Sensitivity of the Aerosol Indirect Effect to Subgrid Variability in the Cloud Parameterization of the GFDL Atmosphere General Circulation Model AM3

JEAN-CHRISTOPHE GOLAZ
NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

MARC SALZMANN
NOAA/Geophysical Fluid Dynamics Laboratory, and Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey

LEO J. DONNER, LARRY W. HOROWITZ, YI MING, AND MING ZHAO
NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 28 July 2010, in final form 15 November 2010)

ABSTRACT

The recently developed GFDL Atmospheric Model version 3 (AM3), an atmospheric general circulation model (GCM), incorporates a prognostic treatment of cloud drop number to simulate the aerosol indirect effect. Since cloud drop activation depends on cloud-scale vertical velocities, which are not reproduced in present-day GCMs, additional assumptions on the subgrid variability are required to implement a local activation parameterization into a GCM.

This paper describes the subgrid activation assumptions in AM3 and explores sensitivities by constructing alternate configurations. These alternate model configurations exhibit only small differences in their present-day climatology. However, the total anthropogenic radiative flux perturbation (RFP) between present-day and preindustrial conditions varies by \( \pm 50\% \) from the reference, because of a large difference in the magnitude of the aerosol indirect effect. The spread in RFP does not originate directly from the subgrid assumptions but indirectly through the cloud retuning necessary to maintain a realistic radiation balance. In particular, the paper shows a linear correlation between the choice of autoconversion threshold radius and the RFP.

Climate sensitivity changes only minimally between the reference and alternate configurations. If implemented in a fully coupled model, these alternate configurations would therefore likely produce substantially different warming from preindustrial to present day.

1. Introduction

The atmospheric component–Atmospheric Model, version 3 (AM3)–of the new Geophysical Fluid Dynamics Laboratory (GFDL) coupled general circulation model (CM3; Donner et al. 2011) includes a prognostic treatment of cloud drop number concentration. This is a significant departure from the previous version, AM2 (GFDL Global Atmosphere Model Development Team 2004), which separately prescribes cloud drop number for land and ocean grid points. The introduction of prognostic drop number was motivated by the desire to simulate the aerosol indirect effect, a major source of uncertainty in global climate models (e.g., Solomon et al. 2007).

Key to predicting cloud drop number is the representation of the activation process that converts aerosols serving as cloud condensation nuclei (CCN) into cloud droplets. At local cloud scales, aerosol activation is governed by the Köhler theory (e.g., Pruppacher and Klett 1997) and depends on aerosol mass, size distribution, chemical composition, and supersaturation. The supersaturation is itself controlled by cloud-scale vertical motions. General circulation models (GCMs) with their horizontal grid spacing of the order of 100 km are unable to resolve cloud-scale motions. As a result, additional
assumptions regarding the subgrid variability are required to implement a local activation parameterization in a GCM grid box.

We describe the AM3 subgrid activation assumptions and contrast them to existing approaches. We also investigate the sensitivity of AM3 to these underlying assumptions. To limit the scope of this work, we solely focus on the subgrid variability aspect of the activation parameterization. We treat the local activation parameterization that links aerosols and vertical velocity to the nucleation of cloud droplets as given. In AM3, it is based on Ming et al. (2006) with modifications as described in Donner et al. (2011). While the present study focuses on subgrid activation assumptions, we note that explicit or implicit subgrid assumptions also impact the representation of other cloud processes in GCMs, such as radiative and microphysical impacts of vertical cloud overlap assumptions, and horizontal cloud condensate variability (e.g., Pincus and Klein 2000; Barker and Räisänen 2004, 2005; Pincus et al. 2006).

2. Background

For clarity, we distinguish between a local activation parameterization and a gridbox activation parameterization. We define a local parameterization as one that is applicable at cloud scale. It predicts the number of activated cloud drops based on aerosol properties and supersaturation following the Köhler theory. The dependence on supersaturation is often recast in terms of cloud updraft velocity as proposed by Twomey (1959). We define a gridbox activation parameterization as a local parameterization that has been modified for GCM application. We call “subgrid assumptions” the set of assumptions used to convert a local parameterization into a gridbox parameterization. Subgrid assumptions typically incorporate the effect of subgrid variability of vertical velocity.

A survey of the literature reveals that the treatment of the subgrid vertical velocity generally falls within one of two distinct approaches (see the appendix for a more in-depth review): (i) an explicit assumption about the shape of the vertical velocity probability density function (PDF) is made, and an integration of the local parameterization is performed to obtain the gridbox parameterization; or (ii) the existence of a characteristic vertical velocity is assumed. This velocity is, in general, different from the gridbox average velocity. The gridbox parameterization is obtained by substituting the velocity with the characteristic velocity.

Mathematically, let us assume that \( N^*(a_1, \ldots, a_n, T, p, w) \) is the number of activated aerosols predicted as a function of aerosol properties \( (a_i, i = 1, \ldots, n) \), vertical velocity \( (w) \), and possibly temperature \( (T) \) and pressure \( (p) \) by a local activation parameterization. Similarly, let \( \bar{N}^* \) denote the number predicted by the gridbox parameterization; \( \bar{N}^* \) represents the cloud drop number that can be nucleated within the grid box. The first approach gives

\[
\bar{N}^* = \int N^*(a_1, \ldots, a_n, T, p, w) \, \text{pdf}(w) \, dw, \tag{1}
\]

where pdf\((w)\) is the subgrid distribution of \( w \). The integration is carried over the cloudy portion of the model grid box. Furthermore, because no nucleation occurs in downdrafts, the integration is typically performed for \( w > 0 \). The second approach translates to

\[
\bar{N}^* = N^*(a_1, \ldots, a_n, T, p, \bar{w}), \tag{2}
\]

with \( \bar{w} \) denoting the characteristic vertical velocity.

To compute cloud microphysical and radiative properties, GCMs must calculate the gridbox mean cloud drop number, \( \bar{N}_d \), based on the number that can be activated, \( \bar{N}^* \). Different GCMs follow different approaches. Some models take a diagnostic perspective and assume local equilibrium (e.g., Boucher and Lohmann 1995; Chuang et al. 1997; Jones et al. 2001; Menon et al. 2002; Dufresne et al. 2005),

\[
\bar{N}_d = f(\bar{N}^*), \tag{3}
\]

while others implement a fully predictive equation for \( \bar{N}_d \) (e.g., Ghan et al. 1997; Lohmann et al. 1999b; Storelmo et al. 2006; Ming et al. 2007; Gettelman et al. 2008). It takes the following general form:

\[
\frac{\partial \bar{N}_d}{\partial t} = \frac{\partial \bar{N}_d}{\partial t}_{\text{adv}} + \frac{\partial \bar{N}_d}{\partial t}_{\text{mix}} + \frac{\partial \bar{N}_d}{\partial t}_{\text{activation}} + \frac{\partial \bar{N}_d}{\partial t}_{\text{sinks}}, \tag{4}
\]

where the first two terms on the RHS represent the change in cloud drop number due to advection and mixing (turbulent and convective). The third term is the source of cloud drop from nucleation. It is a function of \( \bar{N}^* \). The last term denotes various sinks, such as evaporation and conversion to rain and ice.

3. AM3 subgrid nucleation assumptions

We now summarize the subgrid assumptions to implement the local activation parameterization of Ming et al. (2006) as a gridbox parameterization in AM3. A similar approach also applies to a version of AM3 that incorporates a two-moment bulk microphysics (Salzmann et al. 2010).
The maximum number of cloud drops that can be activated, \( N_d^* \), is computed based on an integration over the subgrid \( w \) distribution following Eq. (1). The PDF of \( w \) is assumed to follow a normal distribution. Its mean is identical to the vertical velocity that drives large-scale condensation in the Tiedtke (1993) cloud scheme of AM3, which is the large-scale velocity outside the convective area modified to include the impact of the turbulent and radiative cooling rates. The integration is carried out numerically using a 64-point Gauss–Hermite quadrature. The standard deviation of the distribution, \( \sigma_w \), is related to the eddy diffusivity coefficient \( K_h \) as first suggested by Ghan et al. (1997):

\[
\sigma_w = \frac{2}{\sqrt{3} c_\mu \Delta z} \left( \frac{K_h}{\tau_w, \min} \right),
\]

with \( c_\mu = 0.09 \) derived from turbulence measurements (Duynkerke and Driedonks 1987) and used here to relate \( K_h \) to the turbulence kinetic energy (TKE) using the gridbox height \( \Delta z \) as length scale. The \( \sqrt{3} \) factor converts TKE to vertical velocity variance assuming isotropy. The lower bound, \( \sigma_w, \min \), is set to 0.7 m s\(^{-1}\). It is larger than other studies [0.1 m s\(^{-1}\) in Ghan et al. (1997), 0.2 m s\(^{-1}\) in Ghan et al. (2001b), and 0.3 m s\(^{-1}\) in Storelvmo et al. (2006)], but it is within the range of 0.25–0.75 m s\(^{-1}\) explored by Chuang et al. (1997). We note that in AM3, the frequency of occurrence of the lower bound is 98%. Thus, the activation behaves essentially as if the variance were fixed. The frequency of occurrence is 95% when the minimum is reduced to 0.2 m s\(^{-1}\) (configuration S1 below). One of the motivations for this work is to explore the impact of the minimum velocity variance in AM3.

For large-scale clouds, the source term in the cloud drop number budget Eq. (4) is formulated as in Ming et al. (2007):

\[
\frac{\partial N_d}{\partial t} \bigg|_{\text{activation}} = \left\{ \begin{array}{ll}
N_d^* \frac{\partial C}{\partial t} : & \frac{\partial C}{\partial t} > 0 \\
0 : & \text{otherwise}
\end{array} \right.
\]

where \( \partial C/\partial t \) is the change in cloud amount predicted by the large-scale cloud scheme. This formulation assumes that activation coincides with cloud formation. Condensation in preexisting clouds adds to the growth of existing drops rather than nucleate new ones. This is a departure from some GCM formulations, which allow additional nucleation in preexisting clouds when \( N_d^* \) exceeds \( N_d \) (e.g., Storelvmo et al. 2006; Gettelman et al. 2008):

\[
\frac{\partial N_d}{\partial t} \bigg|_{\text{activation}} = \frac{\max (N_d^* - N_d, 0)}{\Delta t},
\]

where \( \Delta t \) is the model time step. Gettelman et al. (2008) replace \( \Delta t \) with a relaxation time scale of 20 min.

AM3 also includes the nucleation of cloud droplets in its shallow convective parameterization [Bretherton et al. (2004) with modifications described in Zhao et al. (2009) and Donner et al. (2011)]. Nucleation occurs at cloud base using the plume vertical velocity at that level. Cloud drops are transported upward in the plume. When liquid water is detrained, both mass and number are transferred to the large scale. Prediction of cloud drop number is not incorporated in AM3’s deep convective parameterization (Donner 1993).

### 4. Sensitivity experiments

The purpose of this work is to evaluate the impact of the subgrid assumptions embedded in Eqs. (5) and (6). To do so, we construct three alternate model configurations (Table 1). These alternate configurations are designed to test (i) the impact of the lower bound on \( \sigma_w \) by reducing its value [S1 versus reference (REF)]; (ii) the assumption that cloud drops nucleate only in newly formed clouds by allowing nucleation to occur in both new and preexisting clouds, that is, replacing Eq. (6) with (7), (S2 versus REF); and (iii) the difference between subgrid PDF and characteristic velocity formulations (S3 versus S2), where the characteristic velocity, \( \tilde{w} \), is defined similarly to Morrison and Gettelman (2008):

\[
\tilde{w} = \max (\sigma_w, \tilde{w}_{\min})
\]

but with \( \sigma_w \) as in (5) for consistency.

| Config. | \( N_d^* \) | \( \frac{\partial N_d}{\partial t} \bigg|_{\text{activation}} \) |
|---------|-----------|---------------------------|
| REF (AM3) | w PDF, \( \sigma_w, \min = 0.7 \text{ m s}^{-1} \) | New clouds [Eq. (6)] |
| S1 | w PDF, \( \sigma_w, \min = 0.2 \text{ m s}^{-1} \) | New clouds [Eq. (6)] |
| S2 | w PDF, \( \sigma_w, \min = 0.1 \text{ m s}^{-1} \) | New and preexisting clouds [Eq. (7)] |
| S3 | Characteristic \( \tilde{w}, \tilde{w}_{\min} = 0.05 \text{ m s}^{-1} \) [Eq. (8)] | New and preexisting clouds [Eq. (7)] |
Model configuration is similar to the fixed sea surface temperature (SST) simulation presented in Donner et al. (2011), except that interannual variability in boundary conditions and forcings are removed to enable shorter simulations. Interannual monthly SSTs are replaced with monthly climatologies for the period 1980–2000. Emissions of sulfate precursors, organic and black carbon, and greenhouse gas concentrations are held constant at their 1990 values. As in AM3, emissions are based on Lamarque et al. (2010). Impacts of explosive volcanoes are neglected. Results discussed in this section are based on 6-yr experiments following a 1-yr spinup except for some untuned experiments, which are based on 1-yr simulations.

Applied in isolation, formulation changes in Table 1 have a large impact on the predicted cloud drop number and radiation balance, as illustrated in Figs. 1 and 2, from 1-yr simulations. Radiative balance is affected mainly by the absorbed shortwave radiation. Lowering the minimum $\sigma_w$ value from 0.7 to 0.2 m s$^{-2}$ (REF $\rightarrow$ S1u) reduces the cloud droplet number and increases absorbed shortwave by making clouds less reflective. Allowing droplet nucleation to occur in all clouds rather than new clouds only (REF $\rightarrow$ S2u, S3u) increases cloud droplet number because nucleation occurs more frequently. (S1u, S2u, S3u refer to untuned experiments.) This in turn makes clouds more reflective and reduces absorbed shortwave. In light of radiation balance constraints, the relatively high minimum $\sigma_w$ in the reference model is compatible with the assumption that nucleation only occurs in the newly formed clouds, while allowing drop nucleation in all clouds facilitates a lower minimum $\sigma_w$.

It is interesting to note that these formulation changes alone have a radiative impact with a magnitude comparable to the radiative forcing resulting from changes in greenhouse gas concentrations between preindustrial (PI) and present-day (PD) conditions (Solomon et al. 2007). However, configurations S1u, S2u, and S3u could not be used for long-term coupled climate experiments, because the magnitude of their net top-of-the-atmosphere (TOA) radiation fluxes is unrealistically large. Therefore, as in commonly done in GCMs, these configurations are retuned to bring their energy balance in line with the reference configuration. This is accomplished by adjusting parameters from other model components whose exact values are uncertain.

The impact of the retuning process on cloud drop number concentration is illustrated in Fig. 2. The altered model parameters are listed in Table 2. They consist of the autoconversion threshold radius and the erosion constants in the large-scale cloud parameterization.

As surveyed by Rotstayn (2000), the autoconversion threshold is often used to tune GCM models. Values in the range of 4.5–7.5 $\mu$m are common, even though they are unrealistically small. The threshold radius of 10.6 $\mu$m in AM2 is comparable to values in cloud and mesoscale models. This value was reduced to 8 $\mu$m in the atmospheric component AM2.1 of the coupled model CM2.1 (Delworth et al. 2006). The value of the threshold radius in large-scale models, such as GCMs, is often reduced compared to finer-scale models to compensate for the neglect of in-cloud subgrid-scale variability (Pincus and Klein 2000; Larson et al. 2001) and the dispersion effect (increased breadth of the drop size distribution with increased number concentration; Rotstayn and Liu 2005).

The large-scale cloud scheme in AM3 solves prognostic equations for cloud amount, liquid, and ice (Tiedtke 1993). These equations include a sink term (“erosion term”) that governs the rate at which subgrid-scale mixing dissipates clouds in subsaturated grid cells. Tiedtke (1993) suggested a cloud erosion constant of $1 \times 10^{-6}$ s$^{-1}$. AM2 retained this value but also introduced larger values for grid boxes where either convection or vertical turbulence is active. As noted in GFDL Global Atmosphere Model Development
Additional simulation characteristics are shown in Table 2. Model parameters modified in the retuning of alternate configurations (S1, S2, and S3) compared to REF, AM3, AM2, and AM2.1. See GFDL Global Atmosphere Model Development Team (2004) for a discussion of these parameters. AM2.1 refers to the atmospheric component of the coupled model CM2.1 (Delworth et al. 2006). Also shown are values of the TOA net radiation. AM3 and REF are the same model; they differ by the configuration of the simulations and therefore have different TOA radiation balance.

<table>
<thead>
<tr>
<th>Config.</th>
<th>Years</th>
<th>Autoconversion threshold (µm)</th>
<th>Erosion constants (×10^{-6} s^{-1})</th>
<th>TOA net radiation (W m^{-2})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Main</td>
<td>Convective</td>
</tr>
<tr>
<td>AM2</td>
<td>1983–98</td>
<td>10.6</td>
<td>1.0</td>
<td>4.7</td>
</tr>
<tr>
<td>AM2.1</td>
<td>1983–98</td>
<td>8.0</td>
<td>1.0</td>
<td>8.0</td>
</tr>
<tr>
<td>AM3</td>
<td>1980–2000</td>
<td>8.2</td>
<td>1.3</td>
<td>70.0</td>
</tr>
<tr>
<td>REF</td>
<td>6</td>
<td>8.2</td>
<td>1.3</td>
<td>70.0</td>
</tr>
<tr>
<td>S1</td>
<td>6</td>
<td>10.2</td>
<td>1.3</td>
<td>70.0</td>
</tr>
<tr>
<td>S2</td>
<td>6</td>
<td>7.5</td>
<td>3.7</td>
<td>200.0</td>
</tr>
<tr>
<td>S3</td>
<td>6</td>
<td>6.5</td>
<td>20.0</td>
<td>200.0</td>
</tr>
</tbody>
</table>

Team (2004), the AM2 erosion of 4.7 × 10^{-6} s^{-1} under convective conditions might be about 40 times smaller than suggested by the analysis of large-eddy simulations (LESs) of trade wind cumulus clouds (Siebesma et al. 2003). Erosion constants in S2 and S3 are closer to the LES’s estimated value.

We emphasize that the specifics of the retuning process are subjective. Configurations listed in Table 2 are not unique. We also note that the distinction between formulation changes and retuning can be arbitrary. Here, we regard the modifications to the activation as formulation changes, and modifications to the cloud macrophysics (erosion scales) and microphysics (autoconversion threshold) as retuning. One could take an alternate view and regard the cloud macro- and microphysical changes as formulation changes and the modifications to the activation as part of the retuning process. Both reflect uncertainties in GCM parameterizations.

Basic simulation metrics are presented in Table 3. AM3 and REF are identical models but differ by their experiment setup (20 yr with varying boundary conditions versus 6 yr with climatological boundary conditions). AM3 results are from Donner et al. (2011). The REF configuration generally yields lower scores than AM3 owing to the different setup. REF is generally better than alternate configurations S1, S2, and S3, although scores often differ only very slightly.

Additional simulation characteristics are shown in Fig. 3 in the form of annual zonal averages. Despite significant differences in the formulation of cloud drop activation and associated cloud retuning, basic simulation characteristics of these four configurations exhibit remarkably little difference, except for column integrated in-cloud drop number. Indeed, for most fields, intermodel differences are much smaller than differences between model and observations.

Figure 4 illustrates distributions of cloud drop effective radius and number concentration. Similar to Gettelman et al. (2008), distributions are shown for clouds between the surface and 100 hPa and with a minimum in-cloud water content of 2 ppmm. There are large differences between the four configurations. Effective radius distributions for REF and S1, which activate cloud drops only in new clouds, have more occurrences of clouds with low cloud droplet numbers and correspondingly large effective radii. Number distribution peaks at low values for REF and S1, compared to a distinct peak around 15 cm^{-3} for S2 and S3, where activation is allowed to occur in all clouds. From Fig. 4, it is clear that nucleating drops in all or new clouds only has a large impact on global cloud drop distributions. The minimum σ_w value has a comparatively smaller impact (REF versus S1), as does the choice of integrating over a PDF or using a characteristic velocity (S2 versus S3).

It is also interesting to contrast Fig. 4 with Figs. 5a and 5c in Gettelman et al. (2008). The range of drop effective radius is larger in the four configurations tested here, especially for REF and S1. Drop number concentrations tend to be smaller; Gettelman et al. (2008) report a higher frequency of clouds with drop concentrations above 100 cm^{-3}.

Based on the results presented in this section, it would be tempting to conclude that the various activation formulations and related cloud tuning have a decisive impact only on cloud drop number concentration. As we demonstrate in the next section, this is, however, not the case.

5. Impact on RFP and Cess climate sensitivity

We perform additional sensitivity experiments with the four configurations described in the previous section. The first set of experiments uses preindustrial greenhouse gas concentrations and emissions (Lamarque et al. 2010) but with the same sea surface temperatures as the present-day simulations. The difference in global net radiation is defined as the total radiative flux perturbation (Haywood et al. 2009; Lohmann et al. 2010) and can be used to assess the total anthropogenic impact of greenhouse
where $F$ sensitivity is defined as

$$ D = \int_{t_1}^{t_2} F \, dt, $$

for 6 yr after a 1-yr spinup. Both preindustrial and Cess-like experiments integrate the radiation flux in the perturbed sea surface temperature experiments. Between the four configurations, with a range slightly larger than 50% from the reference configuration. Differences are statistically significant between REF and S1 and S3 but not between REF and S2 because of large interannual variability. Compared to RFP, the variation in Cess climate sensitivity $\lambda$ is small. All alternate configurations have a sensitivity that is within 10% of the reference. The sensitivity of the reference configuration (0.73 K W$^{-1}$ m$^2$) is larger than the sensitivity of 0.58 K W$^{-1}$ m$^2$ for AM2. Both values are larger than sensitivities for the Community Atmosphere Model, version 3 (CAM3) with a
semi-Lagrangian dynamical core (0.41 K W$^{-1}$ m$^2$), the CAM3 with an Eulerian dynamical core (0.54 K W$^{-1}$ m$^2$), and the "superparameterized" version of the CAM (SP-CAM) (0.41 K W$^{-1}$ m$^2$) as reported by Wyant et al. (2006).

The surface temperature change index, $\Delta T_{s,i}$, is 0.66 K for REF (AM3), which is larger than the PD – PI temperature change of 0.32 K (1980–2000 minus 1880–1920 global average) in a five-member ensemble of the coupled model CM3 (2011). For AM2, RFP is estimated to be 1.44 W m$^{-2}$ and therefore $\Delta T_{s,i} \approx 0.84$ K, compared to a warming of 0.66 K in a five-member ensemble of CM2.1 (Knutson et al. 2006). Possible hypotheses for the smaller surface temperature change of the coupled model include the lack of equilibrium due to ocean thermal inertia, neglect of explosive volcanoes in our climatological configurations, and—to a lesser degree—model internal variability and simulation lengths. Similar differences have also been observed for the earth when comparing equilibrium temperature change with actual global mean surface temperature change (Schwartz et al. 2010).

RFP values in Table 4 are also plotted in Fig. 6 as a function of the autoconversion threshold selected in the retuning process. Quite remarkably, the four configurations

---

**FIG. 3.** Comparisons of zonal mean model results and observations. Results from the four model configurations (REF, S1, S2, and S3) are represented by the colored curves. Observations are from (a)–(d) Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced And Filled (EBAF) and Earth Radiation Budget Experiment (ERBE) (Harrison et al. 1990); (e) Global Precipitation Climatology Project, version 2 (GPCP v2) and Climate Prediction Center Merged Analysis of Precipitation, version 2 (CMAP v2) (Xie and Arkin 1997); (f) ISCCP (Rossow and Schiffer 1999), and (g) Greenwald et al. (1993) and Weng and Grody (1994).
are almost perfectly aligned along a straight line. To investigate whether this is simply a coincidence or whether there is a real correlation between the autoconversion threshold and RFP, we perform two additional experiments. REF2 is identical to REF but with the threshold reduced from 8.2 to 6.2 μm, and REF1 is with the threshold increased to 10.2 μm. REF2 and REF1 are subject to large TOA energy imbalances and therefore could not be used for coupled climate experiments. RFP values for these two additional configurations are shown in Fig. 7. Again, there is an almost perfect linear correlation between the autoconversion threshold and RFP. Moreover, the linear fit for REF, REF+, REF− (solid line in Fig. 7) is almost identical to the linear fit for REF, S1, S2, and S3 (dashed line in Figs. 6 and 7).

RFP includes the impact of greenhouse gases, and aerosol direct and indirect effects. The autoconversion threshold does not directly affect greenhouse gases or the aerosol direct effect. Although changes in cloud liquid water content could have a small impact on aerosol wet deposition, the correlation between autoconversion threshold and RFP is most likely caused by a modulation of the strength of the indirect effect. This is corroborated by the fact that RFP_{clr} exhibits only small and statistically insignificant differences between the four configurations. The amplitude of the modulation of the indirect effect is of the order of 1 W m^{-2} for the range of threshold values tested here (Table 4). This range approaches the range of −1.8 to −0.3 W m^{-2} for the radiative forcing from the first indirect effect reported by Solomon et al. (2007). Although not unexpected, this is significant given that the autoconversion process is only one of several physical processes that impact the indirect effect. For example, Storelvmo et al. (2009) reported that only changing diagnostic relationships between aerosol mass and cloud drop number can also lead to a large modulation of the strength of the indirect effect from −1.94 to −0.62 W m^{-2}. Hoos et al. (2009) investigated the impact of imposing a minimum cloud droplet or aerosol concentrations in Community Atmosphere Model (CAM)-Oslo. They found a large sensitivity of the indirect effect (−1.88 ± 0.14 to −0.62 ± 0.17 W m^{-2}) to the choice of imposed minimum concentrations. We note that AM3 does not prescribe any minimum cloud droplet or aerosol concentrations.

Examples of autoconversion formulation impact on the magnitude of the indirect effect have been documented before. Lohmann and Feichter (1997) performed sensitivity studies to evaluate the effect of sulfate aerosols on the indirect effect. In one of their tests, they found that by replacing the autoconversion parameterization alone, the magnitude of the indirect effect could change by 2.6 W m^{-2}. Rotstayn (2000) found that introducing a new autoconversion parameterization increased the magnitude of the indirect effect by 0.6 W m^{-2} compared to the original scheme. In a subsequent work, Rotstayn and Liu (2005) incorporated an autoconversion parameterization that accounted for the dispersion effect. They found that this alternate autoconversion led to a reduction in the magnitude of the indirect effect.

Quaas et al. (2009) compared the aerosol indirect effect in 10 different GCMs against satellite observations. Among their findings, they noted that all models overestimate the variations of liquid water path with anthropogenic aerosol optical depth, with the strength of this relationship influenced by the autoconversion parameterization.

The RFP uncertainty range of nearly 1 W m^{-2} documented here is substantially larger than the recent study by Lohmann and Ferrachat (2010). They estimated the impact of parametric uncertainties on the aerosol indirect effect by varying four parameters often used to ensure radiation balance in ECHAM5: rate of rain formation by...
autoconversion, rate of snow formation by aggregation, inhomogeneity factor for ice clouds, and entrainment rate for deep convection. They found an uncertainty range of 25% (0.25 W m\(^{-2}\)) for the total anthropogenic aerosol effect when considering all their experiments and a smaller range of 11% when considering only experiments that satisfied energy balance constraints.

The sensitivity of anthropogenic RFP to the autoconversion threshold can be understood by the formulation of the autoconversion parameterization. In AM3, it is based on Manton and Cotton (1977):

\[
\frac{\partial q_l}{\partial t}_{\text{auto}} = -q_a \frac{0.104 g E \rho^4}{\mu (N \rho)^{7/3}} \left(\frac{q_l}{q_a}\right)^{7/3} H \left(\frac{q_l}{q_a} - q_{\text{crit}}\right),
\]

where \(q_l\) is the gridbox mean liquid water specific humidity, \(q_a\) is the cloud amount, \(N\) is the cloud drop number concentration, \(\rho\) is the air density, \(\rho_l\) is the liquid water density, \(g\) is the acceleration due to gravity, \(\mu\) is the dynamic viscosity of air, and \(E = 0.55\) the mean collection efficiency; \(H\) denotes the Heaviside function. The autoconversion is only active when the in-cloud liquid water is greater than the critical value \(q_{\text{crit}}\) defined as

\[
q_{\text{crit}} = \frac{4}{3} \pi \frac{\rho_l}{\rho} r_{\text{thresh}}^3 N,
\]

where \(r_{\text{thresh}}\) is the threshold radius. Caution must be taken when numerically integrating Eq. (12) with a long GCM time step. Time integration is performed analytically, and a limiter is introduced that prevents in-cloud liquid water from being depleted below \(q_{\text{crit}}\), where the autoconversion would cease to be active (Rotstyn 1997). In practice, the limiter is almost always invoked (S. Klein 2000, personal communication), and the in-cloud liquid water remaining after the autoconversion, \(\tilde{q}_l\), is

\[
\tilde{q}_l = q_{\text{crit}} = \frac{4}{3} \pi \frac{\rho_l}{\rho} r_{\text{thresh}}^3 N.
\]

A perturbation in cloud drop number concentration \(\Delta N\) corresponds to a change in residual in-cloud liquid \(\Delta \tilde{q}_l\):

\[
\Delta \tilde{q}_l = \frac{4}{3} \pi \frac{\rho_l}{\rho} r_{\text{thresh}}^3 \Delta N.
\]

Equation (15) can help understand the RFP sensitivity to \(r_{\text{thresh}}\). The liquid water response to a given perturbation \(\Delta N\) will be the largest when \(r_{\text{thresh}}\) is large. The first aerosol indirect effect is controlled by \(\Delta N\). Assuming that \(\Delta N\) is constrained by emissions changes from PD to PI, the cloud response through the second indirect effect will be strongest (more negative) when \(r_{\text{thresh}}\) is largest, thus offsetting a larger fraction of the greenhouse gases warming and leading to a smaller total anthropogenic RFP value, as observed in Figs. 6 and 7. Differences in liquid water path between PD and PI for the four configurations (Fig. 8) confirm the existence of a linear correlation between \(\Delta LWP_{\text{PD-PI}}\) and \(r_{\text{thresh}}\), as indicated by Eq. (15).

Finally, Fig. 9 shows the surface temperature change index, \(\Delta T_{s,j}\), inferred from RFP and climate sensitivity. Because RFP is much more sensitive to the autoconversion threshold than climate sensitivity, there is again a nearly linear correlation with the threshold radius. To predict a realistic temperature evolution from PI to PD conditions, GCMS must balance RFP and climate sensitivity, often resulting in an inverse correlation between forcing and sensitivity (Schwartz et al. 2007; Kiehl 2007; Knutti 2008). Among our alternate configurations, S1 has the smallest value of \(\Delta T_{s,j}\) at 0.28 K. It is likely that this...

![TOA shortwave absorbed difference (PD-PI)](image)

**FIG. 5.** Zonal mean differences in (a) TOA shortwave absorbed flux and (b) liquid water path between experiments with PD greenhouse gases and emissions, and PI conditions. Differences are relatively small and as a result noisy. Thick lines indicate where the PD – PI differences are statistically significant at the 95% confidence level based on a Student’s \(t\) test.
configuration would not produce a reasonable temperature change in a fully coupled climate model over the PI–PD period.

6. Conclusions

Subgrid activation assumptions—such as how subgrid variability in vertical velocity is treated in the cloud drop nucleation process and how newly nucleated drops are incorporated in clouds—can differ significantly from one GCM to another. These assumptions have a large impact on the predicted cloud drop number. By changing liquid cloud properties, they affect the amount of absorbed shortwave radiation.

Because of their radiative impact, changes in the subgrid activation assumptions are, implicitly or explicitly, accompanied by changes in other aspects of the cloud parameterization to achieve a realistic radiation balance. As an illustration, we built three alternate configurations based on AM3 but with modified subgrid activation assumptions and cloud tuning.

Despite significant underlying formulation changes, impacts on present-day climatology are small between these configurations. It would be difficult to constrain the subgrid assumptions and cloud tuning with current observations. However, there is a very large impact on the total anthropogenic radiative flux perturbation (RFP) between present-day and preindustrial conditions. RFP varies by ±50% (0.42–1.41 W m⁻²) in our alternate configurations compared to the reference AM3 value of 0.91 W m⁻². This range is attributable to a change in the magnitude of the aerosol indirect effect. In the extreme case, the indirect effect appears to counteract the majority of the greenhouse gases effect. The impact on anthropogenic RFP is not caused by the subgrid assumptions themselves but through the required cloud retuning—in particular, the retuning of the autoconversion threshold radius. Adjusting the autoconversion threshold radius for radiation balance is a common practice, but using a different tuning strategy could have resulted in a smaller RFP spread.

Although sensitivities of the indirect effect to details of autoconversion parameterizations have been documented before (e.g., Lohmann and Feichter 1997; Rotstyn 2000; Rotstyn and Liu 2005), we find a linear relationship between the autoconversion threshold radius and the RFP previously not documented. The sensitivity originates from the numerical implementation of the autoconversion, which uses a limiter to avoid over depletion of water. The RFP uncertainty range of nearly 1 W m⁻² documented here is substantially larger than the sensitivity study of Lohmann and Ferrachat (2010) (0.25 W m⁻² for all their experiments and 0.11 W m⁻² for their experiments in energy balance).

To simulate a realistic temperature evolution from preindustrial to present day, a GCM must balance RFP

![Fig. 6](image-url) PD – PI RFP for the four different configurations plotted as a function of the autoconversion threshold. Dashed line is a linear fit. Error bars indicate standard deviation calculated as \( \sigma = \left[ \left( \sigma_{\text{PD}}^2 + \sigma_{\text{RI}}^2 \right) / n \right]^{1/2} \), where \( n \) is the number of years (Lohmann et al. 2010).

![Fig. 7](image-url) As in Fig. 6, but for REF and two additional experiments, REF⁺ and REF⁻, that only differ by the value of the autoconversion threshold. Thin solid line is a linear fit; dashed line is the fit from Fig. 6.

![Fig. 8](image-url) PD – PI liquid water path difference for the four different configurations plotted as a function of the cube of the autoconversion threshold (note the cubic scale of the horizontal axis). Dashed line is a linear fit, indicating a cubic relationship between autoconversion threshold and liquid water path difference as suggested by Eq. (15).
and climate sensitivity. Climate sensitivity estimated from Cess-like experiments does not change substantially in our alternate configurations. This suggests that in the present model setup, the warming from preindustrial to present day would be strongly controlled by the RFP. Some of our alternate configurations would likely result in an unrealistic temperature evolution compared to observations. Our results serve as a reminder that uncertainties in the formulation of individual components of a cloud parameterization can translate into significant uncertainties in the aerosol indirect effect and total anthropogenic RFP. Reducing these uncertainties will necessitate more realistic physical representation of cloud processes and better observational constraints.

Acknowledgments. We acknowledge the Global Atmospheric Model Development Team (GAMDT) for its considerable efforts leading to the creation of AM3. V. Naik is acknowledged for preparing emission input datasets. C. T. Gordon and S.-M. Fan are acknowledged for their suggestions on an early draft. We thank the two anonymous reviewers for their constructive comments on the submitted manuscript.

ERBE and CERES-EBAF data were obtained from the NASA Langley Research Center Atmospheric Science Data Center. GPCP and ISCCP data were obtained from NASA Goddard Space Flight Center. NOAA/NCEP CPC Merged Analysis of Precipitation (CMAP) data were obtained from the IRI/LDEO Climate Data Library. ERA-40 data were supplied by the European Centre for Medium-Range Weather Forecasts (ECMWF).

APPENDIX

Review of Existing Formulations

We review some specific examples of GCM formulations for computing the grid box number of nucleated drops, \( \bar{N}^\xi \) and grid average drop number concentration, \( \bar{N}_d \) in the next subsections.

a. Subgrid distribution of \( w \) with diagnostic \( \bar{N}_d \)

In their assessment of the radiative impact of anthropogenic sulfate, Chuang et al. (1997) were probably the first to implement a PDF-based activation parameterization in a GCM. They assume a normal distribution as underlying \( w \) PDF. Its mean is given by the GCM gridbox vertical velocity, and its standard deviation is set to a constant value of 0.5 m s\(^{-1}\). This value is chosen based on observations from the First International Satellite Cloud Climatology Project (ISCCP) Regional Experiment. They also perform sensitivity tests with standard deviations of 0.25 and 0.75 m s\(^{-1}\). Chuang et al. (1997) diagnosed the cloud drop number with Eqs. (1) and (3) in the form \( \bar{N}_d = \bar{N}^\xi \). They did not incorporate a prognostic cloud drop number equation. Chuang et al. (2002) followed the same approach in a study aimed at estimating cloud susceptibility and the first aerosol indirect effect.

b. Subgrid distribution of \( w \) with prognostic \( \bar{N}_d \)

Ghan et al. (1997) incorporated a prognostic cloud drop equation for large-scale stratiform clouds in the National Center for Atmospheric Research (NCAR) Community Climate Model, version 2 (CCM2). The grid box number of nucleated drops \( \bar{N}^\xi \) is computed based on a normal distribution of subgrid vertical velocity whose mean is the gridbox mean of the GCM modified to account for radiative cooling \([\text{their Eq. (4)}; \text{see also Ghan et al. 1993}]\). The standard deviation is diagnosed from the eddy diffusivity \( K \) and the gridbox height \( D_z \). A lower bound of 0.1 m s\(^{-1}\) is imposed because cloud-top radiative cooling is poorly resolved with the relatively coarse vertical resolution of a GCM:

\[
\sigma_w = \max \left( \sqrt{2\pi K} \frac{\Delta z}{D_z}, 0.1 \right).
\] (A1)

The treatment of the nucleation source term in their prognostic equation of \( \bar{N}_d \) \([\text{Eq. (4)}]\) differs for new and preexisting clouds. For new clouds, the source term adjusts \( \bar{N}_d \) to \( \bar{N}^\xi \) over the course of the time step during which the cloud forms. For preexisting clouds, the source term is only active at cloud base to account for the transport of air at cloud base by advection and turbulence mixing.

The same methodology is also employed in subsequent studies with a different GCM [Model for Integrated Research on Atmospheric Global Exchanges (MIRAGE)] to estimate the radiative forcing from anthropogenic sulfate aerosols (Ghan et al. 2001a) and the aerosol indirect effect (Ghan et al. 2001b). The lower bound on the standard deviation is increased to 0.2 m s\(^{-1}\). In an
updated version of MIRAGE (Easter et al. 2004), cloud drop number is also predicted for subgrid stratiform and convective clouds in addition to large-scale stratiform clouds. Activation in subgrid stratiform clouds is treated similarly to large-scale stratiform clouds. Activation in convective clouds occurs at cloud base using a fixed convective velocity of 1 m s$^{-1}$.

Storelvmo et al. (2006) follow a similar approach in their implementation of a prognostic cloud droplet number equation in the CAM-Oslo. The number of drops that can be activated, $N^*$, is computed by integration of a normal distribution with the mean velocity given by the gridbox average and the standard deviation derived from the eddy diffusivity, but with a minimum value set to 0.3 m s$^{-1}$. In contrast to Ghan et al. (1997), $N^*$ is used to compute the drop source term for both preexisting and new clouds:

$$\frac{\partial N_d}{\partial t}_{\text{activation}} = \max(0, N^* - N_d, 0) \frac{\Delta t}{\text{activation}}.$$ (A2)

GCM implementations of subgrid $w$ variability typically assume that the PDF follows a normal distribution. Studies have shown that normal distributions do not represent shallow cumulus convection well because they are unskewed. Distributions constructed from a mixture of two Gaussians can more faithfully describe the subgrid PDFs in boundary layer clouds (e.g., Larson et al. 2002). A turbulence cloud parameterization based on such a PDF has been developed (Golaz et al. 2002, 2007) and has recently been extended to the prediction of cloud drop number in the single-column version of the GFDL GCM (Guo et al. 2010).

c. Characteristic vertical velocity ($\dot{w}$) and prognostic $N_d$

The first use of the characteristic velocity approach in a GCM emanates from the work of Lohmann et al. (1999a), which describes the implementation of a prognostic cloud drop equation in ECHAM. The number of drops that can be nucleated is calculated from Eq. (2) with the characteristic velocity

$$\dot{w} = \bar{w} + c\sqrt{\text{TKE}},$$ (A3)

where $\bar{w}$ is the gridbox mean GCM vertical velocity, TKE the turbulence kinetic energy, and $c = 0.7$ an empirical factor. The nucleation source term in Eq. (4) is computed as follows. Nucleation occurs at cloud base when condensation occurs. The nucleation source term is assumed to be constant throughout the depth of the cloud. For preexisting clouds, additional drops are nucleated when the number that would be activated at cloud base is larger than the existing drop number. A similar approach is also retained by Takemura et al. (2005) and Goto et al. (2008).

In a subsequent study of the impact of aerosols on ice clouds, Lohmann (2002) employs a characteristic vertical velocity to estimate the maximum number of nucleated cloud drops in both stratiform and convective clouds with an increased scaling factor for the TKE:

$$\dot{w} = \left\{ \begin{array}{ll} \bar{w} + 1.33\sqrt{\text{TKE}}: & \text{stratiform clouds} \\ \bar{w} + 0.5\sqrt{\text{CAPE}}: & \text{convective clouds}. \end{array} \right.$$ (A4)

A characteristic convective velocity based on the convective available potential energy (CAPE) is likely to be substantially larger than the fixed value of 1 m s$^{-1}$ suggested by Easter et al. (2004). The nucleation source term is related to $N^*$:

$$\frac{\partial N_d}{\partial t}_{\text{activation}} = \max(0, 0.1N^*^{1.27} - N_d, 0) \frac{\Delta t}{\text{activation}}.$$ (A5)

where the empirical fit of Leaitch et al. (1996) has been incorporated. Treatment of nucleation in new and preexisting clouds remains unchanged from Lohmann et al. (1999a).

Ming et al. (2007) implemented a prognostic cloud drop number equation in a modified version of AM2. For convective clouds, the characteristic vertical velocity is chosen to be $0.5\sqrt{\text{CWF}}$, where CWF is the cloud work function, as in Lohmann (2002). For stratiform clouds, the characteristic velocity is chosen consistently with the velocity driving large-scale condensation in the Tiedtke (1993) cloud parameterization. It consists of the resolved velocity outside the convective area modified to include the impact of the turbulent and radiative cooling rates. Nucleation only occurs in newly formed clouds—that is, when the cloud fraction is predicted to increase by Tiedtke (1993).

Morrison and Gettelman (2008) and Gettelman et al. (2008) have incorporated a two-moment bulk cloud microphysics parameterization in the CAM3 using a characteristic velocity. They parameterize $\dot{w}$ as

$$\dot{w} = \max\left(\frac{K}{l_c}, 0.1\right).$$ (A6)

where $K$ is the eddy diffusivity and $l_c = 30$ m, a constant mixing length. They also performed a sensitivity experiment with a lower bound of 0.4 m s$^{-1}$. Note that the mixing length $l_c$ is smaller than the equivalent mixing length implied by Eq. (16) wherever the vertical grid spacing is coarser than 190 m. In a typical GCM, this
would include most levels except a few near the surface. Calculation of the source term in the cloud drop budget equation follows Eq. (7), with the exception that $\Delta t$ is replaced with a relaxation time scale of 20 min, which corresponds with the eddy turnover time and time scale for recycling air through cloud base.

Wang and Penner (2009) implemented an alternative formulation of the characteristic velocity:

$$\dot{w} = \overline{w} + c(\sigma_{w})\sigma_{w}, \quad \text{with} \quad (A7)$$

$$\sigma_{w} = \max\left(\frac{K}{L_{c}}, 0.1\right) \quad (A8)$$

and the mixing length $L_{c}$ varying between 300 m in the boundary layer and 30 m in the free atmosphere. The functional form $c(\sigma_{w})$ is empirically fitted such that the cloud drop number predicted with $\dot{w}$ in Eq. (A7) approximately matches the one obtained by integration over a normal distribution of zero mean and $\sigma_{w}$ standard deviation. In their fit, $c(\sigma_{w})$ varies between 0.2 and 0.6. They reported an accuracy of 10% as long as the accumulation mode aerosol concentration is less than 1000 cm$^{-3}$.

Recently, Morales and Nenes (2010) proposed a characteristic vertical velocity framework for activation that provides a droplet number concentration characteristic of the value averaged over the PDF of vertical velocity. They derived an analytical formulation for the characteristic velocity for microphysical processes that can be approximated with a power law of the velocity and a normal distribution of the subgrid velocity. They also evaluated the impact of droplet variability effects on autoconversion and effective radius.

Finally, we also note that to predict realistic cloud drop numbers, some mesoscale modeling studies also include subgrid velocity variability using a characteristic velocity derived from the TKE (e.g., Morrison and Pinto 2005; Morrison et al. 2008; Ivanova and Leighton 2008).

d. Validation studies

A few studies have attempted to validate some of the formulations discussed above against observational data and high-resolution models.

Lin and Leaitch (1997) were probably the first to investigate the application of local activation parameterizations (Ghan et al. 1993; Abdul-Razzak et al. 1998) against observations collected during 14 flights from the 1993 North Atlantic Regional Experiment (NARE). They found that an integration over a normal $w$ distribution with zero mean and observed variance led to an underestimate of the cloud drop number compared to observations. They obtained better agreement by first estimating the maximum cloud drop number, $\overline{N}_{\ast}$, using a characteristic velocity $\dot{w} = \overline{w} + 2\sigma_{w}$, and then relating this number to the average drop number with an empirical fit, $\overline{N}_{d} = 0.1\overline{N}_{\ast}^{1.27}$, from Leaitch et al. (1996).

Peng et al. (2005) compared observed cloud drop number for 11 marine stratus cloud cases with simulations from a parcel model. The observational data were collected during two field experiments in the North Atlantic Ocean [NARE and the Radiation, Aerosol and Cloud Experiment (RACE) in 1995]. Simulated cloud drop numbers were calculated using a characteristic velocity as well as a subgrid distribution. They estimated the maximum cloud drop number, $\overline{N}_{\ast}$, and derived an empirical fit to relate it to the average cloud drop number, $\overline{N}_{d} = 0.062\overline{N}_{\ast}^{1.35}$. For the simulations with characteristic velocity, they found that $\dot{w} = 0.8\sigma_{w}$ yielded the best fit between model and data. For the subgrid-scale approach, they assumed a normal distribution with zero mean and observed variance. In a departure from other works, Peng et al. (2005) did not integrate over the entire PDF, instead they integrated from a variable lower bound $w_{0}$, “defined as the minimum value above which the application of the PDF of the $w$ values leads to simulated $N$ that are in agreement with the observed $N’.” Comparing both approaches for the 11 cases they analyzed, they concluded that “one single characteristic vertical velocity... is a good surrogate for the vertical velocity for the nucleation process,” but they also noted that “the exact value of this characteristic vertical velocity may depend on the aerosol properties and dynamic factors for clouds in different regions.”

Jiang and Cotton (2005) estimated the characteristic vertical velocity using LESs of six boundary layer cloud cases that included stratocumulus and shallow cumulus clouds. Assuming

$$\dot{w} = \overline{w} + c\sigma_{w}, \quad (A9)$$

they derived empirical values for the coefficient $c$ for each of their LESs. They obtained values ranging from 0.1 to 0.54 and suggested the median value of 0.24 as an appropriate choice. Because the lowest reported $c$ value was derived for a shallow cumulus regime, and the largest value a stratocumulus regime, it is likely that $c$ is regime dependent.

REFERENCES


Barker, H. W., and P. Räisänen, 2004: Neglect by GCMs of subgrid-scale horizontal variations in cloud-droplet effective


Dufresne, J.-L., J. Quaas, O. Boucher, S. Denvil, and L. Fairhead, 2005: Contrasts in the effects on climate of anthropogenic sulfate aerosols between the 20th and the 21st century. Geo-


and thermodynamics of numerical models that ignore subgrid-scale variability. J. Atmos. Sci., 58, 1117–1128.


