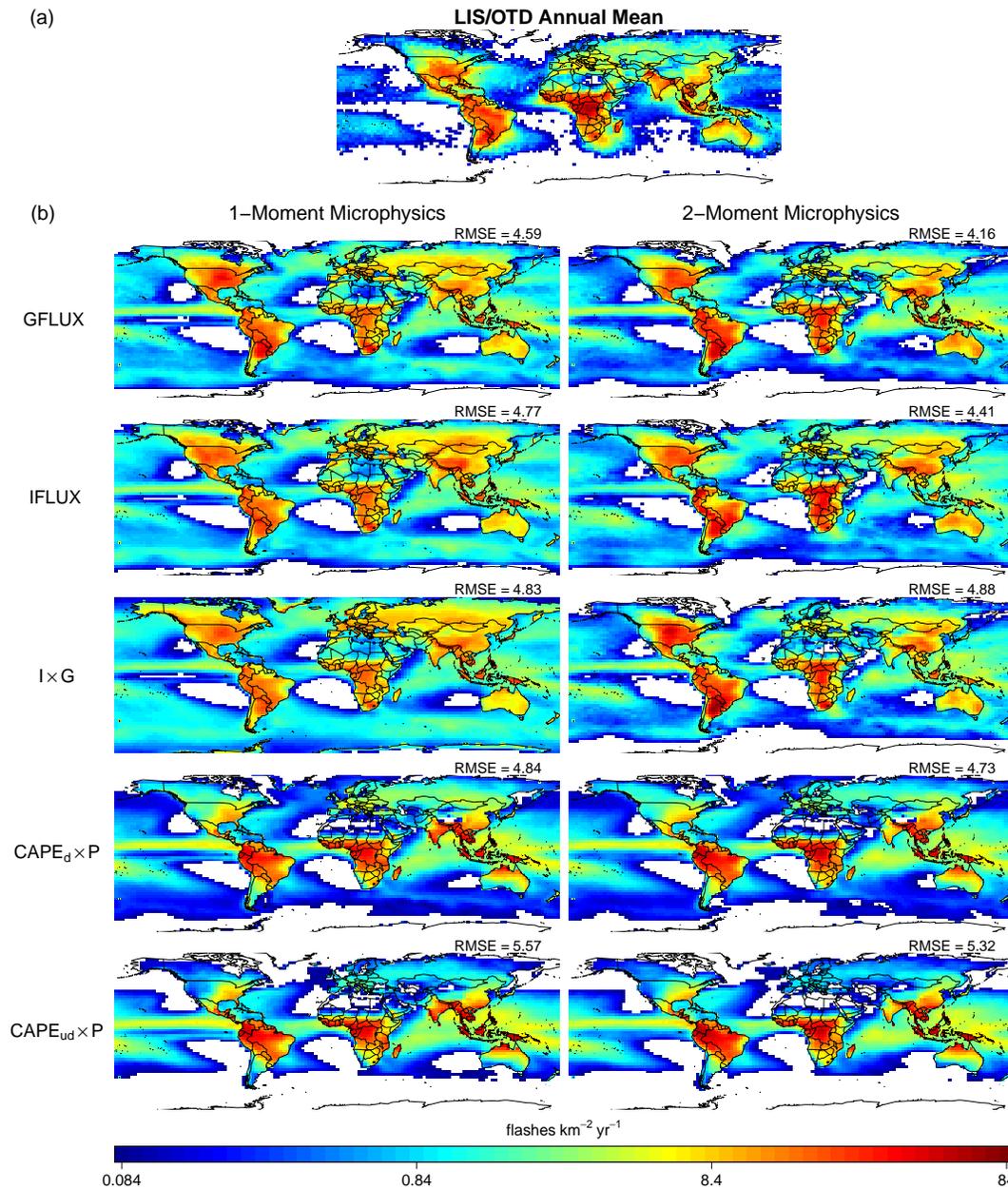
Figure C.1: (a) Annual mean flash rate of LIS/OTD observations and (b) the five lightning parameterizations generated by SPCAM. The proxies were interpolated from the large-scale SPCAM resolution ($1.9°$ latitude $\times 2.5°$ longitude) to the LIS/OTD grid ($2.5° \times 2.5°$). White regions correspond either to numerical values that are below the minimum of the scale (0.084 flashes km^{-2} yr^{-1}) or, in the case of LIS/OTD, possibly to grid cells with no observations.
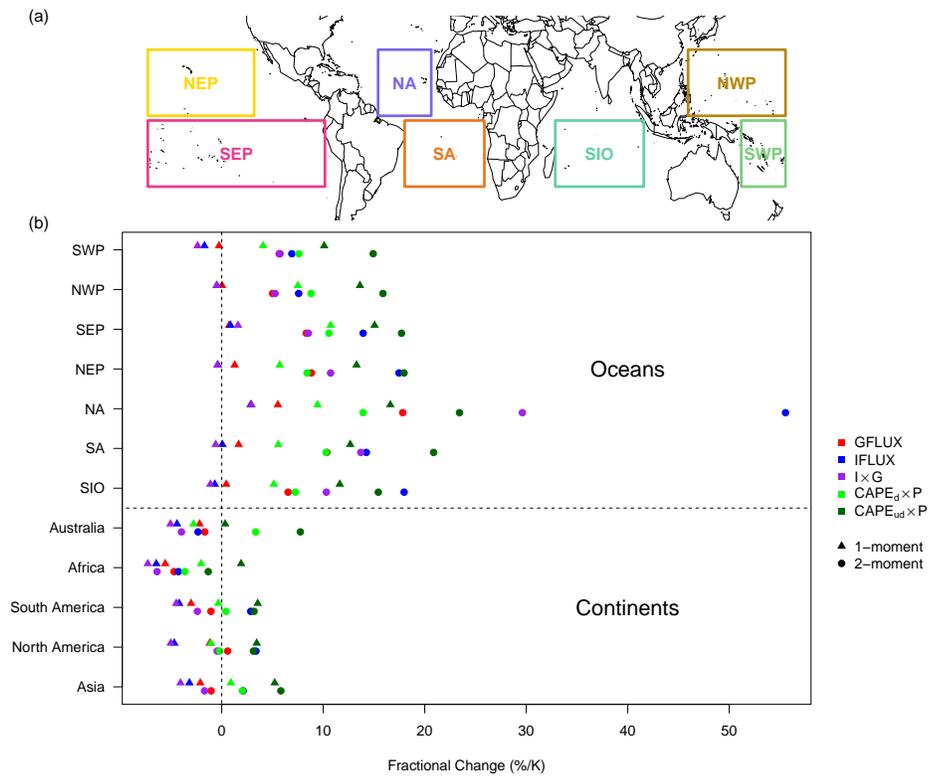
Figure C.2: (a) Demarcation of tropical ocean basins analyzed. (b) Fractional change per degree of local warming at the lowest model level in tropical flash rates from the control climate to a simulation with +4 K SST warming. Tropical continental regions for South America, North America, and Asia are areas between 30°S and 30°N. The entireties of Australia and Africa are considered tropical.
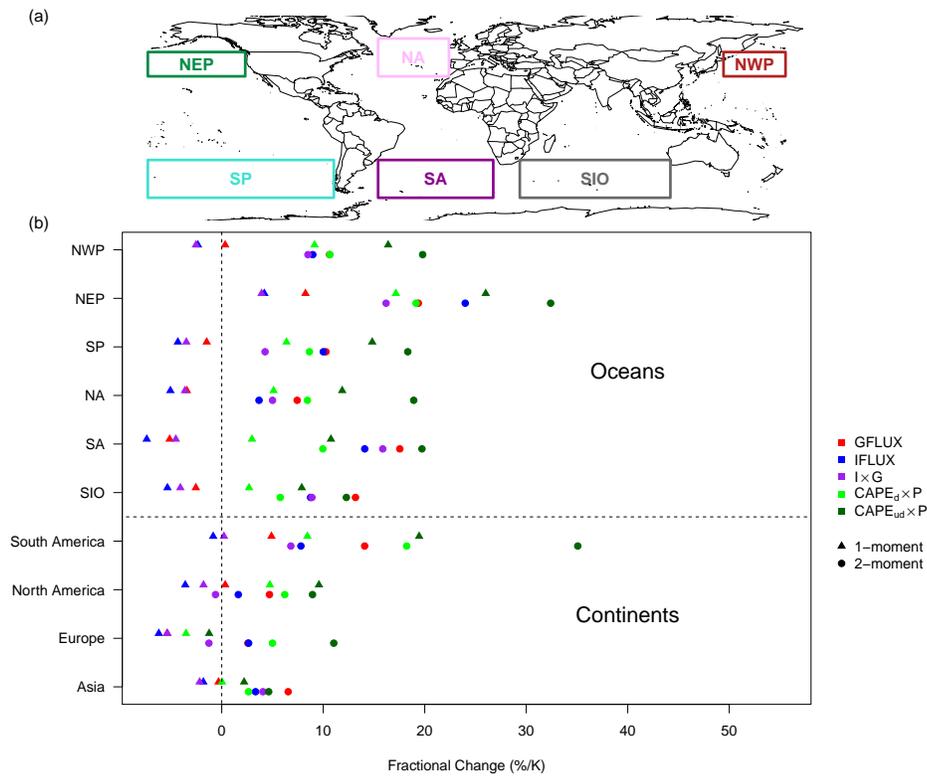
Figure C.3: (a) Demarcation of midlatitude ocean basins analyzed. (b) Fractional change per degree of local warming at the lowest model level in midlatitude flash rates from the control climate to a simulation with +4 K SST warming. Midlatitude continental regions for North America, Europe, and Asia are areas between 30°N and 60°N and for South America areas between 30°S and 60°S. The x-axis matches that in Figure C.2 for ease of comparison.
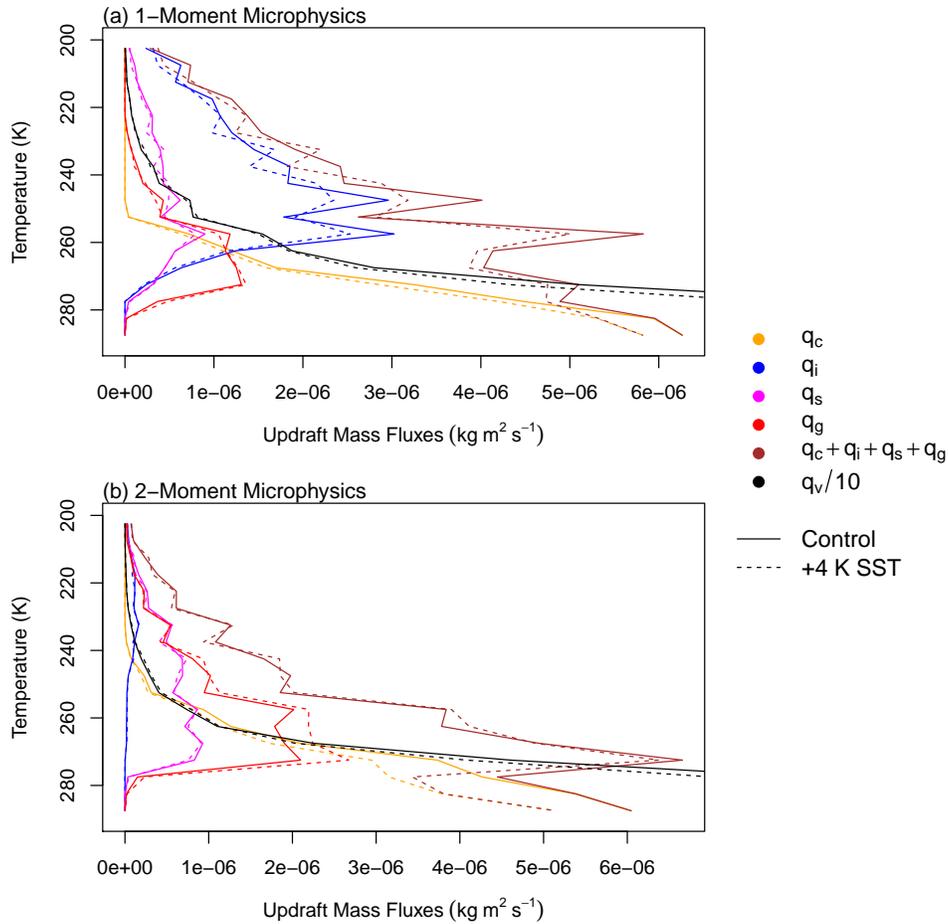
Figure C.4: (a) 1MOM, (b) 2MOM updraft mass fluxes, plotted as tropics-wide (30°S-30°N) means of $\rho q_x w \mathcal{H}(w - 1 \text{ m/s})$ in bins of width 5 K, where q_x denotes mixing ratios of cloud water (q_c), cloud ice (q_i), snow (q_s), graupel (q_g), or vapor (q_v). As in Figure 4.4, vapor fluxes are divided by 10, that is, actual values are 10 times larger, for graphical clarity.