Subgrid Variations of the Cloud Water and Droplet Number Concentration Over Tropical Ocean: Satellite Observations and Implications for Warm Rain Simulation in Climate Models

Zhibo Zhang[1,2*], Hua Song[2], Po-Lun Ma[3], Vincent E. Larson[4], Minghuai Wang[5], Xiquan Dong[6], Jianwu Wang[7]

1. Physics Department, University of Maryland Baltimore County (UMBC), Baltimore, MD, USA
2. Joint Center for Earth Systems Technology, UMBC, Baltimore, MD, USA
3. Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA, USA
4. Department of Mathematical Sciences, University of Wisconsin—Milwaukee, Milwaukee, WI, USA
5. Institute for Climate and Global Change Research & School of Atmospheric Sciences, Nanjing University, Nanjing, China
6. Department of Hydrology & Atmospheric Sciences, University of Arizona, Tucson, AZ, USA
7. Department of Information Systems, UMBC, Baltimore, MD, USA

*Corresponding Author:
Dr. Zhibo Zhang
Email: Zhibo.Zhang@umbc.edu
Phone: +1 (410) 455 6315

Revision for *ACP*

Abstract:

One of the challenges in representing warm rain processes in global climate models (GCM) is related to the representation of the subgrid variability of cloud properties, such as cloud water and cloud droplet number concertation (CDNC), and the effect thereof on individual precipitation processes such as autoconversion. This effect is conventionally treated by multiplying the resolved-scale warm ran process rates by an enhancement factor (E_q) which is derived from integrating over an assumed subgrid cloud water distribution. The assumed subgrid cloud distribution remain highly uncertain. In this study, we derive the subgrid variations of liquid-phase cloud properties over the tropical ocean using the satellite remote sensing products from Moderate Resolution Imaging Spectroradiometer (MODIS) and investigate the corresponding enhancement factors for the GCM parameterization of autoconversion rate. We find that the conventional approach of using only subgrid variability of cloud water is insufficient, and that the subgrid variability of CDNC, as well as the correlation between the two, are also important for the correctly simulating the autoconversion process in GCMs. Using the MODIS data which has the near-global data coverage, we find that E_q shows a strong dependence on cloud regimes, due to the fact that the subgrid variability of cloud water and CDNC is regime-dependent. Our analysis shows a significant increase of E_q from the stratocumulus (Sc) to cumulus (Cu) regions. Furthermore, the enhancement factor E_N due to the subgrid variation of CDNC is derived from satellite observation for the first time, and results reveal several regions downwind of biomass burning aerosols (e.g., Gulf of Guinea, East Coast of South Africa), air pollution (i.e., Eastern China Sea), and active volcanos (e.g., Kilauea Hawaii and Ambae Vanuatu), where the E_N is comparable, or even larger than E_q, suggesting an important role of aerosol in influencing the E_N. MODIS observations suggest that the subgrid variations of cloud liquid water path (LWP) and CDNC are generally positively correlated. As a result, the combined enhancement factor, including the effect of LWP and CDNC correlation, is significantly smaller than the simple product of $E_q \cdot E_N$. Given the importance of warm rain processes in understanding the Earth system dynamics and water cycle, we conclude that more observational studies are needed to provide a better constraint on the warm rain processes in GCMs.

1. Introduction

Marine boundary layer (MBL) clouds are a strong modulator of Earth's radiative energy budget (Klein and Hartmann, 1993; Trenberth et al., 2009). They can interact with other components of the climate system, such as aerosols and precipitations, in various ways. The feedback of MBL clouds to climate change remains one of the largest uncertainties in our understanding of the climate sensitivity (Bony and Dufresne, 2005; Soden and Held, 2006). Despite their importance in the climate system, simulating MBL clouds in general circulations models (GCM) has proved to be extremely challenging. A main difficulty is rooted in the fact the typical grid size of GCM (~100km) is much larger than the spatial scale of many cloud microphysical processes, and as a result these subgrid scale processes, as well as the subgrid cloud variations, have to be highly simplified and then parameterized as functions of resolved, grid-level variables.

Of particular interest in this study is the warm rain processes in MBL clouds, which have fundamental impacts on the cloud water budget and lifetime. Although in reality it is highly complicated and involves multiple factors, warm rain formation in GCMs is usually parameterized as simple functions of only key cloud parameters. For example, the drizzle in MBL cloud is initialized by the so-called autoconversion process in which the collision-coalescence of cloud droplets gives birth to large drizzle drops (Pruppacher and Klett, 1997). In GCMs, for the sake of efficiency, this process is usually parameterized as a power function of liquid water content (LWC or symbol q_c) and cloud droplet number concentration (CDNC or symbol N_c). One of the most widely used parameterization scheme is developed by Khairoutdinov and Kogan (2000) ("KK2000" hereafter), which has the form

$$\frac{\partial q_r}{\partial t} = C(q_c)^{\beta_q}(N_c)^{\beta_N}, \tag{1}$$

where $\frac{\partial q_r}{\partial t}$ is the rain water tendency due to the autoconversion process, q_c has the unit of kg/kg, and N_c of cm^{-3}. The three parameters $C = 1350$, $\beta_q = 2.47$ and $\beta_N = -1.79$ are derived through a simple least-square fitting of the autoconversion rate results from a large-eddy simulation with bin microphysics that can simulate the process-level physics. Even though this is highly simplified, the parametrization scheme still faces a great challenge. The calculation of grid-mean autoconversion efficiency requires the knowledge of subgrid distributions of LWC and CDNC, but in the GCMs only grid-mean quantities $\langle q_c \rangle$ and $\langle N_c \rangle$ are known and available for use in the computation of autoconversion rate. As pointed out by Pincus and Klein (2000), for a process $f(x)$ such as autoconversion that is nonlinearly dependent on subgrid variables, x, the grid-mean value $\langle f(x) \rangle$ is not equal to the value estimated based on the grid-mean $\langle x \rangle$, i.e., $\langle f(x) \rangle \neq f(\langle x \rangle)$. Mathematically, if $f(x)$ is convex, then $f(\langle x \rangle) < \langle f(x) \rangle$ (Larson and Griffin, 2013; Larson et al., 2001). To take this effect into account, a parameter E is often introduced in the GCM as part of the parameterization such that $\langle f(x) \rangle = E \cdot f(\langle x \rangle)$. It is referred to as the "enhancement factor" in many studies and this study too because $E > 1$ for a convex function. Such a nonlinear effect is not just limited to the autoconversion process. Some other examples are the plane-parallel albedo bias (Barker, 1996; Cahalan et al., 1994; Oreopoulos and Davies, 1998a), subgrid cloud droplet activation (Morales and Nenes, 2010) and accretion (Boutle et al., 2014; Lebsock et al., 2013).

The value of E is determined primarily by two factors: the nonlinearity of $f(x)$ and the subgrid probability density function (PDF) $P(x)$. Given the same subgrid variation of LWC, i.e., $P(q_c)$, the nonlinear effect impacts the autoconversion parameterization more than it does on the accretion, because the former is a more nonlinear function of q_c than the latter. For the same $f(x)$, a grid box with a narrow and symmetric $P(x)$ would require a smaller E than another grid box with a broader and non-symmetric $P(x)$. Ideally, the value of the enhancement factor E should be diagnosed from the subgrid cloud PDF $P(x)$. Unfortunately, because this is not possible in most conventional GCMs, the value of E is usually assumed to be a constant for the lack of better options. The E for autoconversion due to subgrid LWC variation is assumed to be 3.2 in the two-moment cloud microphysics parameterization schemes by Morrison and Gettelman (2008) (MG scheme hereafter), which is employed in the widely used Community Atmosphere Model (CAM). This choice of $E = 3.2$ is based on an early study by Barker et al. (1996), in which the mesoscale variation of column-integrated optical thickness of the "overcast stratocumulus", "broken stratocumulus" and "scattered stratocumulus" are studied. The value $E = 3.2$ is derived based on the mesoscale variation of the broken stratocumulus.

Clearly, a simple constant E is not adequate. The following is a list of attempts to better understand the subgrid cloud variations and the implications for warm rain simulations in GCMs. Several previous studies have shown that the mesoscale cloud water variation is a strong function of cloud regime—the subgrid cloud water variation of Sc cloud is much different from that of Cu clouds (Barker et al., 1996; Lee et al., 2010; Oreopoulos and Cahalan, 2005; Wood and Hartmann, 2006). As the first part of a two-part study, Larson and Griffin (2013) first laid out a systematic theoretical basis for understanding the effects of subgrid cloud property variations on simulating various nonlinear processes in GCM, including not only the autoconversion but also the accretion, condensation, evaporation and sedimentation processes. In the second part, using cloud fields from a large-eddy simulation (LES), Griffin and Larson (2013) showed that inclusion of the enhancement factor indeed leads to more rainwater at surface in single-column simulations and makes them agree better with high-resolution large-eddy simulations. Using a combination of in situ measurement and satellite remote sensing data, Boutle et al. (2014) analyzed the spatial variation of both cloud and rain water, as well as their covariation, and developed a simple parameterization scheme to relate the subgrid cloud water variance to the grid-mean cloud fraction. Later, the study of Boutle et al. (2014) was extended by Hill et al. (2015) who developed a cloud regime dependent and scale-aware parameterization scheme for simulating subgrid cloud water variation. Recently, using the ground-based observations from three Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) sites, Xie and Zhang (2015) developed a scale-aware parameterization scheme for GCMs to account for subgrid cloud water variation. Also using ARM measurement, Ahlgrimm and Forbes (2016) analyzed the dependence of cloud water variability on cloud regime. Although these previous studies have shed important light on subgrid cloud variation and the implications for GCM, they lack a global perspective because they are only based on limited data (e.g., LES cases, in situ and ground-based measurement). Currently, satellite remote sensing observation is the only way to achieve a global perspective. Using the observations from the space-borne radar CloudSat, Lebsock et al. (2013) showed that the subgrid cloud water variance is smaller over the Sc region than over the Cu region, and as a result the enhancement factor shows an increasing trend from Sc to Cu region. They also highlighted importance of considering the subgrid co-variability of cloud water and rain water in the computation of the accretion rate. On the modeling side, Guo et al. (2014) investigated the sensitivity of cloud simulation in the Geophysical Fluid Dynamics Laboratory (GFDL) Atmospheric General Circulation Model (AM) to the subgrid cloud water parameterization schemes. A similar study was carried out by Bogenschutz et al. (2013) using the National Center of Atmospheric Research (NCAR) Community Atmospheric Model (CAM). Both studies show that the more sophisticated subgrid parameterization scheme— Cloud Layers Unified by Binormals (CLUBB) (Golaz et al., 2002a; 2002b; Larson et al., 2002)—lead to a better simulation of clouds in the model. However, a more recent study by Song et al. (2018b) reveals that the CLUBB in CAM version 5.3 (CAM5.3) overestimates the enhancement factor in the trade wind cumulus cloud region, which in turn leads to excessive drizzle in the model and "empty clouds" with near-zero cloud water. In addition to CLUBB, the so-called super-parameterization (a.k.a Multiscale Modeling Framework (MMF)), which uses cloud resolving model embeded in the GCM grids to diagnose sub-grid cloud variations (Randall et al., 2003), have also gained increasing popularity. Takahashi et al. (2017) compared the subgrid cloud water variations simulated by a CAM-MMF model with those derived from A-Train observations and found reasonable agreement.

Despite these previous studies, many questions remain unanswered. First of all, all the previous studies, as far as we know, have focused on the impact of subgrid cloud water q_c variation. The potential impact of subgrid variation of N_c and the co-variability of N_c with q_c have been overlooked so far. Given the same amount of q_c, a cloud with a smaller N_c would have larger droplets and therefore larger precipitation efficiency than another cloud with a larger N_c. For the same reason, other things equal, a grid with positive correlation of subgrid N_c and q_c would be less efficient in terms of autoconversion than a grid with negative correlation of the two. Secondly, most of previous studies are based on the assumption that the subgrid cloud property variation follows certain well-behaved distributions, usually either Gamma (e.g., Barker, 1996; Morrison and Gettelman, 2008; Oreopoulos and Barker, 1999; Oreopoulos and Cahalan, 2005) or Lognormal (Boutle et al., 2014; Larson and Griffin, 2013; e.g., Lebsock et al., 2013). However, the validity and performance of the assumed PDF shape are seldom checked. Furthermore, although the study by Lebsock et al. (2013) has depicted a global picture of the enhancement factor for the autoconversion modeling in GCM, the picture is far from clear due to the small sampling rate of CloudSat observations.

In this study, we revisit the subgrid variations of liquid-phase cloud properties over the tropical ocean using 10 years of MODIS cloud observations, with the overarching goal to better understand the potential impacts of subgrid cloud variations on the warm rain processes in the conventional GCMs. Similar to previous studies, we will quantify the subgrid cloud water variations based on MODIS observations. Going one step further, we will also attempt to unveil for the first time the subgrid N_c variation, as well as its correlation with cloud water, and investigate the implications for warm rain simulations in GCM. Moreover, we will take advantage of the wide spatial coverage of MODIS data to achieve a more detailed picture of the enhancement factor for the autoconversion simulation. Last but not least, we will evaluate the two widely used distributions, i.e., Lognormal and Gamma, in terms of their performance and limitations for simulating the enhancement factor. We will first explain the theoretical background in Section 2 and introduce the data and methodology in Section 3. The MODIS observations will be presented and discussed in Section 4. The implications for the autoconversion parameterization in the GCMs will be discussed in 5. The main findings will be summarized in Section 6 with an outlook for future studies.

2. Theoretical Background

2.1. Theoretical Distributions to describe subgrid cloud property variations

In previous studies, the spatial variations of cloud properties, such as cloud optical thickness (COT), cloud liquid water path (LWP) and cloud liquid water content (LWC), are often described using either of two theoretical distributions—the Gamma and Lognormal distribution. The probability density function (PDF) from a Gamma distribution is a two-parameter function as follows (Barker, 1996; Oreopoulos and Davies, 1998b):

$$P_G(x) = \frac{1}{\Gamma(v)} \alpha^{v} x^{v-1} \exp(-\alpha x), \tag{2}$$

where Γ is the Gamma function, v is the so-called inverse relative variance, and α the so-called rate parameter. If x follows the Gamma distribution, its mean value Is given by

$$\langle x \rangle = \int_0^\infty x\, P_G(x) dx = \frac{v}{\alpha}, \tag{3}$$

and variance given by

$$Var(x) = \int_0^\infty (x - \langle x \rangle)^2\, P_G(x) dx = \frac{v}{\alpha^2}. \tag{4}$$

It follows from Eq. *(3)* and *(4)* that the so-called inverse relative variance is

$$v = \frac{1}{\eta} = \frac{\langle x \rangle^2}{Var(x)}, \tag{5}$$

where $\eta = \frac{Var(x)}{\langle x \rangle^2}$ is the relative variance. If x follows the Gamma distribution, for a physical process $M(x)$ that is a power function of x,

$$M(x) = Kx^\beta, \tag{6}$$

then the expected value $\langle M(x) \rangle$ is given by

$$\langle M(x) \rangle_G = K \int_0^\infty x^\beta\, P_G(x) dx = \frac{\Gamma(v+\beta)}{\Gamma(v) v^\beta} K \langle x \rangle^\beta,\ \beta > -v. \tag{7}$$

As explained in the introduction, for a nonlinear process $M(x)$, $\langle M(x) \rangle \neq M(\langle x \rangle)$. The ratio between the two E is by definition the enhancement factor:

$$E(P_G, v, \beta) = \frac{\langle Kx^\beta \rangle}{K \langle x \rangle^\beta} = \frac{1}{\langle x \rangle^\beta} \int_0^\infty x^\beta\, P_G(x) dx = \frac{\Gamma(v+\beta)}{\Gamma(v) v^\beta}, \tag{8}$$

The PDF of a Lognormal distribution is given as follows (Larson and Griffin, 2013; Lebsock et al., 2013):

$$P_L(x) = \frac{1}{\sqrt{2\pi} x \sigma} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \tag{9}$$

where $\mu = \langle \ln x \rangle$ and $\sigma^2 = Var(lnx)$ correspond to the mean and variance of lnx, respectively. The mean value of the Lognormal distribution is given by

$$\langle x \rangle = \int_0^\infty x\, P_L(x) dx = e^{\mu + \frac{\sigma^2}{2}}, \tag{10}$$

and the variance by

$$Var(x) = \int_0^\infty (x - \langle x \rangle)^2\, P_L(x) dx = e^{2\mu + \sigma^2}\left(e^{\sigma^2} - 1\right). \tag{11}$$

It follows from Eq. *(10)* and *(11)* that the inverse relative variance can be derived from the following equation

$$e^{\sigma^2} = 1 + \frac{Var(x)}{\langle x \rangle^2} = 1 + \frac{1}{v}. \tag{12}$$

If x follows the Lognormal distribution, the expected value of $\langle M(x) \rangle$ is

$$\langle M(x) \rangle_L = K \int_0^\infty x^\beta \, P_L(x) dx = \left(1 + \frac{1}{v}\right)^{\frac{\beta^2 - \beta}{2}} K \langle x \rangle^\beta. \tag{13}$$

Evidently, the corresponding enhancement factor is given by

$$E(P_L, v, \beta) = \frac{\langle K x^\beta \rangle}{K \langle x \rangle^\beta} = \left(1 + \frac{1}{v}\right)^{\frac{\beta^2 - \beta}{2}}. \tag{14}$$

Note that Eq. *(7)* and *(8)* are only valid when $\beta > -v$ because Gamma function $\Gamma(v + \beta)$ can run into singular values when $v + \beta$<0. In contrast, Eq. *(13)* and *(14)* are valid for any real value β. This is one advantage of the Lognormal distribution over the Gamma distribution.

An example of the Gamma and Lognormal distributions for q_c is shown in Figure 1a. In this example, both distributions have the same mean $\langle q_c \rangle = 0.5 g/kg$ and also the same inverse relative variance $v_q = 3$. Although the general shapes of the two PDFs are similar, they differ significantly at the two ends: the Gamma PDF is larger than Lognormal PDF over the small values of q_c, and the opposite is true over the large values of q_c. The Gamma and Lognormal distributions can also be used to describe the spatial variation of N_c (Gultepe and Isaac, 2004). An example is given in Figure 1c, in which q_c is a constant of $0.5 g/kg$, $\langle N_c \rangle = 50\ cm^{-3}$, and $v_N = 5.0$. Figure 1 b shows the autoconversion rate based on the KK2000 parameterization scheme for the Gamma $P_G(q_c)$ and Lognormal $P_L(q_c)$ that are shown in Figure 1a. Interestingly, although the cumulative autoconversion rates based on the two types of PDFs are almost identical, the contribution to the total autoconversion rate from the different LWC bins are quite different. As show in Figure 1a, the $P_L(q_c)$ has a longer tail than the $P_G(q_c)$, i.e., the occurrence probability of large q_c (e.g., $q_c > 2.0 g/kg$) is much higher in the Lognormal than in Gamma PDF. This difference is further amplified in the autoconversion rate computation in Figure 1b because the autoconversion rate is proportional to $q_c^{2.47}$.

The enhancement factors based on the Gamma (i.e., $E(P_G, \beta)$ in Eq. (8)) and Lognormal (i.e., $E(P_L, \beta)$ in Eq. (14)) PDF for $\beta_q = 2.47$ are plotted as a function of the inverse relative variance v in Figure 2. When subgrid clouds are more homogenous i.e., $v > 1$, the enhancement factor based on the two PDFs are similar. However, for more inhomogeneous grids with i.e., $v <$

1, the $E(P_L, \beta)$ is significantly larger than that $E(P_G, \beta)$, which is probably because of the longer tail of $P_L(q_c)$ as shown in Figure 1 a and b.

2.2. Impacts of subgrid cloud variations on warm rain parameterization in GCM

The warm rain process in MBL clouds involves many interacting microphysical processes. In this study, we only focus only on the simulation of autoconversion in GCM. Other nonlinear processes, such as accretion and evaporation have been investigated in previous studies (Boutle et al., 2014; Lebsock et al., 2013).

Ideally, if the subgrid variations of q_c and N_c are known, then the grid-mean in-cloud autoconversion rate should be derived from the following integral

$$\langle \frac{\partial q_r}{\partial t} \rangle = \int_0^\infty \int_0^\infty C(q_c)^{\beta_q}(N_c)^{\beta_N} P(q_c, N_c) dq_c dN_c, \tag{15}$$

where $P(q_c, N_c)$ is the joint PDF of q_c and N_c. Unfortunately, most conventional GCMs lack the capability of predicting the subgrid variations of cloud properties, with only a couple of exceptions (Thayer-Calder et al., 2015). What is known from the GCM is usually the in-cloud grid-mean values $\langle q_c \rangle$ and $\langle N_c \rangle$. As a result, instead of using Eq. *(15)*, the autoconversion rate in GCMs is usually computed from the following equation

$$\langle \frac{\partial q_r}{\partial t} \rangle = E \cdot C(\langle q_c \rangle)^{\beta_q}(\langle N_c \rangle)^{\beta_N}, \tag{16}$$

where E is the enhancement factor defined as:

$$E = \frac{\int_0^\infty \int_0^\infty (q_c)^{\beta_q}(N_c)^{\beta_N} P(q_c, N_c) dq_c dN_c}{(\langle q_c \rangle)^{\beta_q}(\langle N_c \rangle)^{\beta_N}}. \tag{17}$$

The value of the enhancement factor depends on the subgrid variations of q_c and N_c. If clouds are homogenous on the subgrid scale, then $E \sim 1$. The more inhomogeneous the clouds are, the larger the E is. In the special case where q_c and N_c are independent, then the joint PDF $P(q_c, N_c)$ becomes $P(q_c, N_c) = P(q_c)P(N_c)$, where $P(q_c)$ and $P(N_c)$ are the PDF of the subgrid q_c and N_c. Consequently, Eq. (15) reduces to

$$\langle \frac{\partial q_r}{\partial t} \rangle = C \int_0^\infty (q_c)^{\beta_q} P(q_c) dq_c \int_0^\infty (N_c)^{\beta_N} P(N_c) dN_c, \tag{18}$$

and Eq.(17) to

$$E = E_q \cdot E_N, \tag{19}$$

where E_q is the enhancement factor due to the subgrid variation of cloud water which has the form,

$$E_q = \frac{\int_0^{\infty} (q_c)^{\beta_q} P(q_c) dq_c}{(\langle q_c \rangle)^{\beta_q}}, \tag{20}$$

and the E_N is the enhancement factor due to the subgrid variation of CDNC which has the form,

$$E_N = \frac{\int_0^{\infty} (N_c)^{\beta_N} P(N_c) dN_c}{(\langle N_c \rangle)^{\beta_N}}. \tag{21}$$

Obviously, if $P(q_c)$ and $P(N_c)$ follow either Gamma or Lognormal distribution, then the above equations reduce to Eq. (*8*) or (*14*), respectively.

If q_c and N_c both have significant subgrid variations and they are not independent, the enhancement factor should ideally be diagnosed from Eq. (*17*). However, the joint PDF $P(q_c, N_c)$ may not be known and the integration can be time-consuming. Some previous studies proposed to approximate the $P(q_c, N_c)$ as a bivariate lognormal distribution as follows:

$$\begin{gathered} P(q_c, N_c) = \frac{1}{2\pi q_c N_c \sigma_q \sigma_N \sqrt{1-\rho^2}} \exp\left(-\frac{\zeta}{2}\right) \\ \zeta = \frac{1}{1-\rho^2}\left[\left(\frac{lnq_c - \mu_q}{\sigma_q}\right)^2 - 2\rho\left(\frac{lnq_c - \mu_q}{\sigma_q}\right)\left(\frac{lnN_c - \mu_N}{\sigma_N}\right) + \left(\frac{lnN_c - \mu_N}{\sigma_N}\right)^2\right], \end{gathered} \tag{22}$$

where ρ is the correlation coefficient between q_c and N_c (Larson and Griffin, 2013; Lebsock et al., 2013). As such, both q_c and N_c follow a marginal lognormal distribution in Eq. (*9*). Substituting Eq. (*22*) into Eq. (*17*), we obtain the enhancement factor for the bivariate lognormal distribution that consists of three terms

$$E = E_q(P_L, v_q, \beta_q) \cdot E_N(P_L, v_N, \beta_N) \cdot E_{COV}(\rho, \beta_q, \beta_N v_q, v_N), \tag{23}$$

where $E_q(P_L, v_q, \beta_q) = \left(1 + \frac{1}{v_q}\right)^{\frac{\beta_q^2 - \beta_q}{2}}$ and $E_N(P_L, v_N, \beta_N) = \left(1 + \frac{1}{v_N}\right)^{\frac{\beta_N^2 - \beta_N}{2}}$ correspond to the impacts of subgrid q_c and N_c variance, respectively (i.e., Eq. (*14*)), and the third term

$$E_{COV}(\rho, \beta_q, \beta_N, v_q, v_N) = \exp(\rho \beta_q \beta_N \sigma_q \sigma_N), \tag{24}$$

corresponds to the impact of the co-variation of q_c and N_c on the enhancement factor. Obviously, Eq. (23) reduces to Eq. (19) when q_c and N_c are uncorrelated (i.e., $\rho = 0, E_{COV} = 1$). If q_c and N_c are negatively correlated (i.e., $\rho < 0$ and $E_{COV} > 1$), clouds with larger q_c would tend to have smaller N_c. The autoconversion rate in such a case would be larger than that in the case where q_c and N_c are positively correlated (i.e., i.e., $\rho > 0$ and $E_{COV} < 1$). A positive correlation would exist, for instance, if all droplets in cloud were the same size, but some parcels had more droplets than other parcels.

Most current GCMs do not have the capability to simulate the subgrid cloud property variations. They usually have to use pre-defined subgrid cloud variations in the computation of grid-mean autoconversion rate instead of using prognostic values. For example, in the MG scheme for the CAM5.3, the subgrid q_c is assumed to follow the Gamma distribution in Eq. *(2)* with a fixed $v_q = 1$ and as a result constant $E_q = 3.2$. Lately, advanced subgrid parameterization schemes, such as CLUBB, have been implemented in several GCMs, including CAM6 and GFDL AM model (Bogenschutz et al., 2017; Guo et al., 2015; 2014), which provides information on the subgrid q_c variation to the host model. The information can then be used to dynamically diagnose the enhancement factor E_q, which will help the model simulate the cloud regime dependence of E_q (Guo et al., 2010; 2014).

However, as explained above, not only the subgrid variation of q_c but the subgrid variation of N_c can also influence the enhancement factor. Unfortunately, this aspect has been ignored by almost all GCMs, even the latest CAM6 with CLUBB. Physically, provided the same q_c, a cloud with smaller N_c would have larger droplet size and therefore larger precipitation efficiency than the cloud with larger N_c. Because the autoconversion rate depends nonlinearly on N_c, the grid-mean autoconversion rate computed based on a skewed PDF of N_c (i.e., $\int_0^\infty (N_c)^{\beta_N} P(N_c) dN_c$) would be different from that computed based on the mean of N_c (i.e., $(\langle N_c \rangle)^{\beta_N}$). The autoconversion enhancement factor based on the Lognormal PDF $E(P_L, \beta)$ for $\beta_N = -1.79$ is given in Figure 2. Interestingly, at the same inverse relative variance v, the enhancement factor based on the same Lognormal PDF $E(P_L, \beta)$ for $\beta_N = -1.79$ is actually larger than that for $\beta_q = 2.47$ because of the formula of the exponent in Eq. *(14)* (i.e., $\frac{\beta^2 - \beta}{2}$). Moreover, the correlation between N_c and q_c can also be important. Going back to Eq.(23), evidently, $E > E_q$ if and only if $E_N \cdot E_{COV} > 1$. After some manipulation, we can show that if $\beta_N < 0$ and $\sigma_N > 0$, then

$$E_N \cdot E_{COV} > 1, if\ \rho < \frac{\sigma_N}{\sigma_q} \cdot \frac{(1-\beta_N)}{2\beta_q}. \quad (25)$$

This equation reveals that when q_c and N_c are weakly or negatively correlated ($\rho \leq 0$), considering only E_q would tend to underestimate E. On the other hand, however, if q_c and N_c are highly positively correlated ($\rho \sim 1$) then considering E_q only would tend to overestimate E.

3. Data and Methodology

To derive the above-mentioned enhancement factors , we will use 10 years (2007 ~ 2016) of the latest collection 6 (C6) *daily mean* level-3 cloud retrieval product from the Aqua-MODIS instrument (product name "MYD08_D3"), which contains the gridded statistics of cloud properties computed from pixel-level (i.e., level-2) retrievals. As summarized in Platnick et al. (2003; 2017), the operational level-2 MODIS cloud product provides cloud masking (Ackerman et al., 1998), cloud top height (Menzel et al., 1983), cloud top thermodynamic phase determination (Menzel et al., 2006), and COT, cloud effective radius (CER) and LWP retrievals based on the bi-spectral solar reflectance method (Nakajima and King, 1990). All MODIS level-2 atmosphere products, including the cloud, aerosol and water vapor products, are aggregated to 1°×1° spatial resolution on a daily, eight-day, and monthly basis. Aggregations include a variety of scalar statistical information, including mean, standard deviation, max/min occurrences, as well as histograms including both marginal and joint histograms. For COT, CER and LWP, the MODIS level-3 product provides both their "in-cloud" grid-mean values ($\langle x \rangle$) and subgrid standard deviations (σ_x). The inverse relative variance ν can then be derived from Eq. *(5)*, i.e., $\nu = \langle x \rangle^2 / \sigma_x^2$. Note that the operational MODIS product provides two CER retrievals, one based on the observation from the band 7 centered around 2.1 µm and the other from band 20 at 3.7 µm. As discussed in several previous studies (Cho et al., 2015; Zhang and Platnick, 2011; Zhang et al., 2012; 2016), the 3.7 µm band CER retrieval is more resilient to the 3-D effects and retrieval failure than the 2.1 µm band retrievals. For these reasons, it is used as the observational reference in this study.

Given the COT and CER retrieval, the operational MODIS product estimates the LWP of cloud using

$$LWP = \frac{2}{3}\rho_w COT \cdot CER, \tag{26}$$

where ρ_w is the density of water. Several studies have argued that a smaller coefficient of 5/9, instead of 2/3, should be used in estimation of LWP (Lebsock et al., 2011; Seethala and Horváth, 2010; Wood and Hartmann, 2006). The choice of coefficient does not matter in this study because it is a common factor in the calculation of ν. The choice of the coefficient has no impact on our study, because we are interested in the relative inverse variance $\nu = \langle x \rangle^2 / \sigma_x^2$. We note here that it is the LWC q_C, instead of the LWP, that is used in the KK2000 scheme. So, the spatial variability of LWC is what is most relevant. However, the remote sensing of cloud water vertical profile from satellite sensor for liquid-phase clouds is extremely challenging even with active sensors. It is why most previous studies using the satellite observations analyzed the spatial variation of LWP, rather than LWC. In fact, even Lebsock et al. (2013), who used the level-2 CloudSat observations, had to use the vertical averaged LWC in their analysis. Airborne in situ measurement faces similar challenge. For example, Boutle et al. (2014) use the LWC observation along “horizontal flight tracks” to study the spatial variability of cloud water, which only samples the LWC at certain levels of MBL clouds. Ground-based observations are much better than satellite and airborne observation in this regard. Recently, Xie and Zhang (2015) analyzed the cloud water profiles retrieved using ground-based radars from the three ARM sites and found no obvious in-cloud vertical dependence of the spatial variability of LWC. Following these previous studies, we assume that the horizontal subgrid variation of LWC is *not* strongly dependent on height and its value can be inferred from the spatial variability of the vertical integrated quantity LWP. The uncertainty caused by this assumption will be assessed in future studies.

The current MODIS level-3 cloud product does *not* provide CDNC retrievals. Following previous studies (Bennartz, 2007; Bennartz and Rausch, 2017; Grosvenor and Wood, 2014; McCoy et al., 2017a), we estimate N_c of liquid-phase clouds from the MODIS retrieved COT (τ) and CER (r_e) based on the classic adiabatic cloud model

$$N_c(\tau, r_e) = \frac{\sqrt{5}}{2\pi k}\frac{\sqrt{f_{ad}\Gamma_w}}{\sqrt{\rho_w Q_e}}\tau^{\frac{1}{2}} r_e^{-\frac{5}{2}} = \frac{\sqrt{15}}{2\pi k}\frac{\sqrt{f_{ad}\Gamma_w}}{\rho_w\sqrt{2Q_e}} LWP^{\frac{1}{2}} r_e^{-3}, \tag{27}$$

where ρ_w is the density of water; $Q_e \approx 2$ is the extinction efficiency of cloud droplets; k is the ratio of r_e to mean volume-equivalent radius; f_{ad} is the adiabaticity of the cloud; Γ_w is the LWC lapse rate. Following previous studies, we assume $k = 0.8$ and $f_{ad} = 1.0$ to be constant and compute Γ_w from the grid mean liquid cloud top temperature and pressure. The theoretical basis and main uncertainty sources of the CDNC estimation based on the adiabatic cloud model from MODIS-like passive cloud retrievals are nicely reviewed by Grosvenor et al. (2018).

Ideally, the values of LWP and CDNC should be estimated on pixel-by-pixel basis from the level-2 MODIS product. However, pixel-by-pixel estimation is highly time consuming, which makes it difficult to achieve a global perspective. Using an alternative method, many previous studies estimate the grid-level CDNC statistics from the joint histogram of COT vs. CER provided in the level-3 MODIS cloud products (Bennartz, 2007; McCoy et al., 2017a; 2017b). For a given 1°×1° grid-box, the liquid-phase COT-CER joint histogram provides the counts of successful cloud property retrievals with respect to 108 joint COT-CER bins that are bounded by 13 COT bin boundaries, ranging from 0 to 150, and 10 CER bin boundaries, ranging from 4 µm to 30 µm. With the joint histogram, which is essentially the joint PDF of COT and CER $P(\tau, r_e)$, we can estimate the grid mean and variance of CDNC from the following equations

$$\langle x \rangle = \int\int x(\tau, r_e) P(\tau, r_e) d\tau dr_e, \tag{28}$$

$$Var(x) = \int\int (x(\tau, r_e) - \langle N_c \rangle)^2 P(\tau, r_e) d\tau dr_e, \tag{29}$$

where x can be either LWP or CDNC. Figure 3a shows the LWP in Eq. *(26)* as a function of the 13 COT bins and 10 CER bins from the MODIS level-3 product. As expected, the largest LWP values are found when both COT and CER are large. Figure 3b shows the CDNC in Eq. *(27)* as a function of the COT and CER bins. As expected, the largest CDNC values are found when both COT is large and CER is small. Figure 3c shows an example of the COT-CER joint histogram from the Aqua-MODIS daily level-3 product "MYD08_D3" on January 09th, 2007 at the grid box 1°S and 1°W. In this particular grid box, a combination of ~2-4 COT and ~10-12 µm CER is the most frequently observed cloud value. Using the joint histogram in Figure 3c, we can derive the mean and variance of both LWP and COT using the Eqs. *(28)* and *(29)*.

The efficiency of using the level-3 MODIS product is accompanied by three important limitations. First of all, as mentioned earlier MODIS provides only LWP retrievals while LWC is needed in the KK2000 scheme. Second, the current level-3 MODIS cloud product has a fixed 1°x1° spatial resolution. Although this resolution is highly relevant to the current generation of GCMs, i.e., Coupled Model Intercomparison Project Phase 6 (CMIP5) (Eyring et al., 2016), future GCMs may have significantly finer resolution. Third, it is difficult to sub-sample the pixels with the best retrieval quality. These limitations will have to be addressed in future studies.

4. Grid-mean and subgrid variations of liquid-phase cloud properties

In this study, we limit our analysis to tropical oceans only where warm rain is frequent and MODIS cloud retrievals have a relatively better quality than over land or over high latitude. The annual mean total cloud fraction (f_{tot}), liquid-phase cloud fraction (f_{liq}), in-cloud COT, CER from the 3.7 µm band, LWP and estimated CDNC over the tropical oceans based on 10 years Aqua-MODIS retrievals are shown in Figure 4. The highest f_{liq} in the tropics is usually found in the stratocumulus (Sc) decks over the Eastern boundary of the ocean, e.g., SE Pacific off coast of Peru, NE Pacific off the coast of California and SE Atlantic off the coast of Namibia. The liquid-cloud fraction reduces significantly toward the open ocean trade wind regions, where the dominant cloud types are broken cumulus (Cu). Close to the continents, the Sc decks are susceptible to the influence of continental air mass with higher loading of aerosols in comparison with pristine ocean environment, which is probably the reason the SC decks have smaller CER and higher CDNC than the open-ocean trade cumulus (Figure 4 d and f). The in-cloud COT (Figure 4 c) and LWP (Figure 4 e) generally increase from the Sc decks to the open-ocean Cu regime, although less dramatically than the transition of cloud fraction. The Sc decks and the Sc-to-Cu transition are the most prominent features of liquid-phase clouds in the tropics. However, as mentioned in the introduction, simulating these features in the GCMs proves to be an extremely challenging task, and most GCMs suffer from some common problems, such as the “too few too bright” problem and the abrupt Sc-to-Cu transition problem (Kubar et al., 2014; Nam et al., 2012; Song et al., 2018a).

Switching the focus now from grid-mean values to subgrid variability, we will show the grid-level inverse relative variances $\nu = \langle x \rangle^2 / Var(x)$) for several key cloud properties. Here, we first derive the daily mean ν and then aggregate the result to monthly mean values. Therefore, for each grid box we have 120 samples (i.e., 10 years x 12 months) of monthly mean ν for analysis and visualization. Because the value of ν can be ill-behaved when $Var(x)$ approaches zero, instead of the mean value, we plot the median value of $\tilde{\nu}$ based on 120 months of MODIS observations in Figure 5. There are several interesting and important features in Figure 5. First of all, the $\tilde{\nu}$ of all four sets of cloud properties (i.e., COT, CER, LWP and CDNC) all exhibits a clear and similar Sc-to-Cu transition, with larger values in the Sc region and smaller value in the broken Cu regions. This indicates that cloud properties, including both optical and microphysical properties, are more homogenous, in terms of spatial distribution within the grid, in the Sc region than in the Cu region. Secondly, the value of $\tilde{\nu}$ of CER (i.e., 10~100 in Figure 5b) is larger than that of the other properties (i.e., 1~10) by almost an order of magnitude, indicating that the subgrid variability of CER is very small. On the other hand, however, it is important to note that the $\tilde{\nu}$ of CDNC (Figure 5d) is comparable with that of COT (Figure 5a) and LWP (Figure 5c). The reason is probably in part because the highly nonlinear relationship between CDNC and CER (i.e., $N_c \sim r_e^{-\frac{5}{2}}$) leads to a stronger variability of CDNC than CER, and also in part because the variability of CDNC is also contributed by the subgrid variation of COT. In some regions, the Gulf of Guinea, East and South China Sea, and Bay of Bengal for example, the $\tilde{\nu}$ of CDNC is close to unity, indicating the subgrid standard deviation of CDNC is comparable to the grid-mean values in these regions. As discussed in the next section, the significant subgrid variability of CDNC in these regions should be taken into account when modeling the nonlinear processes, such as the autoconversion, in GCM to avoid systematic biases due to the nonlinearity effect.

The values of $\tilde{\nu}$ in Figure 5 from this study are in reasonable agreement with previous studies. Barker (1996) selected a few dozens of cloud scenes, each about 100 ~ 200 km in size, from the Landsat observation and analyzed their spatial variability of COT. It is found that the typical value of ν for "overcast stratocumulus", "broken stratocumulus" and "scattered cumulus" is 7.9, 1.2, and 0.7, respectively (see their Table 3), which is consistent with the Sc-to-Cu transition pattern seen in Figure 5. Oreopoulos and Cahalan (2005) derived the subgrid inhomogeneity of

COT on a global scale from the level-3 Terra-MODIS retrievals. Although using a different metric (i.e., their inhomogeneity parameter is defined as $\chi = \exp(\ln\langle\tau\rangle) / \langle\tau\rangle$), they also found systematic increase of inhomogeneity (decreasing value of χ) from the Sc region to cu region. Also using the MODIS cloud property retrievals, Wood and Hartmann(2006) investigated the meso-scale spatial variability of LWP in the NE Pacific and SE Pacific region. The ν of LWP is found to increase systematically with meso-scale cloud fraction and the relationship between the two can be reasonably explained by a simple PDF cloud thickness model in Considine et al. (1997). See also Kawai and Teixeira (2010).

As explained in section 2, the correlation between cloud water and CDNC can also influence the computation of enhancement factor and thereby the grid-mean autoconversion rate. Figure 5e shows the median value of the LWP and CDNC correlation coefficient $\tilde{\rho}$. Similar to the derivation of median $\tilde{\nu}$, we first compute the monthly mean ρ from daily MODIS observations and then derive the median value of $\tilde{\rho}$ for each grid from the 120 months of observation. As shown in Figure 5e, at the subgrid level, the LWP and CDNC tend to be positively correlated almost over all tropical oceans. Mathematically, this is not surprising because as shown in Figure 5b and c, the subgrid variability of r_e is order of magnitude smaller than that of LWP. Since CDNC is proportional to $LWP^{\frac{1}{2}} r_e^{-3}$ according to Eq. (27), the subgrid variability of CDNC is mainly determined by the variability of LWP, leading to the positive correlation. Physically, the correlation can be explained by several mechanisms. For example, Wood et al. {*Wood:2018cx} and O et al. {*O:2018to} found that a large amount of low-level water clouds over the stratocumulus to cumulus transition are “optically thin veil clouds”. These clouds are usually associated with low LWP and low CDNC (therefore positive correlation) and probably caused by the strong precipitation scavenging process in the active cumulus. Note that our definition of ρ is the subgrid spatial correlation of LWP and CDNC. It may be different from the definition used in many aerosol indirect effect studies where the temporal correlation of monthly mean LWP and CDNC is more interested.

5. Implications for warm-rain simulations in GCM

5.1. Influence of subgrid variation of cloud water

As discussed in Section 2.2, most current GCMs only considers the impact of subgrid cloud water variation on autoconversion rate but ignore the impact of subgrid CDNC variation. To make our analysis relevant to the current GCMs, we first analyze E_q in Eq. (20) based on observation. The impacts of subgrid CDNC variation (i.e., E_N) and its correlation with cloud water (i.e., E_{COV}) will be analyzed in the next section.

We derive E_q using two approaches. First, we derive it from the observed LWP PDF based on Eq. *(20)*. As such, we do not have to make any assumption about the shape of LWP PDF although solving the integration in Eq. (20) is time-consuming. In the second approach, we first derive the relative inverse relative variance ν of LWP and then derive the enhancement factor by assuming the subgrid PDF to be either Gamma or Lognormal. This approach is more efficient, but it may be subject to error if the true PDF deviates from the assumed PDF shape. Figure 6a shows the annual mean enhancement factor E_q in the tropical region derived based on Eq. *(20)* (i.e., the first approach) from 10 years of MODIS observation. Figure 6 b and c show the annual mean enhancement factor E_q derived by assuming the subgrid cloud water follows the Lognormal (i.e., Eq. (14)) and Gamma distribution (i.e., Eq. *(8)*), respectively. There are a couple of interesting and important points to note. First of all, similar to the grid-mean quantities in Figure 4, the enhancement factor E_q also shows a clear Sc-to-Cu transition. Over the Sc decks, because clouds are more homogeneous ($\tilde{\nu} > 5$), the enhancement factor E_q is only around 1 ~ 2.5, while over the Cu regions, the more inhomogeneous clouds with $\tilde{\nu} < 1$ leads to a larger enhancement factor E_q around 3~5. As aforementioned, in the current CAM5.3, E_q is assumed to be a constant of 3.2. While this value is within the observational range, it obviously cannot capture the Sc-to-Cu transition. In fact, the constant value 3.2 overestimates the E_q over the Sc region and underestimates the E_q over the Cu region, which could lead to unrealistic drizzle production in both regions and to consequential impacts on cloud water budget, radiation and even aerosol indirect effects on the model. The second point to note is that the E_q based on the Lognormal PDF assumption in Figure 6 b agrees well with the results in Figure 6 a derived directly from the observation. In contrast, the E_q based on the Gamma PDF assumption in Figure 6 c tends to be smaller, especially in the Cu regions. This result seems to suggest that the Lognormal distribution provides a better fit to the observed subgrid cloud water variation than the Gamma distribution, which has rarely been noted and reported in the previous studies.

A flexible, cloud-regime dependent E_q could help improve the simulation of Sc-to-Cu transition in the GCM. If a GCM employs an advanced cloud parameterization scheme, such as CLUBB, that is able to provide regime-dependent information on subgrid cloud variation, i.e., ν, then the enhancement factor E_q could be diagnosed from ν . However, most traditional cloud parameterization schemes do not provide information on subgrid cloud variation. In such case, if one does not wish to use a constant E_q, but a varying regime-dependent scheme, then either ν or E_q need to be parameterized as a function of some grid-mean cloud properties resolved by the GCM. In fact, several attempts have been made along this line. Based on the combination air-borne in situ measurement and satellite remote sensing product, Boutle et al. (2014) parameterized the "fractional standard deviation" (which is equivalent to $1/\sqrt{\nu}$ in our definition) of liquid-phase cloud as a function of grid-mean cloud fraction. This scheme was later updated and tested in a host GCM in Hill et al. (2015), and was found to reduce the shortwave cloud radiative forcing biases in the model. In a recent study, Xie and Zhang (2015) derived the subgrid cloud variations from the ground-based observations from three Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) sites, and then parameterize the inverse relative variance ν as a function of the atmospheric stability.

Figure 7a shows the variation of inverse relative variance ν as a function of the grid-mean liquid-phase cloud fraction f_{liq}. In general, the value of ν increases with the increasing f_{liq}, which is expected from the Sc-to-Cu increase of f_{liq} in Figure 4b and the Sc-to-Cu decrease of ν in Figure 5c. The $\nu(f_{liq})$ pattern in Figure 7a is also consistent with the results reported in Wood and Hartmann (2006) and Lebsock et al. (2013). In the hope of obtaining a simple parameterization scheme for $\nu(f_{liq})$ that can be used in GCMs, we fit the median value of ν as a simple 3rd order polynomial of f_{liq} as follows:

$$\nu(f_{liq}) = 2.38 - 4.95 f_{liq} + 8.74 f_{liq}^2 - 0.49 f_{liq}^3,\ \ f_{liq} \in [0,1]. \tag{30}$$

To test the performance of this simple parameterization, we first substitute the f_{liq} from MODIS daily mean level-3 product into the above equation and then use the resultant v to compute the enhancement factor E_q. Unfortunately, the enhancement factor E_q computed based on the parameterized $v(f_{liq})$ as shown in Figure 8a substantially underestimate the observation-based results in Figure 6, especially over the Cu regions. The deviation is probably because the relationship between E_q and v is highly nonlinear (e.g., Eq. *(8)* and *(14)*) and therefore the above parameterization scheme that only fits the = value of v is not able to capture the variability of E_q. Based on this consideration, we tried an alternative approach. Instead of parameterization of v, we directly parameterize the enhancement factor E_q as a function of f_{liq}. Figure 7b shows the variation of E_q as a function of f_{liq}. As expected, E_q generally decreases with increasing f_{liq}. The median value of E_q is fitted with the following 3rd order polynomial of f_{liq}

$$E_q(f_{liq}) = 2.72{+}7.33 f_{liq} - 19.17 f_{liq}^2 + 10.69 f_{liq}^3,\ \ f_{liq} \in [0,1]. \quad (31)$$

As shown in Figure 8b, the value of E_q based on the above equation clearly agrees with the observation-based values in Figure 6 better than that based on the parameterization of $v(f_{liq})$. The elimination of the middle step indeed improves the parameterization results. While this is encouraging, it should be kept in mind that the Eq. *(31)* has very limited application, i.e., it is only useful for the autoconversion rate computation for a particular value of the autoconversion exponent beta, i.e., $\beta_q = 2.47$. A good parameterization of v could be useful for not only autoconversion, but also for accretion and radiation computations. Another caution is that, if applied to a GCM, the performance of the $E_q(f_{liq})$ parameterization in Eq. *(31)* will be dependent on the simulated accuracy of f_{liq} in the model.

5.2. Influence of subgrid variance of CDNC

Now we will investigate the impacts of subgrid CDNC variation on the autoconversion rate simulation. For the moment, we will consider E_N only. The impact of CDNC and cloud water correlation will be discussed in the next section. Similar to E_q we first derive E_N from the CDNC PDF based on Eq. *(21)*. The annual mean result based on 10 years of MODIS observations is shown in Figure 9a. There are several intriguing points to note. First of all, the value of E_N is actually larger than E_q in Figure 9 such that we even have to use a different color scale for this plot. Secondly, E_N the regions with escalated E_N seem to coincide with the downwind regions of biomass burning aerosols (e.g., Gulf of Guinea, East Coast of South Africa), air pollution (i.e., Eastern China Sea), and, most interestingly, active volcanos (e.g., Kilauea Hawaii and Ambae Vanuatu). We have also checked the seasonal variation of the E_N and the results also support this observation. Another interesting feature to note is that, although the dust outflow regions such as Tropical East Atlantic and Arabian Sea, have heavy aerosol loading, the value of E_N there is only moderate. Figure 9b shows the value of E_N computed based on Eq. *(14)* from the inverse relative variance of v, assuming that the subgrid CDNC follows a Lognormal PDF. Although the overall pattern is consistent with Figure 9a, the assumption of Lognormal PDF seems to underestimate E_N. A closer examination indicates that the Lognormal PDF tend to underestimate the population of clouds with small CDNC, and therefore underestimate the variance of CDNC as well as E_N. We did not compute the E_N based on the Gamma distribution because of the singular value problem aforementioned in Section2.1.

We could not find any previous observation-based study on the global pattern of the subgrid variation of CDNC and the corresponding E_N. So, it is difficult for us to corroborate our results. On one hand, the magnitude of E_N is surprisingly large. As explained in Section 3, the CDNC is estimated based on Eq. *(27)* from the MODIS retrieval of COT and CER. Several previous studies have shown that the sub-pixel level surface contamination, subpixel cloud inhomogeneity, and three-dimensional radiative transfer effects, can cause significant errors in the MODIS CER retrievals especially over broken cloud regions (Zhang and Platnick, 2011; Zhang et al., 2012; 2016). Given the fact that the CDNC retrieval is highly sensitive to CER error as a result of $N_d \sim r_e^{-\frac{5}{2}}$, the influence of retrieval uncertainty on subgrid CDNC variation cannot be ruled out. On the other hand, the pattern of E_N in Figure 9a seems to suggest that there are some underlying physical mechanisms controlling the subgrid variation of CDNC, in which aerosols seem to play an important role. To achieve a better understanding, we analyzed the dependence of E_N on liquid cloud fraction and grid-mean CDNC in Figure 10, which reveals that E_N has a stronger dependence on CDNC than cloud fraction. This result seems to indicate that the pattern of E_N in Figure 9 is largely determined by physical mechanisms rather than retrieval uncertainties. Interestingly, the largest E_N is usually found when liquid cloud fraction is small and CDNC is large and decreases with decreasing CDNC and increasing cloud fraction. This pattern leads us to the following hypothesis: In the regions where aerosol is limited, even weak updraft can activate most cloud condensation nuclei (CCN). As a result, even if there is significant subgrid variation of turbulence at cloud base, the subgrid variation of CDNC remains small. In contrast, in regions where aerosol is abundant, the subgrid variation of turbulence becomes important. The subgrid variation of updraft leads to subgrid variation CDNC and thereby large E_N.

As far as we know, the results in Figure 9 and Figure 10 mark the first attempt based on satellite observations to unveil the global pattern of the subgrid variations of CDNC and investigate the consequential impacts on warm rain simulations in GCMs. Although obscured by satellite retrieval uncertainties, the results still provide valuable insights. First of all, the enhancement factor E_N due to the subgrid variations of CDNC is nonnegligible, even comparable the effect of subgrid cloud water variation (i.e., E_q). Second, the global pattern of E_N in Figure 9 provides a valuable map for future studies.

5.3. The combined effect of subgrid variations of cloud water and CDNC

Finally, in this section we examine the combined effect of subgrid variations of cloud water and CDNC, as well as their correlation, on the autoconversion rate simulation. The annual mean combined enhancement factor E derived based on Eq. *(17)* from 10 years of MODIS COT and CER observation is shown in Figure 11a. Comparing to the E_q in Figure 6 and E_N in Figure 9, the combined enhancement factor is generally larger. It is easy to see that the in some regions (e.g., Gulf of Guinea, East Coast of South Africa and Eastern China Sea) the combined enhancement factor E resembles the E_N while in other regions (i.e., trade wind cumulus regions over open ocean) it resembles more of E_q. Interestingly, because both E_q and E_N are small over the Sc decks, those regions have the smallest combined enhancement factor E. As discussed in Section 2.2, only when the subgrid variation of cloud water is uncorrelated with the subgrid variation of CDNC can the combined enhancement factor E be decomposed into the simple product of E_q and E_N (i.e., Eq. *(19)*). Figure 11b shows the annual mean value of the simple product $E_q \cdot E_N$, without considering the correlation between cloud water and CDNC. Evidently, the simple product substantially overestimates the combined enhancement factor derived from the joint PDF of LWP and CDNC. This result can be explained by the mostly positive subgrid correlation between LWP and CDNC in Figure 5e. As explained in section 2.2, the positive correlation means that clouds with more water also tend to have more CDNC. The autoconversion rate of such configuration is lower than that when LWP and CDNC have no correlation.

Together, the E_q in Figure 6, E_N in Figure 9 and the combined enhancement factor in Figure 11 lead us to the following important conclusion. It is not sufficient to consider only the impact of subgrid variation of cloud water (i.e., E_q) on the autoconversion rate simulation. The influences of subgrid CDNC variation, as well as the correlation between cloud water and CDNC, must also be taken into account to avoid significant error.

Finally, the combined enhancement factor derived based on Eq. (*23*) assuming that the LWP and CDNC follow the bi-variate lognormal distribution is shown in Figure 11c. Despite the tendency of overestimation, the result agrees reasonably well with that based on observed joint PDF in Figure 11a, clearly better than the simple product $E_q \cdot E_N$. This is encouraging as it suggests that the bi-variate lognormal distribution can be used in the future to model the combined effect of cloud water and CDNC on autoconversion rate simulation in GCMs.

6. Summary and Outlook

One of the difficulties in GCM simulation of the warm rain parameterization is how to account for the impact of subgrid variations of cloud properties, such as cloud water and CDCN, on nonlinear precipitation processes such as autoconversion. In practice, this impact is often treated by adding the enhancement factor term to the parameterization scheme. In this study, we derived the subgrid variations of liquid-phase cloud properties over the tropical ocean using the satellite remote sensing products from MODIS and investigated the corresponding enhancement factors for parameterizations of autoconversion rate. In comparison with previous work, our study is able to shed some new light on this problem in the following regards:

1. A theoretical framework is presented to explain the importance of the subgrid variation of CDNC and its correlation with cloud water on the autoconversion rate simulation in GCMs.

2. The wide spatial coverage of the Level-3 MODIS product enables us to depict a detailed quantitative picture of the enhancement factor E_q, which shows a clear cloud regime dependence, i.e., a Sc-to-Cu increase. The constant $E_q = 3.2$ used in the current CAM5.3 model overestimates and estimates the observed E_q in the Sc and Cu regions, respectively.
3. The E_q based on the Lognormal PDF assumption performs significantly better than that based on the Gamma PDF assumption. A simple parameterization scheme is provided to relate E_q to the grid-mean liquid cloud fraction, which can be readily used in GCMs.
4. For the first time, the enhancement factor E_N due to the subgrid variation of CDNC is derived from satellite observation, and the results reveal several regions downwind of biomass burning aerosols (e.g., Gulf of Guinea, East Coast of South Africa), air pollution (i.e., Eastern China Sea), and active volcanos (e.g., Kilauea Hawaii and Ambae Vanuatu). The largest E_N is usually found where CDNC is large and liquid cloud fraction is small and decreases with decreasing CDNC and increasing cloud fraction.
5. MODIS observations suggest that the subgrid LWP and CDNC are mostly positively correlated. As a result, the combined enhancement factor is significantly smaller than the simple product of $E_q \cdot E_N$ (i.e., assuming no correlation). The combined enhancement factor derived assuming LWP and CDNC to follow the bi-variate lognormal distribution agree with the observation-based results reasonably well.

As noted in the previous sections, this study has several important limitations, most of which are a result of using the level-3 MODIS observations. The fixed 1°x1° spatial resolution of MODIS level-3 product makes it impossible for us to investigate the scale-dependence of subgrid cloud variation. Similar to previous studies, we have to make several assumptions when estimating the CDNC from level-3 MODIS product. Furthermore, the retrieval uncertainties associated with the optically thin clouds in MODIS product pose a challenging obstacle for the quantification of subgrid cloud property variations and the corresponding enhancement factors. These limitations have to be addressed using additional independent observations from, for example, ground based remote sensing product and/or in situ measurement from air-borne field campaigns. Recently, a few novel methods have been developed to provide certain information on the subgrid cloud property variations to the host GCM. Most noticeable examples are the super-parameterization method (a.k.a. multi-scale modeling framework) (Wang et al., 2015) and the PDF-based higher-order turbulence closure methods (e.g., Cloud Layer Unified By Binormals, CLUBB (Golaz et al., 2002a; Guo et al., 2015; Larson et al., 2002) and Eddy-Diffusivity Mass-Flux (EDMF) (Sušelj et al., 2013)). The subgrid cloud property variations derived in this study provide the valuable observational basis for the evaluation and improvement of these schemes.

Acknowledgement:

Z. Zhang acknowledges the financial support from the Regional and Global Climate Modeling Program (Grant DE-SC0014641) funded by the Office of Biological and Environmental Research in the US DOE Office of Science. This work is also supported by the grant CyberTraining: DSE: Cross-Training of Researchers in Computing, Applied Mathematics and Atmospheric Sciences using Advanced Cyberinfrastructure Resources from the National Science Foundation (grant no. OAC–1730250). P.-L. Ma was support by the U.S. DOE, Office of Science, Office of Biological and Environmental Research, Regional and Global Model Analysis program. The Pacific Northwest National Laboratory is operated for the DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. V. Larson is grateful for financial support from Climate Model Development and Validation grant DE-SC0016287, which is funded by the Office of Biological and Environmental Research in the US DOE Office of Science. M. Wang was supported by the Minister of Science and Technology of China (2017YFA0604001). The computations in this study were performed at the UMBC High Performance Computing Facility (HPCF). The facility is supported by the U.S. National Science Foundation through the MRI program (Grants CNS-0821258 and CNS-1228778) and the SCREMS program (Grant DMS-0821311), with substantial support from UMBC.

Figures:

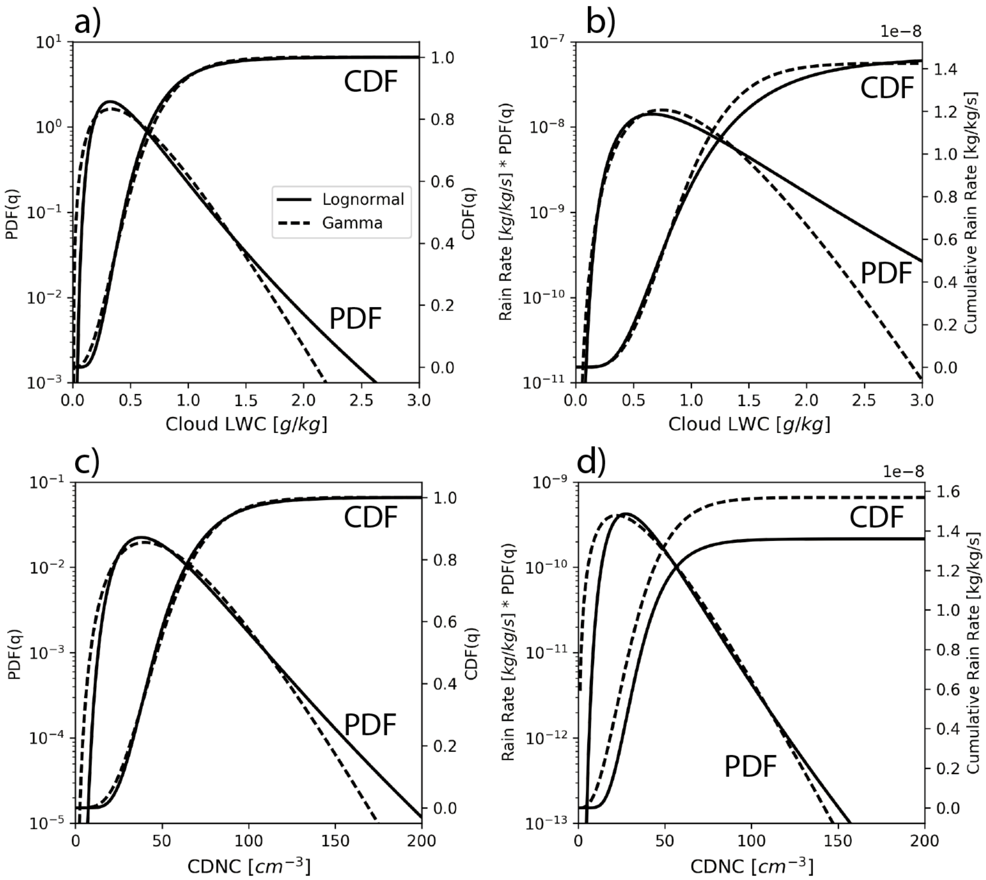

Figure 1 a) The probability density function (PDF) and cumulative distribution function (CDF) of cloud LWC (q_c) that follow the Gamma (dashed) and Lognormal (solid) distribution. For the both distributions, $\langle q_C \rangle = 0.5g/kg$ and $v_q = 3.0$. b) The PDF and CDF of autoconversion rate computed based on the KK2000 scheme in Eq.(15) and the PDF of q_c. In the computation, the N_c is kept at a constant of 50 cm^{-1}. c) The PDF and CDF of N_c that follow the Gamma (dashed) and Lognormal (solid) distribution. For the both distributions, $\langle N_c \rangle = 50cm^{-3}$ and $v_N = 5.0$. d) the PDF and CDF of the autoconversion rate computed based on the KK2000 scheme in Eq. (15) and the PDF of N_c. The q_c is kept at $0.5g/kg$ in the computation.

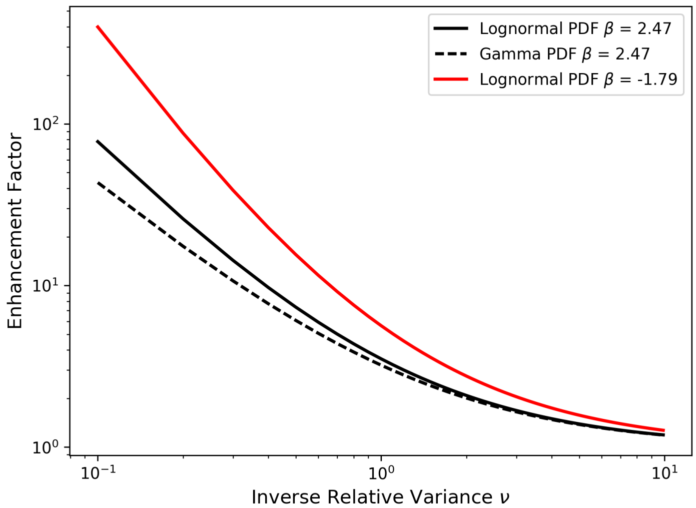

Figure 2 Enhancement factors based on Lognormal $E(P_L, \beta)$ and Gamma $E(P_G, \beta)$ subgrid PDF for different β as a function of the inverse relative variance ν.

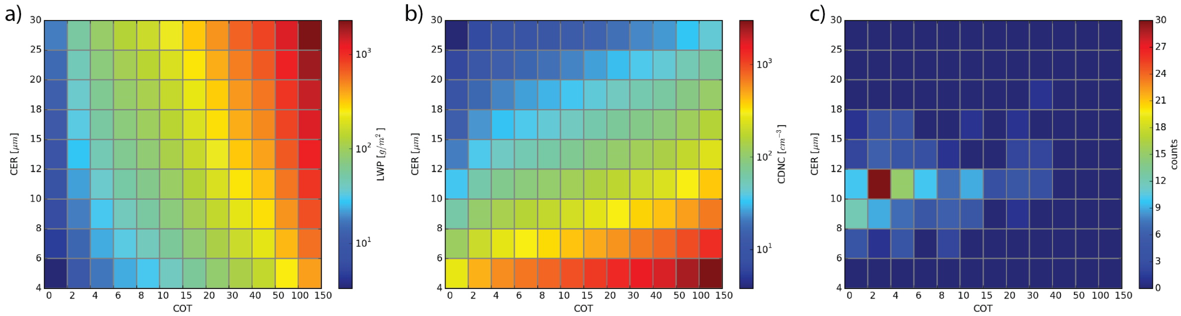

Figure 3 The (a) LWP and (b) CDNC as a function of COT and CER. (c) An exmaple of the COT-CER joint histogram observed by Aqua-MODIS on Jan. 09th, 2007 at 1°S and 1°W.

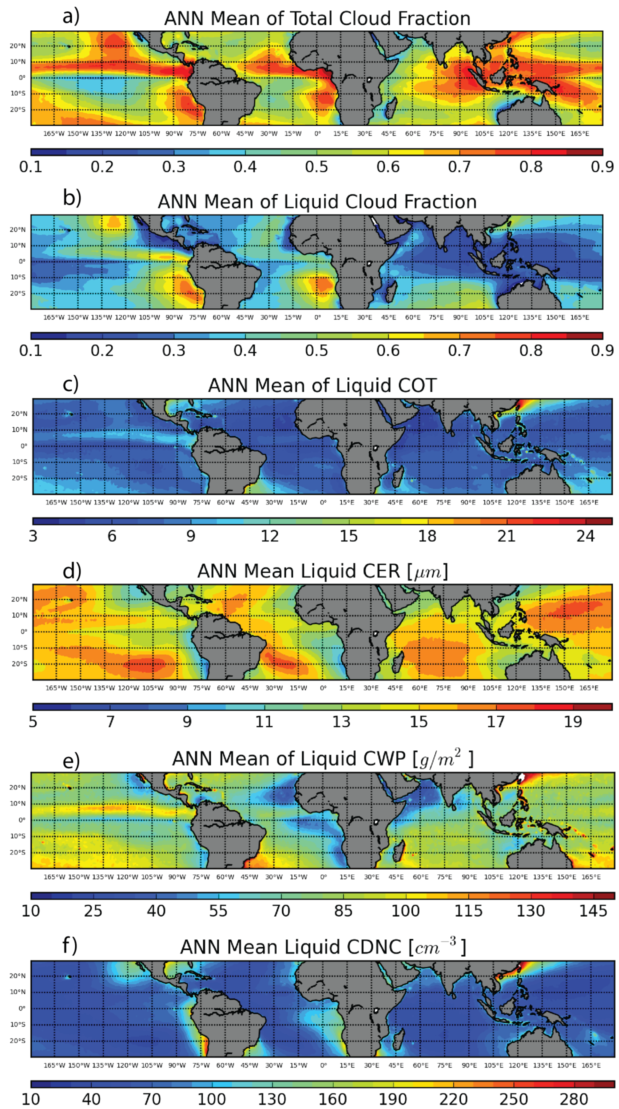

Figure 4 10-year (2007~2016) averaged annual mean a) total cloud fraction, b) liquid cloud fraction, c) cloud optical thickness, d) cloud effective radius retrieved from the 3.7 µm band, e) cloud wather path and f) cloud droplet concentration retrievals from Aqua-MODIS over the tropical (30° S-30° N) oceans. All quantaties are "in-cloud" mean that are averaged over the cloudy-part of the grid only.

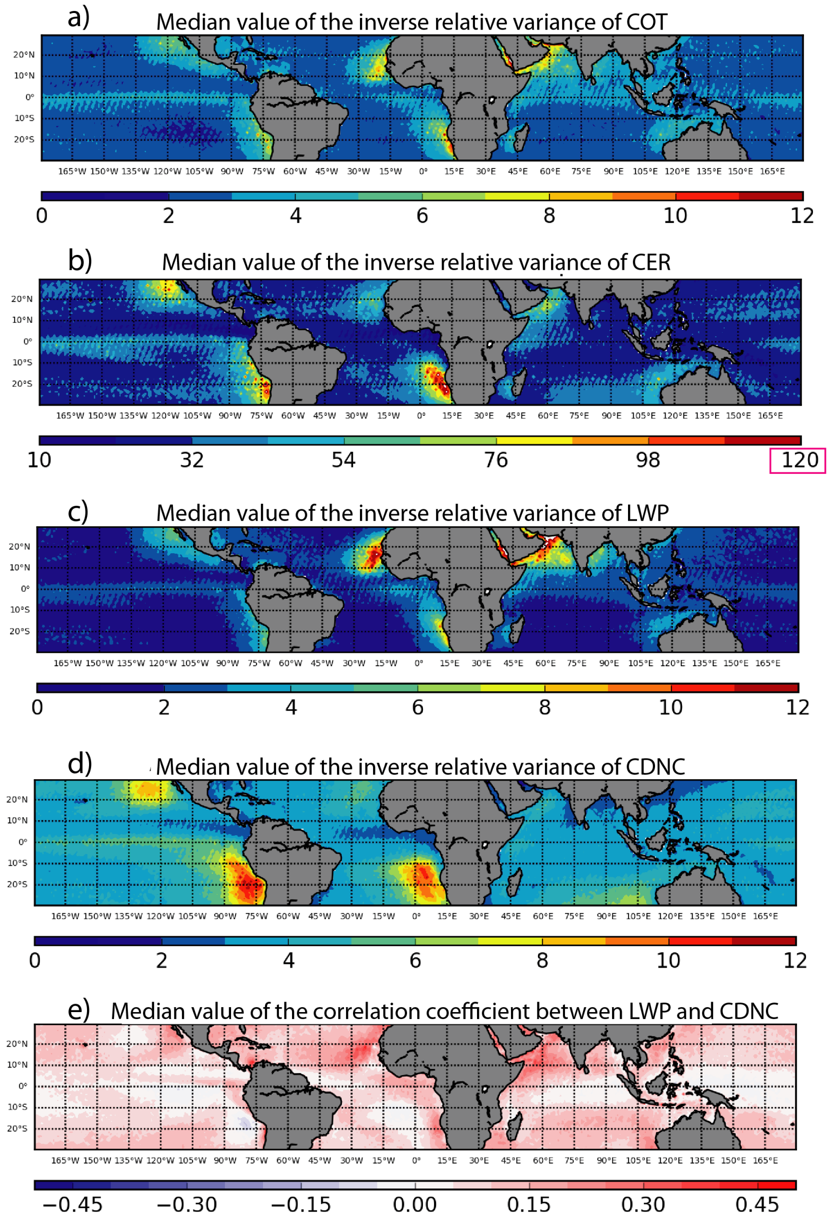

Figure 5 Median value of the inverse relative variance (i.e., $v = \langle x \rangle^2 / Var(x)$) for a) COT, b) CER, c) LWP and d) CDNC, and e) median value of the correlation coefficient between LWP and CDNC derived from 10 years of MODIS observations. Note that the color scale of CER is different from others'.

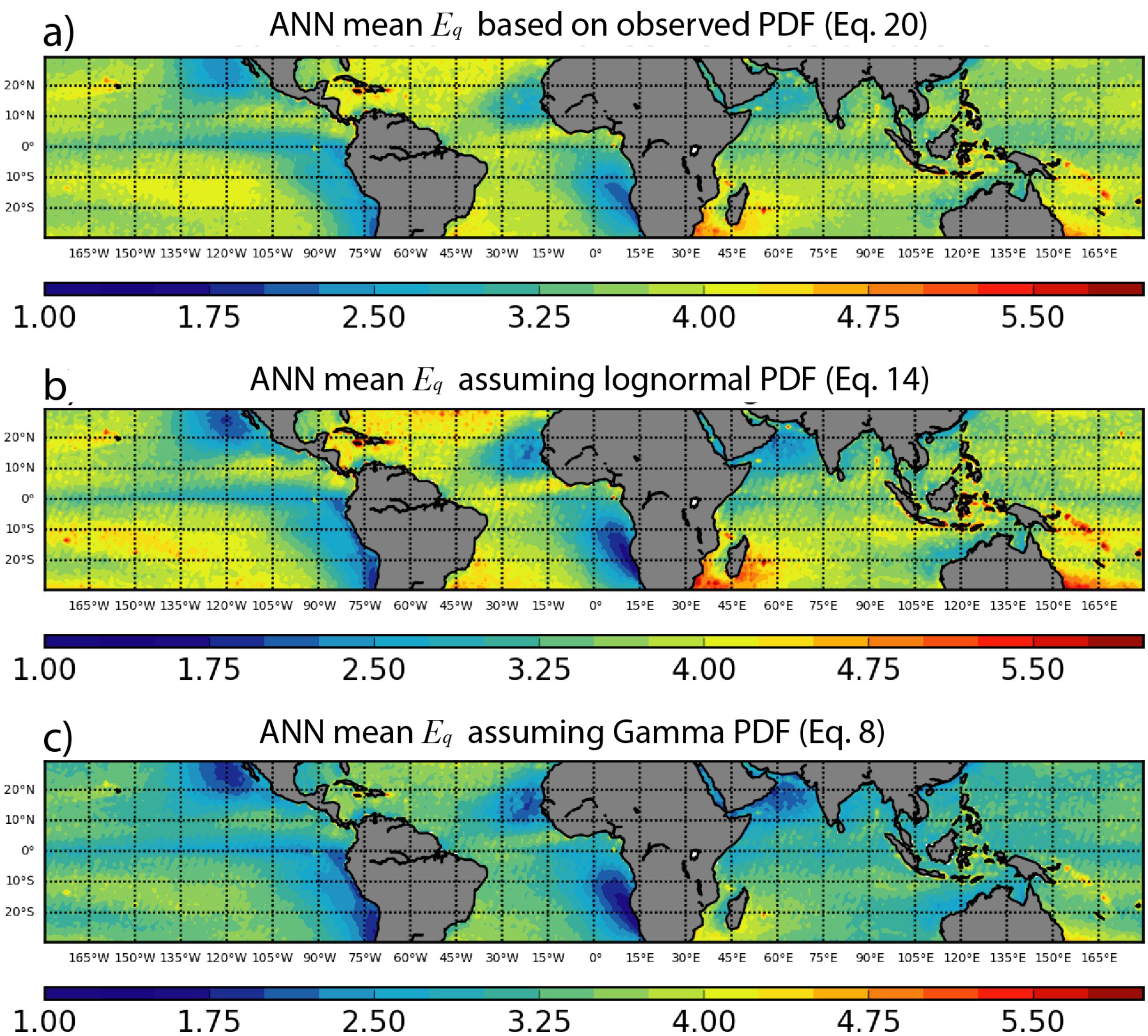

Figure 6 The annual mean factor for the KK2000 scheme due to subgrid variation of LWP computed a) directly from observation, i.e., E_q in Eq.(20), b) from relative variance assuming Lognormal PDF of LWP, i.e., E_q in Eq.(14) and c) from relative variance assuming the Gamma PDF of LWP i.e., E_q in Eq.(8).

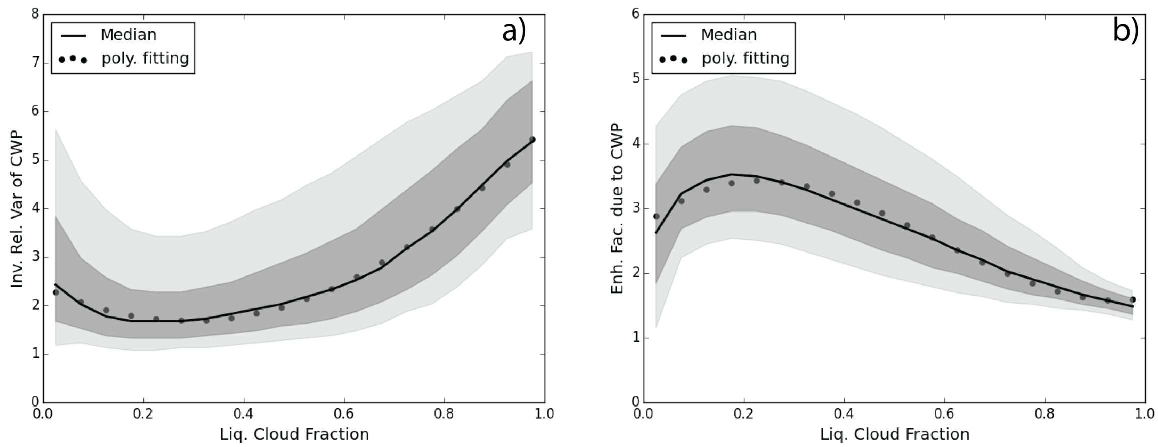

Figure 7 a) The inverse relative variance ν and b) autoconversion enhancement factor due to LWP subgrid variability assuming Log-normal PDF as a function of grid-mean liquid cloud fraction, where the solid line, dark shaded area, and light shaded area correspond to the median value, 25%~75% percentiles, and 10~90% percentiles, respectively. The dotted lines correspond to simple 3-rd order polynomial fitting.

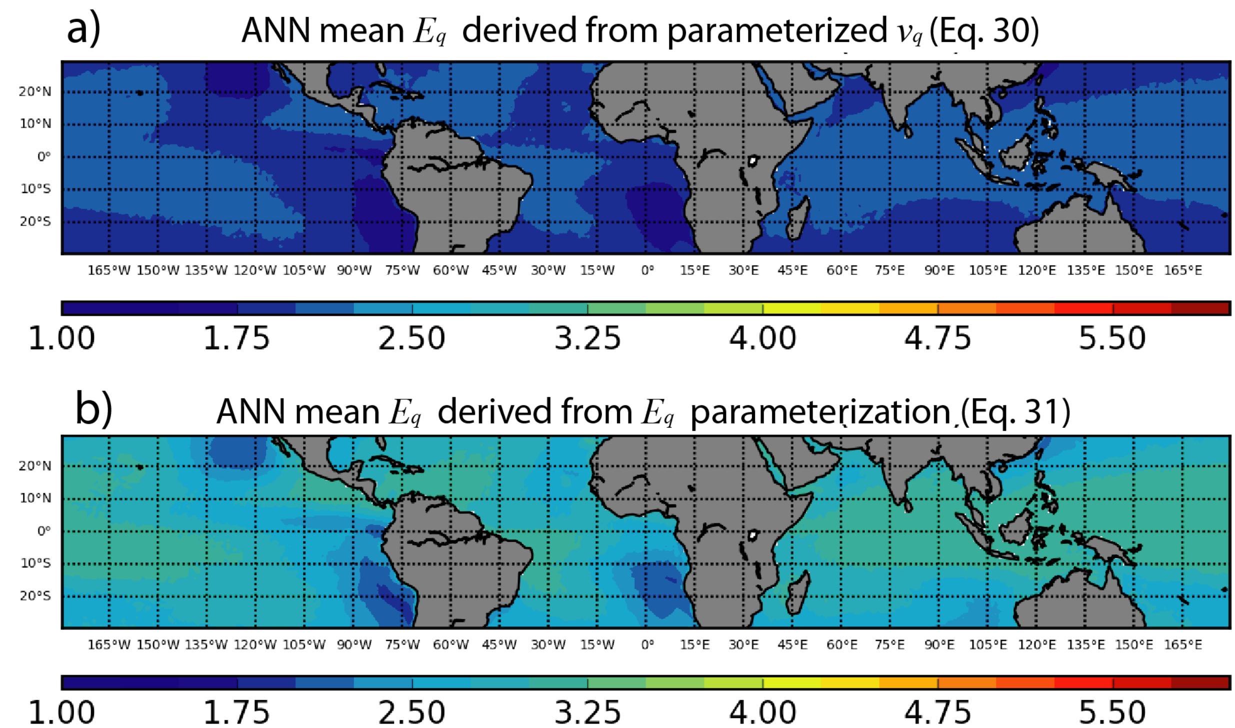

Figure 8 Annual mean value of the enhancement factor E_N *computed based on the a)* $v(f_{liq}) = 2.38-4.95f_{liq} + 8.74f_{liq}^2 - 0.49f_{liq}^3$ *parameterization scheme in Eq.* (30) *and b)* $E_q(f_{liq}) = 2.72+7.33f_{liq} - 19.17f_{liq}^2 + 10.69f_{liq}^3$ *parameterization scheme in Eq.* (31).

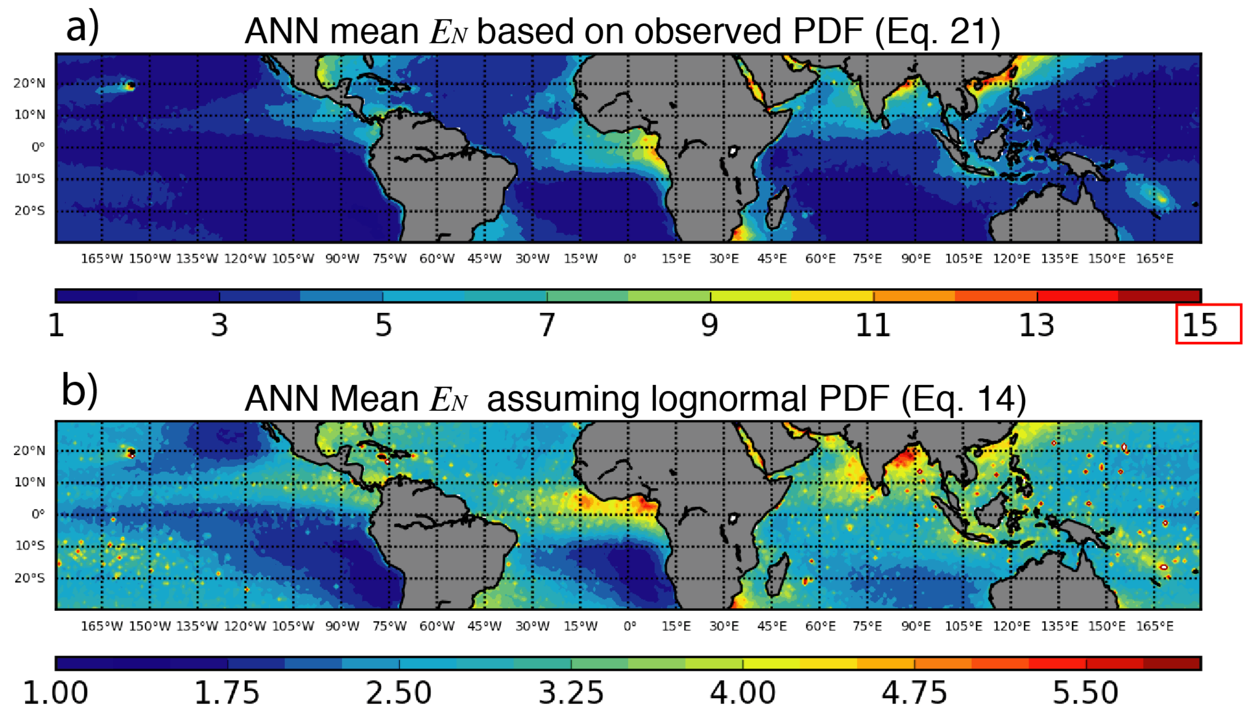

Figure 9 Annual mean value of the enhancement factor E_N *derived from a) observation based on Eq.* (21) *and b) from Eq.* (14) *assuming Lognormal subgrid CDNC distribution.*

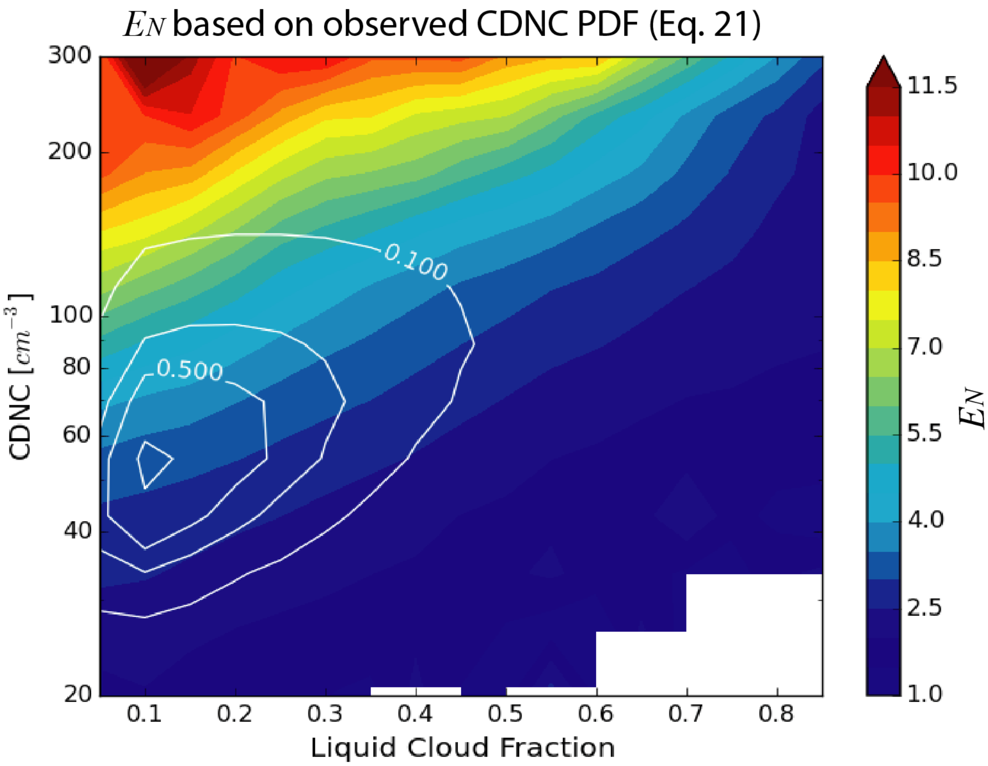

Figure 10 Dependence of E_N on f_{liq} and N_d. The color map corresponds to the mean value of E_N for a given N_d and f_{liq} bin. The white contour lines correspond to the relative sampling frequency of N_d and f_{liq} bins (i.e., the most frequently observed combination is $N_d \sim 50cm^{-3}$ and f_{liq}~0.1).

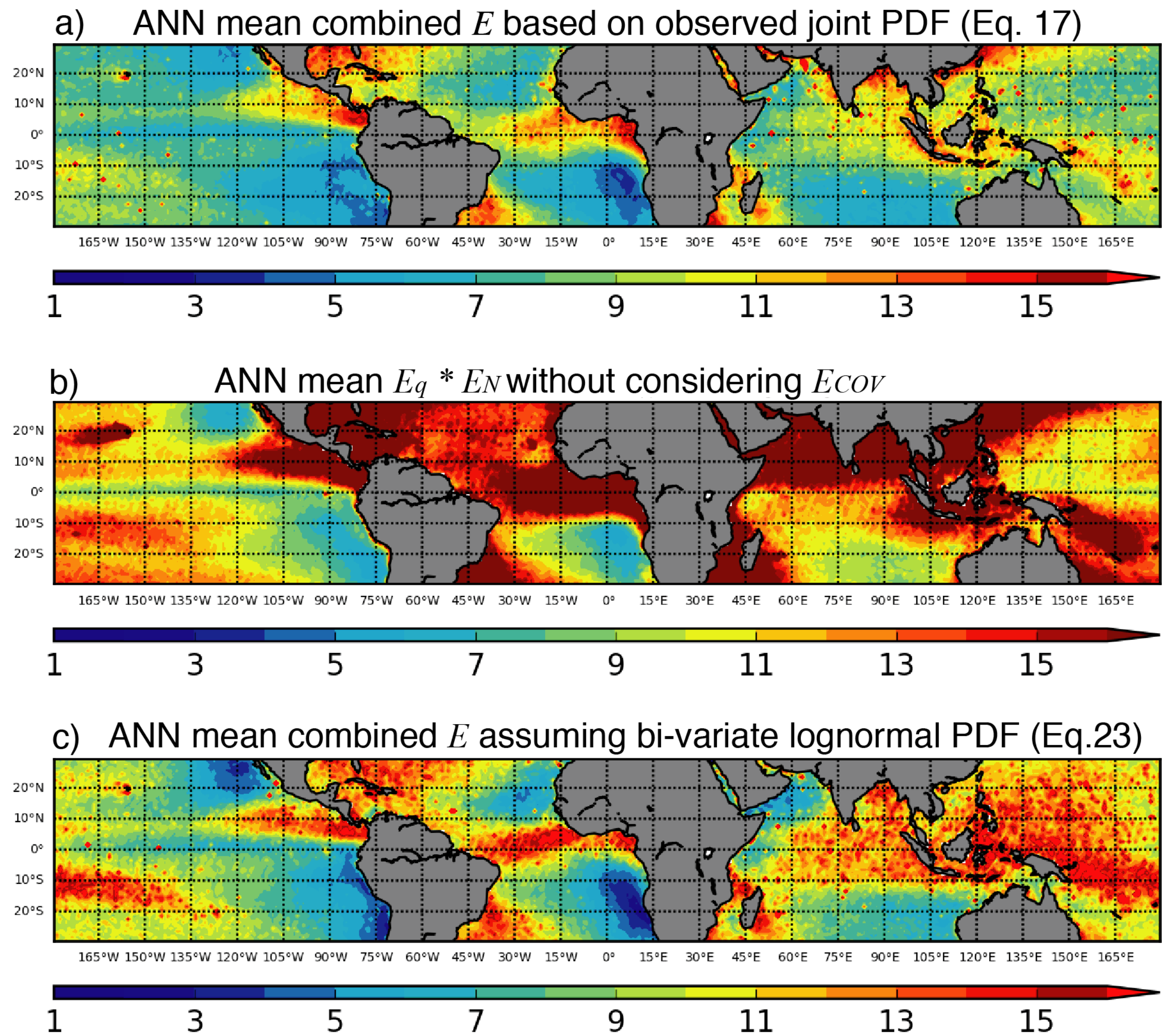

Figure 11 The combined enhancement factor derived a) based on Eq. (17) from the observed joint PDF of LWP and CDNC, b) assuming that subgrid variations of LWP and CDNC are uncorrelated, i.e., $E_q \cdot E_N$ only and c) based on Eq. (23) *assuming that the subgrid LWP and CDNC following the bi-variate lognormal distribution.*

Reference:

Ackerman, S., Strabala, K., Menzel, W., Frey, R., Moeller, C. and Gumley, L.: Discriminating clear sky from clouds with MODIS, Journal of Geophysical Research, 103(D24), 32,141–32,157, 1998.

Ahlgrimm, M. and Forbes, R. M.: Regime dependence of cloud condensate variability observed at the Atmospheric Radiation Measurement Sites, Quarterly Journal of the Royal Meteorological Society, 142(697), 1605–1617, doi:10.1002/qj.2783, 2016.

Barker, H. W.: A Parameterization for Computing Grid-Averaged Solar Fluxes for Inhomogeneous Marine Boundary Layer Clouds. Part I: Methodology and Homogeneous Biases, J. Atmos. Sci., 53(16), 2289–2303, doi:10.1175/1520-0469(1996)053<2289:APFCGA>2.0.CO;2, 1996.

Barker, H. W., Wiellicki, B. A. and Parker, L.: A Parameterization for Computing Grid-Averaged Solar Fluxes for Inhomogeneous Marine Boundary Layer Clouds. Part II: Validation Using Satellite Data, http://dx.doi.org/10.1175/1520-0469(1996)053<2304:APFCGA>2.0.CO;2, 53(16), 2304–2316 [online] Available from: http://journals.ametsoc.org/doi/pdf/10.1175/1520-0469%281996%29053%3C2304%3AAPFCGA%3E2.0.CO%3B2, 1996.

Bennartz, R.: Global assessment of marine boundary layer cloud droplet number concentration from satellite, Journal of Geophysical Research-Atmospheres, 2007.

Bennartz, R. and Rausch, J.: Global and regional estimates of warm cloud droplet number concentration based on 13 years of AQUA-MODIS observations, Atmospheric Chemistry and Physics, 1–32, doi:10.5194/acp-2016-1130, 2017.

Bogenschutz, P. A., Gettelman, A., Hannay, C., Larson, V. E., Neale, R. B., Craig, C. and Chen, C.-C.: The Path to CAM6: Coupled Simulations with CAM5.4 and CAM5.5, Geosci. Model Dev., 1–38, doi:10.5194/gmd-2017-129, 2017.

Bogenschutz, P. A., Gettelman, A., MORRISON, H., Larson, V. E., Craig, C. and Schanen, D. P.: Higher-Order Turbulence Closure and Its Impact on Climate Simulations in the Community Atmosphere Model, J. Climate, 26(23), 9655–9676, doi:10.1175/JCLI-D-13-00075.1, 2013.

Bony, S. and Dufresne, J.-L.: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models, Geophysical Research Letters, 32(20), L20806, doi:10.1029/2005GL023851, 2005.

Boutle, I. A., Abel, S. J., Hill, P. G. and Morcrette, C. J.: Spatial variability of liquid cloud and rain: observations and microphysical effects, Quarterly Journal of the Royal Meteorological Society, 140(679), 583–594, doi:10.1002/qj.2140, 2014.

Cahalan, R. F., Ridgway, W., Wiscombe, W. J., Bell, T. L. and Snider, J. B.: The Albedo of Fractal Stratocumulus Clouds, J. Atmos. Sci., 51(16), 2434–2455, doi:10.1175/1520-0469(1994)051<2434:TAOFSC>2.0.CO;2, 1994.

Cho, H. M., Zhang, Z., Meyer, K., Lebsock, M., Platnick, S., Ackerman, A. S., Di Girolamo, L., C Labonnote, L., Cornet, C., Riedi, J. and Holz, R. E.: Frequency and causes of failed MODIS cloud property retrievals for liquid phase clouds over global oceans, Journal of Geophysical Research-Atmospheres, 120(9), 2015JD023161–n/a, doi:10.1002/2015JD023161, 2015.

Considine, G., Curry, J. A. and Wielicki, B.: Modeling cloud fraction and horizontal variability in marine boundary layer clouds, J. Geophys. Res., 102(D12), 13517–13525, 1997.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J. and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9(5), 1937–1958, doi:10.5194/gmd-9-1937-2016, 2016.

Golaz, J.-C., Larson, V. E. and Cotton, W. R.: A PDF-Based Model for Boundary Layer Clouds. Part I: Method and Model Description, JAS, 59(24), 3540–3551, doi:10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2, 2002a.

Golaz, J.-C., Larson, V. E. and Cotton, W. R.: A PDF-Based Model for Boundary Layer Clouds. Part II: Model Results, http://dx.doi.org/10.1175/1520-0469(2002)059<3552:APBMFB>2.0.CO;2, 59(24), 3552–3571, doi:10.1175/1520-0469(2002)059<3552:APBMFB>2.0.CO;2, 2002b.

Griffin, B. M. and Larson, V. E.: Analytic upscaling of a local microphysics scheme. Part II: Simulations, Quarterly Journal of the Royal Meteorological Society, 139(670), 58–69, doi:10.1002/qj.1966, 2013.

Grosvenor, D. P. and Wood, R.: The effect of solar zenith angle on MODIS cloud optical and microphysical retrievals within marine liquid water clouds, Atmospheric Chemistry and Physics, 14(14), 7291–7321, doi:10.5194/acpd-14-303-2014, 2014.

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., Cairns, B., Chiu, C., Christensen, M., Diamond, M., Feingold, G., Fridlind, A., Hunerbein, A., Knist, C., Kollias, P., Marshak, A., McCoy, D., Merk, D., Painemal, D., Rausch, J., Rosenfeld, D., Russchenberg, H., Seifert, P., Sinclair, K., Stier, P., vanDiedenhoven, B., Wendisch, M., Werner, F., Wood, R., Zhang, Z. and Quaas, J.: **Remote sensing of droplet number concentration in warm clouds: A review of the current state of knowledge and perspectives**, Reviews of Geophysics, (in review), 2018.

Gultepe, I. and Isaac, G. A.: Aircraft observations of cloud droplet number concentration: Implications for climate studies, Quarterly Journal of the Royal Meteorological Society, 130(602), 2377–2390, doi:10.1256/qj.03.120, 2004.

Guo, H., Golaz, J. C., Donner, L. J., Larson, V. E., Schanen, D. P. and Griffin, B. M.: Multi-variate probability density functions with dynamics for cloud droplet activation in large-scale models: single column tests, Geosci. Model Dev., 3(2), 475–486, doi:10.5194/gmd-3-475-2010, 2010.

Guo, H., Golaz, J. C., Donner, L. J., Wyman, B., Zhao, M. and Ginoux, P.: CLUBB as a unified cloud parameterization: Opportunities and challenges, Geophysical Research Letters, 42(11), 4540–4547, doi:10.1002/2015GL063672, 2015.

Guo, Z., Wang, M., Qian, Y., Larson, V. E., Ghan, S., Ovchinnikov, M., Bogenschutz, P. A., Zhao, C., Lin, G. and Zhou, T.: A sensitivity analysis of cloud properties to CLUBB parameters in the single-column Community Atmosphere Model (SCAM5), J. Adv. Model. Earth Syst., 6(3), 829–858, doi:10.1002/2014MS000315, 2014.

Hill, P. G., Morcrette, C. J. and Boutle, I. A.: A regime-dependent parametrization of subgrid-scale cloud water content variability, Quarterly Journal of the Royal Meteorological Society, 141(691), 1975–1986, doi:10.1002/qj.2506, 2015.

Kawai, H. and Teixeira, J.: Probability Density Functions of Liquid Water Path and Cloud Amount of Marine Boundary Layer Clouds: Geographical and Seasonal Variations and Controlling Meteorological Factors, http://dx.doi.org/10.1175/2009JCLI3070.1, 23(8), 2079–2092, doi:10.1175/2009JCLI3070.1, 2010.

Khairoutdinov, M. and Kogan, Y.: A New Cloud Physics Parameterization in a Large-Eddy Simulation Model of Marine Stratocumulus, Mon. Wea. Rev, 128(1), 229–243 [online] Available from: http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(2000)128%3C0229%3AANCPPI%3E2.0.CO%3B2, 2000.

Klein, S. and Hartmann, D.: The seasonal cycle of low stratiform clouds, Journal of Climate, 6(8), 1587–1606, 1993.

Kubar, T. L., Stephens, G. L., Lebsock, M., Larson, V. E. and Bogenschutz, P. A.: Regional Assessments of Low Clouds against Large-Scale Stability in CAM5 and CAM-CLUBB Using MODIS and ERA-Interim Reanalysis Data, J. Climate, 28(4), 1685–1706, doi:10.1175/JCLI-D-14-00184.1, 2014.

Larson, V. E. and Griffin, B. M.: Analytic upscaling of a local microphysics scheme. Part I: Derivation, Quarterly Journal of the Royal Meteorological Society, 139(670), 46–57, doi:10.1002/qj.1967, 2013.

Larson, V. E., Golaz, J.-C. and Cotton, W. R.: Small-Scale and Mesoscale Variability in Cloudy Boundary Layers: Joint Probability Density Functions, J. Atmos. Sci., 59(24), 3519–3539, doi:10.1175/1520-0469(2002)059<3519:SSAMVI>2.0.CO;2, 2002.

Larson, V. E., Wood, R., Field, P. R., Golaz, J.-C., Vonder Haar, T. H. and Cotton, W. R.: Systematic Biases in the Microphysics and Thermodynamics of Numerical Models That Ignore Subgrid-Scale

Variability, J. Atmos. Sci., 58(9), 1117–1128, doi:10.1175/1520-0469(2001)058<1117:SBITMA>2.0.CO;2, 2001.

Lebsock, M. D., L'Ecuyer, T. S. and Stephens, G. L.: Detecting the Ratio of Rain and Cloud Water in Low-Latitude Shallow Marine Clouds, Journal of Applied Meteorology and Climatology, 50(2), 419–432, doi:10.1175/2010JAMC2494.1, 2011.

Lebsock, M., MORRISON, H. and Gettelman, A.: Microphysical implications of cloud-precipitation covariance derived from satellite remote sensing, Journal of Geophysical Research-Atmospheres, 118(12), 6521–6533, doi:10.1002/jgrd.50347, 2013.

Lee, S., Kahn, B. H. and Teixeira, J.: Characterization of cloud liquid water content distributions from CloudSat, J. Geophys. Res., 115(D20), D00A23, doi:10.1029/2009JD013272, 2010.

McCoy, D. T., Bender, F. A. M., Grosvenor, D. P., Mohrmann, J. K., Hartmann, D. L., Wood, R. and Field, P. R.: Predicting decadal trends in cloud droplet number concentration using reanalysis and satellite data, Atmospheric Chemistry and Physics, 1–21, 2017a.

McCoy, D. T., Bender, F. A. M., Mohrmann, J. K. C., Hartmann, D. L., Wood, R. and Grosvenor, D. P.: The global aerosol-cloud first indirect effect estimated using MODIS, MERRA, and AeroCom, Journal of Geophysical Research-Atmospheres, 122(3), 1779–1796, doi:10.1002/2016JD026141, 2017b.

Menzel, P., Frey, R., Baum, B. and Zhang, H.: Cloud Top Properties and Cloud Phase Algorithm Theoretical Basis Document. 2006.

Menzel, W., Smith, W. and Stewart, T.: Improved Cloud Motion Wind Vector and Altitude Assignment Using VAS, Journal of Applied Meteorology, 22(3), 377–384, 1983.

Morales, R. and Nenes, A.: Characteristic updrafts for computing distribution-averaged cloud droplet number and stratocumulus cloud properties, J. Geophys. Res., 115(D18), 1227, 2010.

Morrison, H. and Gettelman, A.: A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, version 3 (CAM3). Part I: Description and numerical tests, Journal of Climate, 2008.

Nakajima, T. and King, M. D.: Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part I: Theory, J. Atmos. Sci., 47(15), 1878–1893, doi:10.1175/1520-0469(1990)047<1878:DOTOTA>2.0.CO;2, 1990.

Nam, C., Bony, S., Dufresne, J. L. and Chepfer, H.: The “too few, too bright”tropical low-cloud problem in CMIP5 models, Geophysical Research …, doi:10.1029/2012GL053421, 2012.

Oreopoulos, L. and Barker, H. W.: Accounting for subgrid-scale cloud variability in a multi-layer 1d solar radiative transfer algorithm, Quarterly Journal of the Royal Meteorological Society, 125(553), 301–330, doi:10.1002/qj.49712555316, 1999.

Oreopoulos, L. and Cahalan, R. F.: Cloud Inhomogeneity from MODIS, Journal of Climate, 18(23), 5110–5124, doi:10.1175/JCLI3591.1, 2005.

Oreopoulos, L. and Davies, R.: Plane Parallel Albedo Biases from Satellite Observations. Part I: Dependence on Resolution and Other Factors, Journal of Climate, 11(5), 919–932, 1998a.

Oreopoulos, L. and Davies, R.: Plane Parallel Albedo Biases from Satellite Observations. Part II: Parameterizations for Bias Removal, J. Climate, 11(5), 933–944, 1998b.

Pincus, R. and Klein, S. A.: Unresolved spatial variability and microphysical process rates in large-scale models, J. Geophys. Res., 105(D22), 27059–27065, doi:10.1029/2000JD900504, 2000.

Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riédi, J. C. and Frey, R. A.: The MODIS cloud products: algorithms and examples from Terra, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 41(2), 459–473, doi:10.1109/TGRS.2002.808301, 2003.

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L. and Riedi, J.: The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 55(1), 502–525, doi:10.1109/TGRS.2016.2610522, 2017.

Pruppacher, H. R. and Klett, J. D.: Microphysics of Clouds and Precipitation: With an Introduction to Cloud Chemistry and Cloud Electricity, 954 pp. 1997.

Randall, D., Khairoutdinov, M., Arakawa, A. and Grabowski, W.: Breaking the Cloud Parameterization Deadlock, Bulletin of the American Meteorological Society, 84(11), 1547–1564, doi:10.1175/BAMS-84-11-1547, 2003.

Seethala, C. and Horváth, Á.: Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds, J Geophys Res, 115(D13), D13202, 2010.

Soden, B. and Held, I.: An assessment of climate feedbacks in coupled ocean–atmosphere models, Journal of Climate, 2006.

Song, H., Song, H., Zhang, Z., Ma, P.-L., Ghan, S. J. and Wang, M.: An Evaluation of Marine Boundary Layer Cloud Property Simulations in the Community Atmosphere Model Using Satellite Observations: Conventional Subgrid Parameterization versus CLUBB, Journal of Climate, 31(6), 2299–2320, doi:10.1175/JCLI-D-17-0277.1, 2018a.

Song, H., Zhang, Z., Ma, P.-L., Ghan, S. J. and Wang, M.: An Evaluation of Marine Boundary Layer Cloud Property Simulations in the Community Atmosphere Model Using Satellite Observations: Conventional Subgrid Parameterization versus CLUBB, Journal of Climate, 31(6), 2299–2320, doi:10.1175/JCLI-D-17-0277.1, 2018b.

Šušelj, K., Teixeira, J. and Chung, D.: A Unified Model for Moist Convective Boundary Layers Based on a Stochastic Eddy-Diffusivity/Mass-Flux Parameterization, J. Atmos. Sci., 70(7), 1929–1953, doi:10.1175/JAS-D-12-0106.1, 2013.

Takahashi, H., Lebsock, M., Suzuki, K., Stephens, G. and Wang, M.: An investigation of microphysics and subgrid-scale variability in warm-rain clouds using the A-Train observations and a multiscale modeling framework, Journal of Geophysical Research-Atmospheres, 138(669), 2151, 2017.

Thayer-Calder, K., Gettelman, A., Craig, C., Goldhaber, S., Bogenschutz, P. A., Chen, C. C., Morrison, H., Höft, J., Raut, E., Griffin, B. M., Weber, J. K., Larson, V. E., Wyant, M. C., Wang, M., Guo, Z. and Ghan, S. J.: A unified parameterization of clouds and turbulence using CLUBB and subcolumns in the Community Atmosphere Model, Geosci. Model Dev., 8(12), 3801–3821, doi:10.5194/gmd-8-3801-2015, 2015.

Trenberth, K. E., Fasullo, J. T. and Kiehl, J.: Earth's Global Energy Budget, Bull. Amer. Meteor. Soc., 90(3), 311–323, doi:10.1175/2008BAMS2634.1, 2009.

Wang, M., Larson, V. E., Ghan, S., Ovchinnikov, M., Schanen, D. P., Xiao, H., Liu, X., Rasch, P. and Guo, Z.: A multiscale modeling framework model (superparameterized CAM5) with a higher-order turbulence closure: Model description and low-cloud simulations, J. Adv. Model. Earth Syst., n/a–n/a, doi:10.1002/2014MS000375, 2015.

Wood, R. and Hartmann, D.: Spatial variability of liquid water path in marine low cloud: The importance of mesoscale cellular convection, Journal of Climate, 2006.

Xie, X. and Zhang, M.: Scale-aware parameterization of liquid cloud inhomogeneity and its impact on simulated climate in CESM, Journal of Geophysical Research-Atmospheres, 120(16), 8359–8371, doi:10.1002/2015JD023565, 2015.

Zhang, Z. and Platnick, S.: An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands, J Geophys Res, 116(D20), D20215, doi:10.1029/2011JD016216, 2011.

Zhang, Z., Ackerman, A. S., Feingold, G., Platnick, S., Pincus, R. and Xue, H.: Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: Case studies based on large-eddy simulations, J Geophys Res, 117(D19), D19208–, doi:10.1029/2012JD017655, 2012.

Zhang, Z., Werner, F., Cho, H. M., Wind, G., Platnick, S., Ackerman, A. S., Di Girolamo, L., Marshak, A. and Meyer, K.: A framework based on 2-D Taylor expansion for quantifying the impacts of sub-pixel reflectance variance and covariance on cloud optical thickness and effective radius retrievals based on the bi-spectral method, Journal of Geophysical Research-Atmospheres, 2016JD024837, doi:10.1002/2016JD024837, 2016.