Impact of Aerosols on Convective Clouds and Precipitation

Wei-Kuo Tao\textsuperscript{1}, Jen-Ping Chen\textsuperscript{2}, Zhanqing Li\textsuperscript{3}, Chien Wang\textsuperscript{4}, Chidong Zhang\textsuperscript{5}

\textsuperscript{1}Laboratory for Atmospheres
NASA/Goddard Space Flight Center
Greenbelt, MD 20771

\textsuperscript{2}Department of Atmospheric Sciences
National Taiwan University
Taipei, Taiwan 10617

\textsuperscript{3}ESSIC and Department of Atmospheric & Oceanic Sciences
University of Maryland
College Park, MD 20740

\textsuperscript{4}Massachusetts Institute of Technology
Cambridge, MA 02139

\textsuperscript{5}Rosenstiel School of Marine and Atmospheric Science (RSMAS)
University of Miami
Miami, FL 33149

Reviews of Geophysics

July 8 2011

Corresponding author address: Dr. Wei-Kuo Tao, Mesoscale Atmospheric Processes Branch, Code 613.1, NASA GSFC, Greenbelt, MD 20771
E-mail: wei-kuo.tao-1@nasa.gov
ABSTRACT

Aerosols are a critical factor in the atmospheric hydrological cycle and radiation budget. As a major reason for clouds to form and a significant attenuator of solar radiation, aerosols affect climate in several ways. Current research suggests that aerosol effects on clouds could further extend to precipitation, both through the formation of cloud particles and by exerting persistent radiative forcing on the climate system that disturbs dynamics. However, the various mechanisms behind these effects, in particular the ones connected to precipitation, are not yet well understood. The atmospheric and climate communities have long been working to gain a better grasp of these critical effects and hence to reduce the significant uncertainties in climate prediction resulting from such a lack of adequate knowledge.

The central theme of this paper is to review past efforts and summarize our current understanding of the effect of aerosols on precipitation processes from theoretical analysis of microphysics, observational evidence, and a range of numerical model simulations. In addition, the discrepancy between results simulated by models, as well as that between simulations and observations will be presented. Specifically, this paper will address the following topics: (1) fundamental theories of aerosol effects on microphysics and precipitation processes, (2) observational evidence of the effect of aerosols on precipitation processes, (3) signatures of the aerosol impact on precipitation from large-scale analyses, (4) results from cloud-resolving model simulations, and (5) results from large-scale numerical model simulations. Finally, several future research directions on aerosol – precipitation interactions are suggested.
1. **Introduction**

Aerosols, and especially their effect on clouds and precipitation, are one of the key components of the climate system and the hydrological cycle. Yet the aerosol effect on clouds and precipitation remains poorly known. A recent report published by the National Academy of Science of the United States states that, "The greatest uncertainty about the aerosol climate forcing - indeed, the largest of all the uncertainties about global climate forcing - is probably the indirect effect of aerosols on clouds NRC [2005]." This "aerosol indirect effect" (AIE) includes the traditional "indirect", or "Twomey" effect, on cloud droplet size and thus reflectance for a constant liquid water path [Twomey, 1977; Twomey et al., 1984] and the "semi-indirect" effect on cloud extent and lifetime [Albrecht, 1989; Hansen et al., 1997; Ackerman et al., 2000].

Enhanced aerosol concentrations can also suppress warm rain processes by reducing particle sizes and causing a narrow droplet spectrum that inhibits collision and coalescence processes [e.g., Squires and Twomey, 1961; Warner and Twomey, 1967; Warner, 1968; Rosenfeld, 1999].

The aerosol effect on precipitation processes, also known as the second type of aerosol indirect effect [Albrecht, 1989], is even more complex, especially for mixed-phase convective clouds.

Continued advancement of instrumental technology provides the community with much improved research tools to reveal insights concerning aerosol-cloud-precipitation interactions. For instance, a combination of cloud-top temperature and effective droplet sizes, estimated from the Advanced Very High Resolution Radiometer (AVHRR), has been used to infer the suppression of coalescence and precipitation processes for smoke [Rosenfeld and Lensky, 1998] and desert dust [Rosenfeld et al., 2001]. Multi-sensor (passive/active microwave and visible and infrared) satellite observations from the Tropical Rainfall Measuring Mission (TRMM) have also been used to identify the presence of non-precipitating super-cooled liquid water near the cloud top due to over-seeding from both smoke over Indonesia [Rosenfeld, 1999] and urban pollution over Australia [Rosenfeld, 2000]. In addition, aircraft measurements have provided evidence of sustained super-cooled liquid water down to -37.5°C in continental mixed-phase convective clouds [Rosenfeld and Woodley, 2000]. These findings further suggest that the increase of continental aerosols can reduce the mean size of cloud droplets, suppressing coalescence and warm-rain processes, permitting more freezing of cloud droplets and associated latent heat release above the 0°C isotherm, and enhancing the growth of large hail and cold-rain processes.
[Rosenfeld and Woodley, 2000]. Andreea et al. [2004] analyzed in situ observations made during the LBA-SMOCC (the Large-Scale Biosphere-Atmosphere Experiment in Amazonia-Smoke, Aerosols, Clouds, Rainfall, and Climate) campaign and found that increases in smoke and surface heat due to biomass burning tend to lead to higher cloud-top heights and the enhancement of cold-rain processes over the Amazon basin. Lin et al. [2006] examined multi-platform satellite data over the Amazon basin and found that high biomass burning-derived aerosols are correlated with elevated cloud-top heights, large anvils, and more rainfall. Koren et al. [2005] examined cloud properties derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) and found strong evidence that aerosols from pollution, desert dust and biomass burning systematically invigorate convective clouds over the Atlantic Ocean. Using long-term integrated TRMM-derived precipitation data, Bell et al. [2008] found a significant mid-week increase in summer-time afternoon thunderstorms over the southeast U.S., which coincides with a mid-week increase in ground-measured aerosol concentration.

These findings are consistent with the notion that aerosols have a major impact on the dynamics, microphysics, and electrification properties of continental mixed-phase convective clouds. In addition, high aerosol concentrations in urban environments could affect precipitation variability by providing a significant source of cloud condensation nuclei (CCN). Such pollution effects on precipitation potentially have enormous climatic consequences both in terms of feedbacks involving the land surface via rainfall as well as the surface energy budget and changes in latent heat input to the atmosphere. Table 1 summarizes observations of cloud system features in high-aerosol and low-aerosol continental environments. Basically, aerosol concentrations can influence cloud droplet size distributions, the warm-rain process, the cold-rain process, cloud-top heights, the depth of the mixed-phase region, and the occurrence of lightning. The subsequent review of observation-modeling studies of the AIE follows the general lifecycle of a cloud from the germination of cloud droplets, warm-rain, cold rain and mixed-phase rain processes, and deep convective clouds to thunderstorms. In each phase, aerosols exert different influences on the development of clouds that ultimately affect precipitation.

The main objective of this paper is to review our current understanding of the interactive processes between aerosols, clouds and deep-convective precipitation systems. Recent results from observations and modeling (from cloud-resolving models to global models) will be reviewed. The paper has seven sections. It opens in the following section with a brief description
of fundamental theories regarding the effects of cloud condensation nuclei and ice nuclei on microphysics and precipitation processes. The observational evidence of the aerosol effects on precipitation processes are then reviewed in section 3. Section 4 presents the signatures of aerosol impacts on precipitation processes from large-scale analyses. Sections 5 and 6 summarize recent results from cloud-resolving (-processes) model and large-scale numerical model simulations, respectively. In section 7, current and future research on aerosols and their interactions with precipitation are discussed.

2. Fundamental Theories of Condensation Nuclei Effects on Cloud and Precipitation Processes

2.1. Activation of Condensation Nuclei into Cloud Drops

Aerosol particles composed of hygroscopic materials are good sites for water vapor to condense, so they are called condensation nuclei (CN). CCN are the portions of CN that are capable of initiating cloud drop formation. Such a capability is strongly related to the mass and composition of its water-soluble component. Köhler [1936] explained this mechanism with the Köhler curve which describes the equilibrium saturation ratio over the solution drop surface, $S_d$, as a function of the drop radius. For each droplet containing a fixed amount of dissolved salt, the solution (Raoult) effect reduces the water activity and thus surface vapor pressure, whereas the curvature (Kelvin) effect does the opposite. These two effects combine to form the Köhler curve which shows a peak value in $S_d$ at a unique drop size called the critical radius $r^*$ (see Figure 1). The maximum $S_d$ appearing on a curve corresponding to a given particle with dry size $r_d$ or equilibrium size $r_e$ is commonly called the critical (or activation) saturation ratio, $S^*$, for that particle, presumably with a known composition and fixed solute mass. During cloud formation, once the ambient saturation ratio $S_\infty$ exceeds the $S^*$ of a particle, that particle can be activated into a cloud drop. Those particles that never meet the condition of $S_\infty > S^*$ will stay in a “haze state” and are called interstitial aerosols. This also means that whether CN can be regarded as CCN depends not only on the properties of the particle but also the ambient condition (i.e., $S_\infty$). It is important to recognize that aerosols have a broad size range, and the sizes of CN are important in determining their role in cloud and precipitation formation. From Figure 1, one can see that larger aerosols have lower $S^*$, so they have the advantage of easier activation. Note that
the solute amount in an aerosol particle may vary with time during its uptake of water either by
absorbing gases (such as nitric acid or ammonia) from the ambient air or by dissolving its solid
cores (e.g., mineral dust) [Kulmala et al., 1993; Kulmala et al., 1997]. Besides solute mass (or
particle dry size), the Köhler curve also varies with the solute chemical composition, which
influences not only the Raoult effect (by altering the water activity) but also the Kelvin effect (by
altering the surface tension). Temperature is another factor that determines Köhler curve
characteristics. An aerosol particle may become warmer than its environment due to either latent
heating from vapor condensation or absorption of solar radiation by its soot or mineral dust
component. Such heating affects all three factors that determine the Köhler curve (saturation
vapor pressure, surface tension and solution activity) so the Köhler curve should not be
calculated using the ambient temperature [Conant et al., 2002; Nenes et al., 2002]. The insoluble
portion of the aerosol particle, such as soot and mineral dust, may also affect the Köhler curve,
simply by adding to the particle size and thus reducing the Kelvin effect.

The relationship between cloud drop number concentration (CDNC), \( N_C \), and aerosol
number spectra is an important issue in studying aerosol effects on cloud microphysics and
radiation with regional and global atmospheric models. The Köhler theory is a first step in
understanding this relationship. However, the Köhler theory can only tell us whether a
condensation nucleus can be activated into a cloud drop under a particular ambient
supersaturation, \( ss_\infty \) \( (\equiv S_\infty - 1) \). The maximum \( ss_\infty \) (hereafter termed \( ss_{\infty,max} \)) that can be reached
must be known in order to determine the number of cloud drops activated at the cloud base. The
ambient supersaturation is controlled by two factors, as indicated by the saturation development
equation [Squires, 1952]:

\[
\frac{dss_\infty}{dt} = A_1 \frac{dz}{dt} - A_2 \frac{dq_w}{dt}
\]

where the coefficients \( A_1 \) and \( A_2 \) are properties of the ambient air, \( z \) is the height and \( q_w \) is the
cloud water content. In the above equation, the source term indicates an increase in \( ss_\infty \) due to
expansion cooling (thus a decrease in saturation vapor pressure) in the updrafts, and the sink
term is due to water vapor removal by condensation onto cloud drops. The condensation rate,
\( dq_w/dt \), is roughly proportional to \( ss_\infty \), cloud drop size, and \( N_C \); \( N_C \), in turn, is determined by the
CN size spectrum and the variation in \( ss_\infty \). As more and more CCN turn into cloud drops, the
sink term quickly surpasses the source term after which \( ss_\infty \) starts to decrease. Typically, this
peak supersaturation, \( ss_{\infty,\text{max}} \), is reached within a few tens of meters above the cloud base. No more CN activation may occur deeper into the cloud unless \( ss_{\infty} \) rises again and surpasses the original \( ss_{\infty,\text{max}} \) due to enhanced updrafts or depletion of cloud drops by rain accretion [Chen, 1994a]. Although the relationship between \( N_C \) and the CN spectrum is quite convoluted, a unique solution for \( N_C \) at the cloud base exists when the CN spectrum and updraft speed \( (dz/dt) \) are given. In the following, we briefly describe the common form of the CN spectrum representation and how it can be used to derive \( N_C \).

One common type of CN spectra measurement is based on the Köhler theory discussed above. The aerosol particles to be measured are introduced into a diffusion chamber with controlled \( ss_{\infty} \), and then the number of cloud drops formed at that condition is counted. This specified \( ss_{\infty} \) represents the critical supersaturation, \( ss^* (= S^*-1) \), of the smallest CCN that can be activated. By sweeping through a range of \( ss_{\infty} \) values in the diffusion chamber, one can obtain a relationship between the number of CCN \( (N_{CCN}) \) and \( ss_{\infty,\text{max}} \). Earlier studies found that the measured CCN spectra often can be approximated empirically as:

\[
N_{CCN} = C \cdot ss_{\infty,\text{max}}^k
\]

where \( C \) and \( k \) are time- and location-dependent coefficients of the CN spectra [Twomey, 1959a]. Figure 2 shows examples of the measured relationship between \( N_{CCN} \) (normalized by the number of CN, \( N_{CN} \)) and \( ss_{\infty,\text{max}} \), as well as that using Eq. (2) for several \( k \) values. Note that \( ss_{\infty,\text{max}} \) here is often expressed in terms of percentage of supersaturation. Such a power-law formulation is simple and effective, but it has the weakness of producing unlimited aerosol number concentrations with ever-increasing \( ss_{\infty} \) while in reality, the number of aerosols normally diminishes when the size (supersaturation) gets too small (high).

Another common type of aerosol measurement is done by segregating aerosol particles according to their aerodynamic sizes then counting the particle number in each size segment to obtain a size spectrum. Earlier studies using such techniques found that the number density distribution (or the size distribution) of aerosol particles with a dry size, \( r_a \), greater than 0.1 \( \mu \text{m} \) can be approximated by a power-law relationship [e.g., Junge, 1952, 1955]:

\[
n(r_a) \equiv \frac{dN(r_a)}{dr_a} = a_1 r_a^{-a_2}
\]
where \( N(r_a) \) is the total number of CN with a dry size greater than \( r_a \), and \( a_1 \) and \( a_2 \) are constants. Note that \( ss_{\infty, max} \) in Eq. (2) actually represents the critical supersaturation of the smallest CCN, and it can be related to the \( r_a \) of that particle by the approximation \( s^* \approx \alpha \cdot r_a^{-3/2} \), where \( \alpha \) depends on the chemical composition. With this relationship, Eq. (2) can be converted into a size spectrum:

\[
\frac{dN_{CCN}}{dr_a} = \alpha \cdot C \cdot r_a^{-(1/3)}.
\]

One can easily recognize the equivalency between Eqs. (2) and (4). As mentioned before, these formulas might not apply to \( r_a < 0.1 \) \( \mu m \), yet under some conditions, such as in strong convection, very small CN do form cloud drops. To better describe the full size spectrum, many studies applied the tri-modal log-normal function to represent the observed aerosol spectra [e.g., Whitby, 1978; Jaenicke, 1993; Brechtel et al., 1998]:

\[
n(\ln r_a) = \sum_{i=1,3} \frac{N_i}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(\ln(r_a/r_0^i))^2}{2\sigma_i^2}\right]
\]

where \( i \) is the modal index, and for each mode, \( N, \sigma \) and \( r_0 \) are the number concentration, standard deviation and modal radius, respectively. Examples of the tri-modal log-normal size distribution are illustrated in Figure 3. Presenting the aerosol population in terms of size spectrum such as Eqs (3) or (5) has a disadvantage that chemical composition must be known in order to calculate \( NC \), but such information is often not provided. In this regard, Eq. (2) may be a better formulation for CN spectra because it does not require additional information. Even if chemical compositions are known, the numerous chemical species in natural aerosols makes it difficult for model calculations. For this problem, Petters and Kreidenweis [2007] suggested a single hygroscopicity parameter, \( \kappa \), which can be measured relatively easily and used effectively to model CCN activity.

Quite a few cloud resolving models (CRMs) are capable of resolving \( ss_{\infty} \) and thus obtain \( NC \) directly by applying the Köhler theory to a specified CN spectrum such as those given in Eqs. (2) and (5). These models include those using sophisticated bin microphysics [e.g., Clark, 1973; Kogan, 1991; Khain et al., 2004; Tao et al. 2007; Morrison and Grabowski, 2010], as well as those using bulk microphysics [e.g., Feingold et al., 1994; Ekman et al., 2006, 2007; Cheng et al., 2007; Morrison and Grabowski, 2008]. But many global models and regional cloud models,
such as those adopting the “saturation adjustment” assumption, do not resolve $ss_\infty$ explicitly. Even for CRMs that do resolve $ss_\infty$, the peak supersaturation very often is under-predicted because it occurs within roughly 50 m above the cloud base, which is at a level that CRMs cannot resolve [cf. Clark, 1973; Chen and Liu, 2004]. Nevertheless, Cheng et al. [2007] showed that the calculation of supersaturation in such models can be improved by imbedding a Lagrangian parcel scheme. The Lagrangian-type calculation can also be applied in global models by running a parcel model, either online or offline, with given updraft speeds in order to derive $N_C$ [Nenes and Seinfeld, 2003; Reolofs et al., 2006]. But such calculations not only demand more computational resources but also require detailed calculations of thermodynamics and the aerosol activation process according to the Köhler theory. Alternatively, $N_C$ may be derived by using diagnostic formulas introduced below.

Twomey [1959b] was the first to link $N_C$ to the CN spectrum in a diagnostic way. By inserting Eq. (2) into Eq. (1) and applying a simplified condensation equation, he derived an asymptotic solution for $N_C$ as a function of $C$, $k$, updraft speed and air properties. As shown in Figures 4(a) and 4(b), Twomey’s solution predicts that $N_C$ increases monotonically with $N_{CN}$. Recall that Eqs. (2) and (4) will yield an infinite number of $N_{CN}$ if no upper limit is set for $ss_{\infty,\text{max}}$ in Eq. (2) or no lower limit is set for $r_a$ in Eq. (4). Yet, if the 0.1 $\mu$m cutoff size mentioned in Junge [1952, 1955] is applied to Eq. (2), the predicted $N_C$ would unrealistically exceed $N_{CN}$. So, a cutoff radius of 0.01 $\mu$m is applied to obtain the $N_{CN}$ in Figures 4(a) and 4(b), and the chemical composition is assumed to be ammonium sulfate. Twomey’s solution compares well with the observational results summarized by Ramanathan et al. [2001] (Figure 4(c)). The trends and the magnitudes are quite similar, and the regional differences in the observations seem to correspond to the variations in coefficient $k$ in Eq. (2). One feature that Twomey’s prediction does not capture well is the leveling off of $N_C$ toward higher $N_{CN}$. Chen and Liu [2004] examined Twomey’s quasi-analytical solution by incorporating Eq. (2) into a parcel model with detailed microphysics to obtain $N_C$. They found that Twomey’s solutions significantly deviate from model calculations for large $k$ and $C$ values, partly because Twomey’s solution is not exact and partly because its derivation omitted the solute effect and gas kinetic effect on condensation growth. Modification of Twomey’s method has been performed by Sedunov [1967], Ming et al. [2006], Khvorostyanov and Curry [2009] and others, and the improvements include a refined asymptotic solution as well as considering more physical constraints such as soluble fraction,
surface tension, and condensation (accommodation) coefficient. Parameterizations of CCN activation were also developed using results from a parcel model that are fitted into various empirical functions [e.g., Abdul-Razzak et al., 1998; Fountoukis and Nenes, 2005] or by compiling results into lookup tables [Saleeby and Cotton, 2004]. Besides the theoretical approaches, prediction of $N_C$ can also be obtained empirically by correlating $N_C$ observed just above the cloud base with aerosol concentration (usually the mass of sulfate) observed below the cloud base. This method has been applied commonly in global climate models [e.g., Boucher and Lohmann, 1995; Roelofs et al., 1998; Jones et al., 2001; Tsai et al., 2010].

Past studies on CCN focused mainly on the inorganic component of the aerosol. Yet around the globe, organic carbon is the second-most abundant component of fine aerosols after sulfates [Novakov and Penner, 1993; Kanakidou et al., 2005; Liu and Wang, 2010]. Soluble organic compounds can decrease water activity and the surface tension of water [Facchini et al., 1999; Abdul-Razzak and Ghan, 2004] just like inorganic solutes, but the effect is weaker for organics on the same mass basis. Such organic aerosol effects have been included in a few of the aforementioned studies [e.g., Nenes and Seinfeld, 2003]. One particular type of organic carbon called “surfactants” consists of polar (hydrophilic) and non-polar (hydrophobic) segments, so these molecules stay preferentially at the droplet surface and form a monolayer film [Langmuir and Langmuir, 1927; La Mer, 1962]. Such a monolayer film does not influence the Köhler curve but can retard the condensation process through the gas-kinetic effect by reducing the water accommodation coefficient. Derjaguin et al. [1985] applied a 1-dimensional model to show significant retardation of cloud drop condensation growth by the monolayer film. Feingold and Chuang [2002] further demonstrated that, because of slower growth of the activated cloud drops (i.e., the sink term in Eq. (1)), the monolayer film caused a higher ambient supersaturation $ss_\infty$ and thus promoted activation of the smaller CN. Such a surface film effect is generally ignored in current cloud models due to its complexity as well as a lack of information on the characteristics of surfactants in aerosols.

2.2. CCN Effect on Rain Formation

Rain formation processes can be categorized into two major pathways: warm rain processes that do not involve ice crystals, and cold-rain processes that do. CN have a strong influence on warm
rain formation, mainly through their effect on $N_C$. As discussed above, $N_C$ tends to increase with $N_{CN}$. Because of competition of water vapor among cloud drops, a higher $N_C$ necessarily leads to smaller cloud drops. Small cloud drops are ineffective in collision, partly due to lower fall speeds and partly due to smaller collision efficiencies. It is difficult for a cloud drop to grow into a raindrop merely by condensation, and collision-coalescence is considered a necessary process for warm-rain initiation. Once a raindrop is initiated, its further growth proceeds mainly by collecting cloud drops, and the efficiency of this process again depends heavily on cloud drop sizes.

Another effect of increasing $N_{CN}$ on warm rain formation is the narrowing of the cloud drop size spectrum. If we assume that the shape of a CN spectrum, as defined by the coefficient $k$ in Eq. (2) or $\sigma$ and $r_0$ in Eq. (5), remains fixed, then the increase in CN number concentration will lead to a lower $s_{ss,\infty,max}$ and thus a larger cutoff size for CN activation. This means that the CN being activated, and therefore the cloud drops formed, have a narrower size range. The difference in cloud drop sizes (and thus fall speeds) is important to the collision process as it requires that two colliding droplets have significant relative motions. Therefore a narrowed spectral width also leads to suppression of warm rain formation.

Note that the warm rain suppression effect might not occur if the increase in $N_{CN}$ comes with a change in the shape of the CN spectrum. An extreme example is the addition of a few very large CNN, which are often termed giant CCN (GCCN), with radii in the micrometer range or larger. One natural GCCN is large sea salt particles generated by the breakup of air bubbles in the ocean or by wave tearing. Because of their inherently large sizes, GCCN may be activated directly into rain embryos and readily initiate the warm rain process [Johnson, 1982; O’Dowd et al., 1997; Feingold et al., 1999; Lasher-Trapp et al., 2001; Cheng et al., 2007]. In addition, as they also compete for water vapor, GCCN can lower $s_{ss,\infty,max}$ and thus deprive the chance of small CN to activate into cloud drops [Ghan et al., 1998; Feingold et al., 1999; Cheng et al., 2007]. Furthermore, as there are less cloud drops, each cloud drop can grow larger and thus is more efficient in converting into rain by collision coalescence. So, GCCN counteract both Twomey’s first and second indirect effects. Large insoluble particles such as mineral dust have a similar capability especially when they are coated with hygroscopic materials. The GCCN effect has been utilized for artificial rain enhancement. In the so-called warm-cloud seeding procedure, GCCN-sized salt particles are introduced into clouds to accelerate rain initiation [Mather et al.,
However, the effectiveness of GCCN depends on how efficient rain can be initiated without them. In clean environments (e.g., low $N_C$), cloud drops are able to grow large enough by condensation to reach the so-called Hocking limit, after which the cloud drop can grow effectively by collision coalescence and convert into raindrops. So, GCCN tend to be less important in such situations. As found out by Feingold et al. [1999] with detailed model simulations, higher concentrations of GCCN are required at higher CCN concentrations for marine stratocumulus to produce drizzle.

The discussions above suggest that more CN will lead to more cloud drops and less precipitation. Yet there are many examples showing that an increase in CCN does not lead to higher cloud water content or longer cloud lifetimes. For instance, Ackerman et al. [2004] demonstrated that the entrainment of overlying dry air above a deck of boundary layer stratocumulus cloud may result in a decrease in cloud water when the humidity of the overlying air is very low. Also, more and smaller cloud drops may enhance evaporation in the downdraft region of cloud edges, which leads to lower cloud fraction, cloud size, and cloud depth [Teller and Levin, 2006; Xue and Feingold, 2006]. More detailed discussion on these effects will be given in Section 5.

In cold clouds, such as the upper part of deep convective or high latitude clouds, CN’s effects on cloud and precipitation are much more complicated than in warm clouds. Due to the Raoult effect of freezing point depression, soluble compounds tend to inhibit the freezing of liquid aerosols and cloud droplets and thus the formation of ice crystals. Yet, liquid droplets normally cannot form without these soluble chemicals. So, when the atmosphere is short of ice nuclei (IN), the only viable way to form cloud ice is to freeze the water in aerosol particles or cloud drops. Such homogeneous nucleation occurs quite commonly in cirrus formed by slow lifting or in orographic wave flow at high altitudes [Heymsfield and Sabbin, 1989; Heymsfield and Miloshevich, 1993], as well as in the cirrus anvil generated in convective storms [Knollenberg et al., 1993; Heymsfeld et al., 2005]. Since $N_C$ is largely controlled by $N_{CN}$, the number of ice crystals in such clouds inevitably is influenced by the amount of CN entering the clouds. Several model simulations showed that aerosols have strong influences on not only the microphysical structure but also the extent and lifetime of cirrus [Chen et al., 1997; Phillips et al., 2005; van den Heever et al., 2006; Carrió et al., 2007; Ekman et al., 2007; Fan et al., 2010a]. In regions where cloud drop freezing is slow (i.e., not all cloud drops can freeze), the
number of cloud ice particles formed by cloud drop freezing depends not on $N_C$ but on the cloud-drop mass mixing ratio according to the classical nucleation theory. In this case, the number of cloud ice particles may increase with increasing $N_{CN}$ due to the suppression of warm-rain processes thus elevating cloud water content. But each cloud ice particle that formed would be smaller because of smaller cloud drop sizes. It is interesting to note that cloud ice particles with smaller initial sizes may not necessarily grow slower than larger ones due to the shape effect [Sheridan et al., 2009]. Another way to initiate cloud ice is by heterogeneous nucleation with the help of a particular type of aerosol called ice nuclei. Water may directly deposit onto ice nuclei as ice (called deposition nucleation) or as liquid then freeze (called condensation-freezing nucleation). Such a process is very sensitive to $s_{\infty}$. As can be realized from Eq. (1), $s_{\infty}$ is strongly modulated by $N_C$ and the size of cloud drops, so CN can influence ice nucleation indirectly through their effect on cloud drops.

Besides nucleation, CN may also influence the growth of ice particles. When cloud ice crystals are first formed, they grow mainly by vapor deposition. The water vapor they need comes either from a continuous cooling of air by lifting or by the evaporation of cloud drops. The latter is called the Wagner-Bergeron-Findeisen (WBF) conversion, and this process is limited by the rate of cloud drop evaporation. With more cloud drops, the WBF conversion proceeds faster, so the deposition growth of cloud ice is influenced indirectly by CN. When cloud ice crystals are large enough, they can also grow by accreting cloud water, a process called riming. Just like the collision between cloud drops, the efficiency of riming is strongly dependent on the cloud drop size. Again, since cloud drop sizes are influenced by $N_{CN}$, so will the riming process. In general, the WBF conversion is enhanced but the riming process is suppressed with more CN. Therefore, the formation of snow, graupel and hail, as well as the so-called cold rain that forms from their melting, will all be affected [Khain et al., 2004; van den Heever et al., 2006; Tao et al., 2007; Cheng et al., 2010]. The net outcome will vary with the overall dynamic and thermodynamic structures of the clouds, as well as the availability of natural GCCN and IN. Note that all the CN effects on ice-phase precipitation and cold rain formation mentioned above are related to cloud particle size and number concentration. These effects can be best resolved with binned microphysical schemes [e.g., Khain et al., 2004; Tao et al., 2007; Morrison and Grabowski, 2010], and to some extent, by multi-moment bulk microphysical schemes [e.g., Feingold et al., 1994; Morrison and Pinto, 2005; Cheng et al., 2007; Li et al.,
Single-moment bulk microphysical schemes are not suitable because they do not keep track of particle sizes.

### 2.3 Ice Nuclei Effect on Cold-Cloud Processes

Cloud glaciation can be initiated by either homogeneous freezing that requires very low temperatures, or by heterogeneous nucleation which works at higher temperatures with the aide of IN. The latter is considered very important to cloud physical processes, precipitation formation and global radiation balances [Pruppacher and Klett, 1997; Cantrell and Heymsfield, 2005]. Just like CCN, the effect of IN on cloud microphysics comes mainly from their impact on cloud ice number concentration, $N_{CI}$, which is particularly important in the mixed-phase zone of clouds where the balance between ice and liquid is sensitive to cloud ice number [Korolev and Field, 2008]. Because of insufficient knowledge about ice nucleation, mixed-phase zone processes are difficult to model [Klein et al., 2009; de Boer et al., 2010]. Nevertheless, relevant information about IN and ice nucleation rates is accumulating.

The most common type of IN is probably mineral dust particles transported into the atmosphere by strong winds from bare lands, such as deserts, semi-arid areas, agricultural uplands and dry river beds. They have been implicated as possible atmospheric IN for over 50 years [Isono et al., 1959]. More recently, emerging evidence has suggested that mineral particles may reach high into the upper troposphere where they serve as IN for cirrus cloud formation [Heintzenberg et al., 1996; Sassen, 2002; Sassen et al., 2003]. DeMott et al. [2003] detected IN concentrations up to 100 times higher than typical background values in and above the marine boundary layer in Florida during Saharan dust episodes. These IN were within the air layer that feeds thunderstorm development and subsequent cirrus anvil formation. Measurements of residual particles from anvil cirrus [Cziczo et al., 2004] and cloud model simulations [van den Heever et al., 2006] all supported the hypothesis that mineral dust modifies cloud microphysics and dynamics. Soot particles from anthropogenic or natural burning processes can also act as IN but their ice nucleating efficiency is much weaker. Another source of IN are biological aerosols such as bacteria, pollen, fungi, spores and even phytoplankton. These bio-aerosols, especially some species of bacteria, are very efficient IN but their number concentrations are generally low.
IN may initiate ice formation via four different modes. Two were mentioned earlier: deposition nucleation and condensation-freezing nucleation, which are controlled mainly by supersaturation. IN may also cause cloud drops or raindrops to freeze by either immersion freezing or contact freezing, and temperature is their main controlling factor. Immersion freezing occurs when an ice nucleus enters a cloud drop (either by acting as a CCN or by collection inside the cloud) and keeps it unfrozen for some time, later initiating freezing when the air gets cold enough. Contact freezing occurs when an ice nucleus collides with a supercooled drop and immediately freezes. *Hoose et al.* [2010a] estimated that, globally, immersion freezing by mineral dust is the dominant ice formation process, followed by immersion and contact freezing by soot. Ice formation from biological aerosols is significant only over certain areas over land (such as forests), where their concentrations are greater than anywhere else in the atmosphere.

Similar to the activation of CN into cloud drops, not all IN can nucleate into cloud ice in a specific cloud environment. Two hypotheses have been applied to explain such a phenomenon. The “stochastic hypothesis” states that nucleation is a probabilistic and time-dependent process, i.e., only a portion of a specific kind of IN can be nucleated during a unit length of time, and the rate of nucleation depends on both particle properties and environmental conditions [cf. *Fletcher*, 1962; *Mason*, 1971]. A contrasting view is called the “singular hypotheses” which considers that nucleation is controlled by impurities or by the number of active sites on the surface of ice nuclei, where each site has a characteristic temperature (for freezing nucleation) or supersaturation (for deposition nucleation) [cf. *Dorsey*, 1948; *Vali and Stansbury*, 1966; *Fletcher*, 1969; *Vali*, 2008]. This means that singular nucleation is time-independent, so a particular kind of IN are nucleated all together once their threshold temperature or supersaturation is reached. Ice nucleation occurring at different temperatures or supersaturations would imply the presence of different IN which can be of the same species and of the same size but possess different surface properties. More detailed discussion about these hypotheses are given by *Vali and Stansbury* [1966], *Vali* [2008], *Chen et al.* [2008] and *Niedermeier et al.* [2010].

These two views of nucleation form the basis for the parameterization of ice nucleation from measurements and for use in cloud microphysical schemes. For example, IN concentrations ($N_{IN}$) in the atmosphere are often measured by sampling air in a cloud chamber
and counting \( N_{CI} \) that form under a given temperature, pressure and humidity. The measured \( N_{CI} \) are then used to represent \( N_{IN} \) by expressing it as a time-independent function of temperature or supersaturation [e.g., Fletcher, 1962; Huffmann, 1973]. Note that such measurements usually cannot identify the type of IN so the functional dependence can be considered as empirical. There are also laboratory measurements of heterogeneous nucleation targeted on specific types of IN [e.g., Yankofsky et al., 1981; Diehl, 2001; Archuleta et al., 2005; Field et al., 2006; Möhler, 2006]. Such results can be used in cloud microphysical models only when the types and size spectrum of IN is known or given. Empirical parameterization of heterogeneous nucleation according to the singular concept has been adopted in cloud microphysical schemes, such as the freezing parameterization from Heymsfield et al. [1993], condensation/deposition freezing from Cooper [1986] and Meyers et al. [1992], and nucleation in the contact mode from Meyers et al. [1992]. Examples of measurements and fitting functions for these ice nucleation modes are given in Figure 5. Note that the above \( N_{IN}-N_{CI} \) relationships have been converted into pseudo-rate equations by taking the rate change of temperature or supersaturation as the functional variable, such as that applied in the spectral bin models of Chen and Lamb [1994] and Khain et al. [2000]. The stochastic approach has also been adopted for interpreting measurement results and for application in regional microphysical schemes. For example, measurements of IN and their immersion freezing capability are often expressed as rate equations [e.g., Bigg, 1953; DeMott, 1990], which were used in several cloud schemes [e.g., Reisin et al., 1998; Diehl and Wurzler, 2004; de Boer et al., 2010; Fan et al., 2010a,b].

The empirical type of ice nucleation parameterization is popular because they can be used without knowing IN types and even \( N_{IN} \). These advantages are also the main weakness of such parameterizations because they cannot respond to the large spatial and temporal variations of natural IN species. The empirical treatment of IN is therefore one of the main uncertainties in current cloud microphysical simulations. Progress has been made toward the improvement of ice nucleation processes and their representation in cloud models. For example, many models have the capability of simulating the emission and transport of mineral dust particles and soot particles. A few recent efforts to simulate bacteria or other bio-aerosols in the atmosphere have also been made [Burrows et al., 2009; Hoose et al., 2010a, b]. New parameterizations of ice nucleating rates based on laboratory measurements are also emerging [e.g., Marcolli et al., 2007; Chen et al., 2008; Fornea et al., 2009; Kanji and Abbatt, 2006; Lüönd et al., 2010].
example, Chen et al. [2008] presented a generalized parameterization which keeps the original mathematical form of classical nucleation theory and includes thermodynamic parameters of the IN (i.e., contact angle and activation energy) derived from laboratory measurements. Furthermore, effects not considered in the measurements, such as solute or curvature effects on heterogeneous freezing, can be easily added by modifying the water activity in the classical nucleation rate formula with the Raoult or Kelvin equation. A comparison of such parameterizations against original measurements for several IN species is shown in Figure 6. Preliminary implementation of this parameterization has been done by Hoose et al. [2010a] to estimate the relative importance of different IN species on a global scale. Integration of such schemes into cloud microphysical models is likely to occur in the near future.

3. Observational Evidence of Processes and Mechanisms behind the AIE

While various mechanisms have been proposed to describe how aerosols affect precipitation, they are essentially tied to two fundamental processes through which aerosols alter energy and water cycles and ultimately influence precipitation: scattering and absorption of solar radiation and serving as CCN or IN. Most, if not all, mechanisms proposed so far [e.g., IPCC, 2007] can be traced either directly or indirectly to these fundamental processes but in different stages of a chain of actions, illustrated in Figure 7. Transformation of aerosol particles into CCN is the core for the majority of the effects.

Among the seven mechanisms illustrated in Figure 7, only the last one is linked to the radiative effect of aerosols, namely its absorption of solar radiation to burn off cloud droplets. All remaining mechanisms originate from the Twomey effect: increases in aerosol concentration leading to increases in CDNC [Twomey, 1977]. For a fixed liquid water path, this leads to smaller cloud particle sizes and thus suppression of precipitation and prolonging of cloud lifetime [Albrecht, 1989]. Cloud height increases [Pincus and Baker, 1994], as does the chance of freezing, which releases extra latent heat to invigorate the development of deep convective clouds and enhance precipitation [Andreae et al., 2004; Rosenfeld et al., 2008]. The aerosol effect on precipitation processes is generally referred to as the second aerosol indirect effect [Albrecht, 1989] which is much more complex than the first indirect effect.
Comparing the latest Intergovernmental Panel on Climate Change (IPCC) report to earlier editions, the number of mechanisms governing aerosol-cloud-precipitation interactions has increased. For example, 5 mechanisms were described in the 2001 IPCC report [IPCC, 2001] and 7 were presented in the 2007 IPCC report [IPCC, 2007]. These mechanisms are essentially contingent upon a series of constraints applied. By adding or changing the constraints, one can, in principle, find more types of aerosol indirect effects. However, identifying them would become an increasingly more daunting task because nature rarely behaves so orderly. Many idealized coincidences are required to observe the hypothesized processes, although some ideal conditions sometimes prevail, allowing us to capture them. In general, cloud and precipitation processes are dictated primarily by the dynamic and thermodynamic settings of the atmosphere and by cloud microphysics as well. Virtually all investigations concerning the AIE, especially observation-based ones, have been fraught with difficulties and uncertainties in removing or decoupling the influences of atmospheric dynamics and thermodynamics.

To comprehend this, Figure 8 illustrates the major processes occurring in an atmospheric column. From the viewpoint of observation studies, we have snap-shots of atmospheric variables including aerosol, cloud and precipitation in a column and we try to relate them as a measure of the AIE. For satellite remote sensing, aerosols below clouds are obscured by clouds. Cloud and precipitation at any point in a spatial-temporal domain results from all priori processes taking place over a much larger domain. Other influential factors need to be taken into account in order to isolate and quantify aerosol effects on precipitation. For example, advection may move a cloud from outside to inside of the (observation) domain, convection that links aerosol with locally-generated clouds, vertical motion and the moisture supply that dictate the growth from cloud droplets to rain drops. The essence of studying the AIE is to establish relationships between newly generated and/or the new growth of existing hydrometers with local aerosols. For stratiform clouds associated with large-scale atmospheric features such as frontal systems, local aerosols would have little to no bearing on clouds and precipitation. For convective clouds, the relationship would be closer.

Despite the complex interaction between meteorological systems and aerosols, with adequate measurements and sound approaches, it is still possible to reveal the AIE by identifying favorable conditions, i.e. enhancing “signals” and suppressing “noise”. Selection and study of dramatic events whose effects on cloud and precipitation exceed those exerted by other factors,
such as ship tracks [Ferek et al., 2000; Coakley et al., 2000], fire smoke [Kaufman and Fraser, 1997; Rosenfeld, 1999; Koren et al., 2004; Andreae et al., 2004], heavy air-pollution [Rosenfeld, 2000], and dust storms [Rosenfeld et al., 2001], can help in this pursuit. Another kind of approach would be to track complete aerosol episodes/cycles during which aerosol loading goes up and down, as does the associated cycle of changes in cloud microphysics (e.g., effective radius) [Schwartz et al., 2002; Feingold et al., 2003].

3.1. Impact on Cloud Droplet Size Distributions and Warm Rain Processes

As stated above, the most fundamental mechanism is the first type of aerosol indirect effect: the decrease in cloud droplet size with increasing aerosol number concentration for a fixed liquid water path (LWP) [Squires, 1958; Twomey, 1977]. Cloud albedo is enhanced in a cloud with constant LWP because there are a greater number of small cloud particles [Twomey, 1977]. This effect has been supported with ample evidence from satellite remote sensing [Han et al., 1994; Wetzel and Stowe, 1999; Nakajima et al., 2001; Liu et al., 2003]. Satellites collect a multitude of aerosol measurements over the entire globe more readily than in situ or ground-based instruments. However, they suffer from several inherent retrieval problems (see Yuan et al., 2008, Li et al., 2009 for recent reviews of the satellite retrieval of aerosol parameters) and a general limitation that cloud and aerosol properties cannot be obtained at the same time over the same location. The majority of previous studies employed large-scale and long-term data, which are elaborated in more detail in the following section. Within a sufficiently large spatial and temporal domain, clear and cloudy scenes can be found but at different times and/or different locations. Given the large-scale variation in aerosols, significant uncertainties can be incurred due to both data mis-matching and small-scale changes in cloud properties.

These limitations can be overcome, or lessened, by ground and in situ observations [Leaitch et al., 1996; Feingold et al., 2003; Penner et al., 2004]. Nevertheless, a large range of variation (by a factor of 3 or more) was found concerning the sensitivity of cloud microphysics to aerosols, as measured by the ratio of the change in cloud particle size to aerosol parameters. While some of the differences are related to the use of different analysis methods and/or observational data [Rosenfeld and Feingold, 2003], natural variation is also observed [Feingold et al., 2003; Kim et al., 2003; Yuan et al., 2008]. Figure 9 shows the changes in cloud droplet
size with aerosol extinction coefficient from ground-based measurements collected at the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) central facility site [Feingold et al., 2003].

The vast majority of previous studies showed a negative dependence of cloud particle size on aerosol loading. Some satellite-based studies showed an opposite relationship [Sekiguchi et al., 2003; Storelmo et al., 2006]. Yuan et al. [2008] conducted an extensive study by examining the relationship between cloud droplet effective radius (DER) and aerosol optical depth (AOD) for a large number of convective cloud scenes acquired around the world by the Moderate Imaging Spectroradiometer (MODIS) onboard the Terra satellite. While inverse relationships between cloud particle size and aerosol loading for different types of aerosols predominated over most of the world, exceptional positive relationships were found in two regions: the Gulf of Mexico and the South China coast during summer when the atmospheric environment over both regions is predominantly influenced by monsoon circulations that blow moist air from oceans to polluted inland areas. The slope of the regression between DER and AOD changes systematically from negative to positive with increasing precipitable water, as shown in Figure 10.

The aerosol effect on precipitation processes is the second type of aerosol indirect effect [Albrecht, 1989]. Aerosols alter warm rain processes by changing cloud microphysics and the LWP. Enhanced aerosol concentrations are generally believed to suppress warm rain processes by inducing small DERs and a narrow droplet spectrum; both inhibit collision and coalescence processes [e.g., Squires and Twomey, 1966; Warner and Twomey, 1967; Warner, 1968; Radke et al., 1989; Rosenfeld, 1999; Liu et al., 2003].

A combination of cloud-top temperatures and effective droplet sizes, estimated from the Advanced Very High Resolution Radiometer (AVHRR), has been used to infer the suppression of coalescence and precipitation processes for smoke [Rosenfeld and Lensky, 1998] and desert dust [Rosenfeld et al., 2001].

It should be pointed out that the above finding, as well as those from all previous satellite-based studies of the AIE, was based on retrievals of DER at the cloud top. This is due to the rapid saturation of the radiometric signal at mid-IR channels [Platnick, 2000; Chang and Li, 2002]. Cloud-top DER is driven by many factors, such as atmospheric stability, available water vapor, and cloud thickness. For adiabatic processes, DER increases linearly with cloud-top
height, which is further linked to potential energy [Rogers and Yau, 1989]. DER profiles can be retrieved using multiple IR channels of varying absorption strength [Chang and Li, 2002, 2003]. Matsui et al. [2004] employed column-mean DERs derived from microwave-based retrievals of LWP and cloud optical depth (COD) and showed a negative correlation between these DERs and the aerosol index (AI). They also investigated the impact of the lower tropospheric static stability (LTSS) on the AIE using observational data (Figure 11). Chen et al. [2007] found systematic biases in the retrieval of LWP when using the conventional visible-IR approach, which casts a doubt on the use of optical remote sensing products alone to investigate the AIE.

It should be stated that an increase in DER does not necessarily lead to an increase in precipitation, i.e., the first indirect effect is not a necessary cause of the second type of aerosol indirect effect, nor does it lead to it. Precipitation is controlled by many factors that are dynamical, thermo-dynamical and microphysical in nature.

3.2. Impact on Ice and Mixed-Phase Clouds – Aerosol Invigoration Effect

While most research into aerosol effects on clouds has focused on low-lying stratiform clouds, some investigators have noted possible aerosol effects on cirrus clouds as well [e.g., Sassen et al., 1995; Stro¨m and Ohlsson, 1998]. However, the findings are much less conclusive. Sherwood [2002] analyzed changes in ice particle effective radius at the top of tropical cumulonimbus clouds (using data from the International Satellite Cloud Climatology Project) with respect to the AI from the Total Ozone Mapping spectrometer and the two quantities were negatively correlated. They attributed this to the smaller size of water droplets that were frozen and lifted to high altitudes by the deep convective cloud system. Chylek et al. [2006] found an opposite trend using MODIS data over the Indian Ocean during winter. Over the same region, Massie et al. [2007] found little change in ice particle size with increase in aerosol loading. By combining MODIS ice cloud particle size, Aura carbon monoxide content (a proxy of aerosols) and TRMM precipitation, Jiang et al. [2008] found that the influence of aerosols on ice clouds is contingent upon moisture content, which is somewhat similar to the finding of Yuan et al. [2008] for warm clouds. In a dry environment, aerosols tend to reduce ice crystal size and suppress precipitation, whereas they have little impact on ice crystal size and precipitation in a wet
environment. Based on this finding, ice effective particle size was parameterized as a function of ice water content and aerosol optical depth for different geographic regions [Jiang et al., 2011].

TRMM data were used to infer the presence of non-precipitating super-cooled liquid water near the cloud top due to over-seeding from smoke over Indonesia [Rosenfeld, 1999] and urban pollution over Australia [Rosenfeld, 2000]. Aircraft measurements have also provided evidence of sustained supercooled liquid water (down to -37.5°C) in continental mixed-phase convective clouds [Rosenfeld and Woodley, 2000]. These findings further suggest that continental aerosols reduce the mean size of cloud droplets, suppressing coalescence and warm-rain processes, permitting more freezing of cloud droplets and associated latent heat release above the 0°C isotherm, and enhancing the growth of large hail and cold-rain processes [Rosenfeld and Woodley, 2000]. This is now known as the aerosol invigoration effect [Andreae et al., 2004], which is illustrated in Figure 12.

Andreae et al. [2004] analyzed in situ observations made during the Large-Scale Biosphere-Atmosphere Experiment in Amazonia-Smoke, Aerosols, Clouds, Rainfall, and Climate campaign and found that increases in smoke and surface heat due to biomass burning tend to lead to higher cloud-top heights and the enhancement of cold-rain processes over the Amazon basin. This is shown in Figure 13 with in situ measurements made over the Amazon and Thailand. The “Blue” and “Green” cases in Figure 13 represent clean clouds over the ocean and the Amazon, respectively. The two “Pyro” cases are the most extreme forms of smoky clouds. All other cases except for the two “Thai” cases are clouds with moderate concentrations of smoke. Substantial differences in the cloud drop size, $D_L$, are evident. The $D_L$ at 1000 m above the cloud base is only 10 μm for extremely polluted cases, compared to 25 μm for clean cases. Precipitation occurs at 1000-1500 m above the cloud base for clean cases. But the precipitation threshold is over 6000 m for pyro-clouds. Observations made in Thailand on May 2, 1998 show that clouds had to exceed 6400 m to produce warm rain under smoky conditions. Once these clouds extended above the freezing level, they become quite vigorous with intensive showers, thunderstorms, and sometimes hail occurred. These observations indicate that the shift of the onset of precipitation to a higher level under polluted conditions leads to enhanced formation of ice hydrometers and more vigorous convection. Koren et al. [2005] examined cloud properties derived from the MODIS and found strong evidence that aerosols from pollution, desert dust and biomass burning systematically invigorate convective clouds over the Atlantic Ocean.
A theoretical explanation for such an effect is given by Rosenfeld et al. [2008] using a parcel model to demonstrate the buoyancy changes under different conditions (Figure 14). The buoyancy of the following scenarios is shown: (1) suppressing rainfall and keeping all condensed water load, without freezing; (2) precipitating all condensed water, without freezing; (3) precipitating all condensates, with freezing at $T < -4^\circ C$; (4) suppressing precipitation until $T = -4^\circ C$, and then freezing and precipitating all condensed water above that temperature. This model result suggests that the suppression of warm rain but enhancement of cold rain processes can greatly invigorate the convection.

The importance of cloud top and base heights in determining the strength of the invigoration effect can also be seen in this figure. If cloud particles never reach above the freezing level, there will be no invigoration effect. If the cloud base is high, for example, very close to the freezing level, there will not be very much liquid water to release latent heat since the distance for cloud particles to grow before reaching the freezing level is too short. Freezing level is thus a key factor whose height can be determined with a newly proposed method using passive satellite imaging data like the MODIS [Yuan et al. 2010], while it remains a challenge to determine cloud-base heights using passive satellite data.

The invigoration effect is driven by latent heat release which is rooted to the first indirect effect. This effect, in turn, further depends on aerosol hygroscopic properties, which may be collectively referred to as the aerosol microphysical effect. For absorbing aerosols like smoke, the thermodynamics of the atmosphere can be altered through heating of the aerosol-laden atmosphere and cooling of the surface. This enhances atmospheric stability and reduces the moisture content due to evaporation [Ackerman et al., 2000], thus inhibiting clouds and precipitation. For smoke aerosols that have the capability of serving as CCN, the two competing effects lead to boomerang-shaped responses of cloud-top height and cloud fraction as theorized and confirmed by Koren et al. [2008] over the Amazon region (Figure 15).

3.3. Impact of Aerosols from Urbanization on Thunderstorms and Lightning

As the major emission source of anthropogenic aerosols, urban areas may disturb the atmospheric environment enough so that regional climate is modulated. There are two types of major disturbances: land cover changes and atmospheric emissions. The latter includes both
aerosol and greenhouse effects. The greenhouse effect generally has a much larger spatial and long-term impact than the aerosol effect. Aerosol and land cover changes can leave strong footprints, but it is a non-trivial task to separate the effects of these two changes.

Enhancement of rainfall downstream of some major metropolitan areas has been documented for a long time, such as Chicago [Changnon et al., 1981], Tokyo [Ohashi and Kida, 2002], Houston [Shepherd and Burian, 2003 – see Figure 16], and New York City [Jin et al., 2005]. High aerosol concentrations in urban environments could affect precipitation variability by providing an enhanced source of CCN. Hypotheses have been developed to explain the effect of urban regions on convection and precipitation [van den Heever and Cotton, 2007; Shepherd, 2005]. It is unclear if the urban heat island effect or aerosols, or both, cause the enhancement. With the theory of aerosol invigoration, we may now take a new look at this phenomenon. Heavily polluted cities are an ideal setting for inducing/enhancing invigorated clouds, especially cities near lakes and along coastlines in summertime. During this season, warm moist air can blow from seas/lakes to land, providing the following conditions favorable to cloud invigoration: (1) a plentiful supply of highly hygroscopic anthropogenic aerosols, (2) the heat island effect which fuels initial convection, and (3) moist, hot air that can provide large amounts of potential energy to sustain strong convection.

This seems to corroborate the finding of a mid-week anomaly in precipitation and convection, and intensified lightening over urban areas. Using long-term integrated TRMM-derived precipitation data, Bell et al. [2008] found a significant mid-week increase in summertime afternoon thunderstorms over the southeastern U.S., which coincides with a mid-week increase in ground-measured aerosol concentration, wind convergence at 1000 hPa and divergence at 300 hPa, and vertical velocity at 500 hPa (see Figure 17). These coincidental findings are consistent with the notion that aerosols have a major impact on the dynamics of continental mixed-phase convective clouds [Rosenfeld and Woodley, 2000; Williams et al., 2002].

Enhancement of the lightning flash density was found to be centered over and downwind of the metropolitan area [Orville et al., 2001], as shown in Figure 18. Possible causes include enhanced convergence associated with the urban heat island effect and altered microphysical processes associated with anthropogenic pollution. Steiger and Orville [2003] compared 14 years’ worth of cloud-ground flash density distributions with the locations of PM10 (particulate matter less
than 10 µm in diameter) sources in Louisiana. They concluded that pollution plays a key role in lightning enhancement, while urban and sea breeze effects are negligible. Yuan et al. [2011b] examined 2005 lightning data from the lightning-imaging sensor onboard the TRMM satellite and MODIS AOD (Fig. 219). They found anomalously high lightning activity in the presence of high aerosol loading, which was traced to volcanic activity; no connection was found with any anomalies in meteorology. They further quantified the response of lightning to AOD: a ~60% increase in aerosol loading leads to more than a 150% increase in lightning flashes. Aerosols could influence lightning activity through modification of cloud microphysics. Cloud ice particle sizes are reduced and cloud glaciation is delayed to colder temperatures when aerosol loading is increased. Despite the seemingly convincing evidence linking aerosols with lightning activity, their real connection is far more complex, especially in light of many other influential factors associated with urban regions, as demonstrated by simulations with a state-of-the-art land surface model and a cloud-resolving model [see Section 5 and van den Heever and Cotton, 2007].

4. Large-scale Observational Diagnostics of the Aerosol Effect on Precipitation

In general, a main purpose of large-scale observational analyses of aerosol-precipitation relationships is to reveal statistically significant signals that may suggest aerosol effects on the large-scale distribution and variability of precipitation (i.e., through their radiation and CCN effects). This is different from those discussed in section 3 that seek physical processes governing the AIE. On large scales (especially the global scale), total precipitation is constrained by the moisture budget, which is mainly dictated by surface evaporation. Unless aerosol effects alter surface evaporation, they may only modify the timing and location of precipitation, but not its total amount. A local signal of aerosol effects on precipitation is usually a manifestation of such changes in timing and location of precipitation.

4.1. Isolating Aerosol Effects

Large-scale observational analyses of aerosol effects on precipitation face no less challenges than theoretical, modeling and field observational approaches discussed elsewhere in this review. All large-scale observational studies on any subject, they must adequately address issues of data
quality, analysis methods, robustness of the results (i.e., statistical significance), and causality. The issue of causality manifests itself in observational studies of aerosol effects on precipitation in two problems: aerosol effects on precipitation vs. precipitation effect on aerosols (i.e., rainout and washout), and effects on precipitation by other meteorological factors co-varying with aerosol. The first problem makes any local (point-wise) instantaneous relationship between aerosols and precipitation useless as evidence of aerosol suppressing effect on precipitation. The second problem invalidates any approach using aerosols to investigate its effects on precipitation without considering other factors. Dealing with both problems demands extra care in data analysis beyond the usual statistical correlation and map discussions. Some previous observational studies on aerosol effects on precipitation did not present sufficient evidence to attribute their findings revealing true aerosol-cloud-precipitation interaction (ACPI). As stated above, separating the ACPI from other effects has always been a challenge.

The simplest way to remove the rainout/washout effect on aerosol and cloud contamination from an aerosol-precipitation analysis is to avoid local (point-wise) relationships between the two. Aerosols upwind can be included in an analysis based on back-trajectory calculations [Hui et al., 2008; Huang et al., 2009c]. An equivalent approach is to including time lags with aerosol signals leading precipitation [Lau and Kim, 2006; Bollasina and Nigam, 2009]. Another method is to exclude aerosol data near cloud scenes or within a grid box with a large cloud fraction [Koren et al., 2010a]. In a target precipitation area that is remote from aerosol sources, fluctuations in aerosols independent of precipitation should be the same as those outside and upwind of the precipitation region. Based on this, Huang et al. [2009a, b] used aerosols averaged over a large domain containing rain bands over the tropical Atlantic Ocean and West African monsoon region and broad areas of no precipitation to both the south and north to represent fluctuations in aerosols in the rain band regions (Figure 20a).

Several methods have been used in previous studies to isolate the aerosol effect on precipitation from others. One is to examine the variability of other factors (e.g., water vapor) that may potentially cause the observed fluctuations in precipitation. Based on this method, if there is co-variability between precipitation and aerosols, aerosol effects on precipitation are only suggested. But if there is no coherent variability between precipitation and other variables that may suggest their roles in the observed variability in precipitation, observed aerosol-
precipitation co-variability might indicate aerosol effects on precipitation. The following examples illustrate some approaches recently taken.

Qian et al. [2009] observed a decreasing trend in light rain (< 10 mm day\(^{-1}\)) and an increasing trend in heavy rain (>50 mm day\(^{-1}\)) over East China from 1956 to 2005. This is the same period of large increases in pollution in the same part of China (Figure 21). The possibility that the trend in light rain has been caused by changes in large-scale meteorological conditions was ruled out because there is no correlation between the decreasing trend in light rain with water vapor transport. They subsequently concluded that the increase in aerosols due to pollution observed during the same period has likely caused the observed decreasing trend in light rain through the known suppressing effect and shifted the rain distribution from light to heavy rain. They didn’t discuss, however, whether the change in the total amount of water corresponding to the observed trend in light rain over the period is within the uncertainties in the data of water vapor transport (based on a global reanalysis).

Wilcox et al. [2010] observed a northward shift of rainfall in the Intertropical Convergence Zone (ITCZ) over the tropical Atlantic Ocean during African dust outbreaks. They examined differences in temperature profiles and lower-tropospheric zonal wind over the tropical Atlantic Ocean between African dust outbreaks and low dust conditions. They concluded that the northward shift of rainfall is caused by lower-tropospheric warming that is attributable to advection of the warm Saharan Air Layer (SAL), enhanced subtropical subsidence, and radiative heating of dust.

Similarly, Lau and Kim [2006] compared composites and lag-regressions of aerosols, wind, temperature and precipitation over the Indian summer monsoon region. They claimed that they have found observational support of their “elevated heat pump” hypothesis, namely, the anomalous high concentration of absorbing aerosols during the pre-monsoon season is associated with anomalous warming that is responsible for changes in the large-scale circulation leading to an increase in monsoon rainfall over India. In these two studies, possible contributions of aerosols to the observed variability of rainfall was compared to but not separated from other factors. Therefore their hypothesis is plausible but not without ambiguity, even though the observed signals are consistent with those numerical simulations (see section 6), where physical processes are much clearer. Same sets of observations can be subject to different interpretations and result in contradictory conclusions [Nigam and Bollasina, 2010].
The inadequacy of examining the co-variability between aerosols, precipitation and other factors has been explicitly explored. In a re-examination of the rainfall trend over Israel, Alpert et al. [2008] pointed out that urban dynamics (e.g., the heat island effect), if not fully taken into consideration, may lead to erroneous interpretations of data and the incorrect conclusion that an observed decreasing rainfall trend is due to the suppression effect of increasing urban pollution from upwind regions. Through comparing sub-monthly evolution patterns of aerosols, precipitation, and other meteorological variables during the pre-monsoon period over South Asia, Bollasina and Nigam [2009] reproduced the earlier monthly observations [Bollasina et al., 2008] that anomalous aerosol build-up in May can delay the monsoon rainfall onset due to the radiative effect (i.e., cooling of the land surface). But they acknowledged that the pervasive influence of advection precludes a robust analysis of the aerosol impact.

Various statistical tools have been used either to remove the dependency of precipitation on other factors before its relationship with aerosols is established, or to determine the importance of different factors in the observed variability in precipitation to evaluate the role of aerosols vs. others. The following are examples of such a practice.

In an attempt to identify the aerosol effect on rainfall over the Amazon basin during the biomass-burning season (September 2006), Jones and Christopher [2010] applied the principle component analysis (PCA) method, a commonly-used statistical tool in geoscience. This method encompasses precipitation and 23 other parameters including meteorological environmental variables (winds, humidity, and temperature), cloud information (optical thickness, top pressure, liquid water path, fraction, and droplet effective radius), aerosol optical thickness, and locations. They interpreted the weighting of each parameter and its sign relative to precipitation in each principle component (PC) as a signal of the potential physical connection between the two. Based on this interpretation, they found atmospheric environmental variables (e.g., zonal wind, relative humidity, lapse rate, and vertical motion) are the dominant factors for rainfall variability, with their large weighting for PC1 that explains 25% of total rainfall variance. Following the same argument, they identified the radiative effect of aerosols as the second dominant factor for rainfall that is represented by PC2 (15% of total rainfall variance). Their results suggest a suppressing effect of smoke on rainfall. The orthogonality requirement for each PC may, however, obscure the physics intended to be disclosed.

In a study of the aerosol effect on precipitation over the tropical Atlantic Ocean and in the
West African monsoon region, Huang et al. [2009a, b, d] first used a multi-variable linear regression method to remove the variability in aerosols and precipitation that is coherent with water vapor and SST related to known climate modes [El Niño / Southern Oscillation (ENSO), North American Oscillation (NAO), and Tropical Atlantic Variability (TAV)]. The resulting anomalous time series of aerosols and precipitation are considered linearly independent of other weather and climate factors. They then compared the spatial distribution and seasonal cycle in anomalous precipitation between months of high vs. low aerosol anomalies (top and bottom terciles of all months) averaged over a large domain (Figure 20a). The comparison revealed decreases in precipitation during months of high aerosol anomalies at the southern edge of the rain bands during boreal winter, the season of frequent biomass burning over equatorial and southern Africa (Figure 20b and c). The spatial pattern and the seasonal cycle of the rainfall reduction do not match those of expected aerosol wet deposition. Numerical simulations with only the radiative effect of black carbon aerosol included have reproduced this observation (see Section 6). They concluded that the reduction in rainfall is due to the radiative effect of smoke from biomass burning. The linear regression can account only simultaneous relationships. So the removal of the coherent variability between precipitation and other factors may not be complete.

A highly recommended method of isolating the aerosol effect from others is to examine aerosol-precipitation relationships in different regimes [Stevens and Feingold, 2009]. Regimes can be defined in various ways depending on data and analysis targets. They include convective characteristics (e.g., cloud working function), environmental conditions for cloud (e.g., static stability), or cloud microphysics (e.g., liquid water path, ice path), as demonstrated in the following examples.

In a study of the aerosol effect on rainfall during the biomass burning seasons (August–October of 2000 and 2003) in the Amazon, Lin et al. [2006] examined relationships between aerosol optical depth versus precipitation and clouds stratified by bins of cloud working function (CWF). CWF is defined as a vertical integration of work by buoyancy per unit mass flux, or the kinetic energy generation per unit mass flux [Arakawa and Schubert, 1974]. Instigation of convection in some numerical models is determined by CWF passing a given threshold. CWF was used in this study to define regimes of the large-scale meteorological forcing for rainfall and convection. The CWF-stratified analysis indicates that elevated AOD was associated with increases in precipitation, occurrence of intense rainfall events, cloud cover, cloud-top height,
water path, and formation of ice. This is in contrast to the result from Jones and Christopher [2010] discussed above.

Sorooshian et al. [2009] proposed to stratify the environmental effect in terms of LWP in a concept of precipitation susceptibility, which is defined as $S = -\frac{d \ln R}{d \ln N_d}$, where $R$ is the precipitation rate and $N_d$ is the CDNC or a sub-cloud CCN proxy. Using data from satellite, a cloud parcel model and large eddy simulations, they demonstrated different vulnerabilities of precipitation to aerosols in three regimes: (1) low LWP, where clouds do not precipitate and are relatively insensitive to $N_d$, (2) intermediate LWP, where precipitation forms by the interaction of cloud drops and is progressively more effectively suppressed by increasing $N_d$, and (3) high LWP, where susceptibility begins to decrease because there is ample LWP to drive the precipitation process. To remove ambiguity in whether the microphysical response is a result of a change in aerosols or a change in meteorological conditions (e.g., available liquid water), Sorooshian et al. [2010] further proposed to separate the precipitation susceptibility into parts: $S_r = - \left( \frac{\partial \ln R}{\partial \ln r} \right) \left( \frac{\partial \ln r}{\partial \ln N_d} \right)$ and $S_\tau = - \left( \frac{\partial \ln R}{\partial \ln \tau} \right) \left( \frac{\partial \ln \tau}{\partial \ln N_d} \right)$, where $r$ is the cloud effective radius and $\tau$ is the cloud optical depth. They concluded that the separation provided more confidence in the results from the original precipitation susceptibility.

Panicker et al. [2010] calculated the aerosol indirect effect, $AIE = \frac{d \ln r}{d \ln t}$ [Feingold et al., 2003], for July and September 2001 – 2004 in four Indian summer monsoon regions. Similar to Yuan et al. [2008], they used MODIS AOD data of fine mode fraction (FMF), cloud liquid water path (CLWP), cloud ice path (CIP), cloud ice radius (CIR) and cloud water radius (CWR). They used binned CWLP and CIP to define regimes. They found a positive AIE (the Twomey effect) for fixed CWLP and CIP in years when the monsoon rainfall was below normal, but a negative AIE (the anti-Twomey effect) in years of normal and above-normal rainfall. They suggested that the AIE depends on large-scale circulation patterns that may advect different types of aerosols from different sources. Similar to earlier discussions, without a careful water budget analysis, ambiguity in the aerosol effect on precipitation on the seasonal timescale would always be present.

From the above discussion, it is obvious that each approach of isolating the aerosol effect on precipitation from other factors has its pros and cons. Large-scale diagnostics of the aerosol effect on precipitation, through either radiation or CCN or both, are incomplete without considering microphysics (clouds, aerosols, and precipitation), mesoscale cloud dynamics, and
the large-scale background environment (aerosol transport, water budget). This requires unprecedented collaboration and coordination among sub-disciplines within the atmospheric science field.

4.2. Applications of Satellite Data in Revealing ACPI

Satellite observations have the unique advantage of global coverage (in comparison to in situ observations). Along with analyses of ACPI using more conventional large-scale dynamical and thermodynamical data together with aerosol data demonstrated above [Ramanathan et al., 2001; Menon et al., 2002; Chuang et al., 2002; Lau et al., 2006], observational evidence has been obtained from satellite observations following somewhat different approaches. Satellite observations generate a large ensemble of data that can effectively filter out the day-to-day weather variability in order to single out signals of the AIE, as demonstrated using earlier satellite data [Han et al., 1998; Wetzel and Stowe, 1999; Nakajima et al., 2001, Liu et al., 2003]. The advent of multiple advanced EOS sensors aboard the A-Train [e.g., Koren et al., 2008; L’Ecuyer et al., 2009; Niu and Li, 2011a, b] has substantially increased this capability. The most tangible and direct evidence of the AIE is linked to the microphysical properties of clouds, which are more readily measured by satellite observations despite various retrieval uncertainties and limitations [Li et al., 2009].

Among the handful of microphysical variables retrievable from satellites (DER, LWP, COD), LWP plays the most significant role in precipitation [Sorooshian et al., 2009], even though these variables are not independent. Using matched rainfall data from CloudSat and cloud microphysics from MODIS aboard the A-Train, Chen et al. [2011] examined the predictability of rainfall from warm clouds. They found that LWP is by far the dominant factor, followed by DER and COD, as illustrated by the prediction skills in Table 2. As such, detection of any effects of aerosols on precipitation may begin with their effects on key cloud variables.

Given the dominant influence of LWP (or its primary proxy, COD) on precipitation, it is critical to understand whether and how aerosols impinge on LWP, which is much less conclusive than the effect on DER. Controversial findings have been reported showing either positive relationships [Sekiguchi et al., 2003; Storelvmo et al., 2006; L’Ecuyer et al., 2009], negative relationships [Twohy et al., 2005; Matsui et al., 2006; Brenguier et al., 2000], or mixed results.
[Nakajiman et al., 2001; Coakley and Walsh, 2002; Han et al., 2002; Kaufman et al., 2005]. Several mechanisms have been postulated. The enhancement of entrainment and droplet evaporation by aerosols reduces LWP [Ackerman et al., 2004], whereas suppression of coalescence inhibits precipitation and thereby increases LWP [Lebsock et al., 2008]. When in situ aerosol and cloud measurements are utilized, the scavenging effect of drizzle can remove aerosols, leading to a seemingly negative relationship that undermines any real causal relation [Twohy et al., 2005].

Taking advantage of the unprecedented superior quality and quantities of A-Train sensors [Stephens et al., 2002], thorough and in-depth investigations of the AIE for warm clouds have been conducted under a wide range of conditions on a global scale [Lebsock et al., 2008; Berg et al., 2008; L’Ecuyer et al., 2009; Sorooshian et al., 2009; Niu and Li, 2011a, b]. Using precipitation data from CloudSat, cloud microphysics from MODIS and LWP and column water vapor (CWV) from AMSR-E, as well as meteorological analysis data from the ECMWF, Lebsock et al. [2008] and L’Ecuyer et al. [2009] attempted to sort out various factors influencing warm rain or drizzle using a combination of satellite data, a global reanalysis product LTSS), and a global aerosol transport model (distributions of aerosol species in both cloudy/raining and clear-sky scenes). As shown in Figure 22, the probability of precipitation increases with increasing LWP and CWV, and is moderately affected by atmospheric stability. For the same LWP, CWV and stability, dirty clouds are less likely to rain than clean clouds. Not all types of aerosol particles inhibit precipitation. One exception is sea salt aerosols.

Analyzing MODIS aerosol products (optical depth over land and aerosol index over oceans) over the entire global tropics (between 20°N and 20°S) where convective clouds are more dominant than in the mid- to high-latitudes, Niu and Li [2011a] revealed a more general pattern of the AIE. Aerosols were found to suppress light rain associated warm clouds but enhanced heavy rain associated with mixed-phase clouds, as shown in Figure 23. This is because aerosols significantly invigorate convection mainly through ice processes, while precipitation from liquid clouds is suppressed through aerosol microphysical processes, consistent with their findings based on long-term ground-based measurements [Li et al., 2011c]. It is worth noting that CloudSat is particularly good at detecting drizzle, but not as capable of detecting moderate to heavier rainfalls due to its high-frequency radar of 95 GHz. This is remedied by the TRMM satellite. It carries two rain sensors: the TRMM Microwave Imager (TMI) and the Precipitation
Radar (PR) at 35 GHz with sensitivities to different rain attributes. The TMI is most sensitive to LWP, while the PR is sensitive to large drops. By virtue of these differences in sensitivity, Berg et al. [2006, 2008] discovered a region where the largest discrepancies between the two sets of rain estimates was found, which was conjectured to be a smoking-gun of the strongest AIE on precipitation. It is located around the coast of East Asia and the western Pacific (Figure 24), the outlet of pollutants and dust from the world’s most-populated region: Asia. While continental-scale transport can bring some aerosols generated over Europe and Asia to this region [Chin et al., 2004], local emissions are the largest source of aerosols, as shown by extensive measurements of aerosol optical, physical, chemical properties collected during two major field experiments conducted in China [e.g., Li et al., 2007, 2011a]. Using the 2008 intensive field campaign data acquired by the ARM Mobile Facility in southeast China [Li et al., 2011a], Fan et al. [2011] confirmed the significant role played by aerosols in precipitation. This was corroborated by the analysis of Qian et al. [2009].

5. Cloud-resolving Model Simulations

The statistical significance of precipitating mesoscale convective cloud systems (MCS) has been quantified from space in terms of its correlation with rainfall estimated from TRMM precipitation radar and microwave imager retrievals. Figure 25 shows that MCSs are the dominant heavy-rain producers in the tropics and subtropics; they generate more than 50% of in the rainfall in most regions. The average annual rainfall exceeds 3 mm day$^{-1}$ and up to 90% of the rainfall occurs over certain continental areas, e.g., the La Plata Basin of South America. This section will mainly review cloud resolving model (CRM) simulations associated with MCSs to examine the sensitivity of aerosol concentrations on precipitation processes (rainfall). Other cloud-scale modeling simulations (including large-eddy simulations) associated with warm clouds, stratocumulus, orographic clouds and Arctic stratus clouds can be found in Levin and Cotton [2009].

5.1 Major Methodology

5.1.1. Cloud-Resolving Model
One of the most promising methods to test the representation of cloud processes used in climate models is to use observations together with CRMs [see a recent review by Tao and Moncrieff, 2009]. The CRM has to be non-hydrostatic, and its flow can either filter out (anelastic) [Ogura and Phillips, 1962] or allow (compressible) [Klemp and Wilhelmson, 1978] the presence of sound waves. Sound waves are not important in thermal convection, but their presence can place severe restrictions on the time step in numerical integrations. For this reason, most CRMs use an anelastic system of equations in which sound waves have been removed by neglecting the local variation of air density with time in the mass continuity equation. A 3-D diagnostic (elliptic) pressure equation can be solved using direct (e.g., Fourier transform) or iterative methods. In the compressible system, a very small time step (2 sec for a 1000-m spatial resolution) is needed for time integration due to the presence of sound waves. However, Klemp and Wilhelmson [1978] developed a semi-implicit time-splitting scheme, in which the equations are split into sound-wave and gravity-wave components to achieve computational efficiency. One advantage of the compressible system is its numerical code remains a set of explicit prognostic equations and adjustments due to factors such as surface terrain can be incorporated into the numerical model without complicating the solution procedure.

Compared to typical global circulation and climate models, CRMs use more sophisticated and relatively realistic representations of cloud microphysical processes, and they can reasonably well resolve the structure and life cycles of clouds and cloud systems (with sizes ranging from about 2-1000 km). CRMs also allow for explicit interaction between clouds, outgoing longwave (cooling) and incoming solar (heating) radiation, as well as ocean and land surface processes [see a review by Tao, 2003, 2007; Tao and Moncrieff, 2009]. Observational data are required to initialize CRMs and to validate their results. The GCSS Science Team [1993] and Randall et al. [2003] recommended that improved CRMs should be used as a test bed to develop and evaluate cloud parameterizations in large-scale models.

There are several major advantages in using CRMs to study the interactive processes between clouds, precipitation and aerosols. For example, the use of a fully explicit microphysics scheme (liquid and ice) and a fine horizontal resolution can provide relatively realistic cloud optical properties, which are crucial for determining radiation budgets. With a high spatial resolution, each atmospheric layer is considered either completely cloudy (overcast) or clear; no
partial cloudiness is assumed. In addition, the applied microphysics scheme is realistic enough to simulate the life cycle of clouds, precipitation and aerosols. The CRM can also have better cloud dynamic processes in terms of transport of aerosols [see the review by Thompson et al., 1997; Ekman et al., 2004, 2006; Yin et al., 2005].

5.1.2. Microphysics Used in CRMs

In order to use a CRM to study the impact of aerosols on cloud and precipitation processes, CCN activation and IN nucleation need to be considered in its microphysical scheme (see Section 2). Two-moment bulk and spectral bin microphysics schemes are required to study the impact of the CCN, GCCN and IN on cloud and precipitation formation. One of the major differences between the two-moment bulk and spectral bin microphysics schemes concerns the representation of cloud particle sizes (Figure 26). Two-moment schemes typically combine the main features of single-moment schemes by calculating the mixing ratios of hydrometeors, and then adding additional variables for the number concentrations of all particles. Warm-cloud (ice-free) microphysics assumes a bimodal population of water particles. One population is for small cloud droplets whose terminal velocities are negligible compared to the vertical velocity of air, and one is for large raindrops that have significant fall speeds. Ice microphysics typically assumes three types of particles: small cloud ice whose terminal velocity is negligible, snow whose terminal velocity is about a few tens of centimeters, and large graupel or hail that fall even faster. Only recently have some CRMs included frozen drops/hail as a fourth kind of particle. Graupel has a low density and a high intercept in the Marshal-Palmer distribution (high number concentration) while hail has a high density and a low intercept. Only raindrops, snow, graupel and hail have a chance to reach the ground level. Graupel is representative for tropical oceanic convection and hail for mid-latitude storms [McCumber et al., 1991]. More than 25 transformations occur among water vapor, liquid particles and ice particles, such as growth of ice crystals by vapor deposition and riming, the aggregation of ice crystals, the formation of graupel and hail, the growth of graupel and hail by the collection of super-cooled rain drops, the shedding of water drops from hail, the rapid growth of ice crystals in the presence of super-cooled water, and the melting and sublimation of all forms of ice (see Figure 27). Certain mathematical functions such as Khrgian-Mazin distribution function for cloud drops and Marshall-Palmer distribution
function for raindrops, snow and graupel/hail are assumed to represent size distributions.

With increasing computing power, explicit bin-microphysical schemes have been developed for CRMs to study cloud-precipitation-aerosol interactions. The formulation of explicit microphysical processes is based on solving equations for mass advection by condensation growth and stochastic collision kinetics for the size distribution functions of water droplets (cloud droplets and raindrops together as one category). Ice particles are much more complicated due to their different shapes, so they are often classified into various growth habits such as columnar, plate-like, dendrites, snowflakes, graupel, and frozen drops. Each type is described by a discretized size distribution containing as many as 30 or more categories (bins). Spectral bin microphysics include the following processes: (1) activation (nucleation) of cloud droplets, (2) nucleation of ice particles [Pruppacher and Klett, 1997; Meyers et al., 1992], including homogeneous freezing, deposition nucleation, condensation-freezing nucleation, immersion freezing [Vali, 1994] and contact freezing, (3) ice multiplication [Hallett and Mossop, 1974; Mossop and Hallett, 1974], (4) detailed melting [Khain et al., 2004], (5) condensation/evaporation of liquid drops [Pruppacher and Klett, 1997; Khain et al., 2000], (6) deposition/sublimation of ice particles [Pruppacher and Klett, 1997; Khain et al., 2000], (7) drop/drop, drop/ice, and ice/ice collision/coalescence [Pruppacher and Klett, 1997; Pinsky et al., 2001], (8) turbulence effects on liquid drop collisions [Pinsky et al., 2000, 2008], and (9) collisional breakup [Low and List, 1982; Seifert et al., 2005]. Sedimentation of liquid and ice particles is also considered, and for the latter, crystal shape is an important factor. Spectral bin microphysics is specially designed to take into account the effect of atmospheric aerosols on cloud development and precipitation formation by calculating the activation of aerosols explicitly according to the Köhler theory [Pruppacher and Klett, 1997]. Section 2 describes in more detail processes associated with the initialization of CCN, GCCN and IN.

5.2. **CCN Effect on Clouds**

It is commonly believed that clouds in a clean environment (low CCN concentration) produce fewer droplets but with larger sizes due to greater condensational and collectional growth, leading to a broader size spectrum in comparison to the high CCN situation. Figure 28 shows simulated and observed cloud drop size distributions under low and high CCN concentrations
[Tao et al., 2007]. van den Heever et al. [2006] and Carrio et al. [2007] also found that smaller cloud droplets are found with a narrow spectrum under dirty conditions. The numerical results are in good agreement with observations, indicating that the microstructure of clouds depends strongly on cloud-aerosol interactions. The Twomey effect was well simulated. The width of the drop size distribution could have an impact on precipitation processes. For example, smaller cloud droplets would reduce the chance to form raindrops from cloud-drop coagulation. This result is in agreement with observations [i.e., Twomey et al., 1984; Albrecht, 1989; Rosenfeld, 1999].

The effects of increases in aerosol concentration on cloud lifetime for warm convective clouds have been studied by using CRMs and large eddy simulations (a special type of cloud model with a resolution of 100 m or smaller). Teller and Levin [2006] showed that a polluted cloud produces less precipitation and increases the lifetime of the cloud under the same meteorological conditions. This result is in agreement with the second aerosol indirect effect on cloud lifetime [Albrecht, 1989; Ackerman et al., 2000]. However, a separate modeling study shows that an increase in aerosol concentration from very clean to very polluted conditions does not increase cloud lifetime (Figure 29), even though precipitation is suppressed. In rare cases (two in Figure 29), long-lived polluted clouds may have longer lifetimes than clean clouds due to the merging of individual clouds. This result is contrary to the observation that increases in aerosol concentration leads to prolonging of cloud lifetime [Albrecht, 1989]. But the model result does agree with observations of precipitation suppression under polluted conditions. Jiang et al. [2006] proposed that the small changes in cloud lifetime are due to competing effects of precipitation suppression and enhanced evaporation for shallow clouds. The differences between these two numerical modeling studies could be due to the differences in aerosol concentration (mildly polluted versus highly polluted) and/or environmental conditions.

5.3. CCN Effect on Precipitation

5.3.1. Impact of CCN on Surface Precipitation

Recently, many CRMs have been used to examine the role of aerosols on mixed-phase convective clouds (Table 3). These modeling studies have many differences in terms of model configuration (two- or three-dimensional), domain size, grid spacing (150 – 3000 m),
microphysics (i.e., two-moment bulk, simple or sophisticated spectral-bin), turbulence (1st or 1.5 order turbulent kinetic energy (TKE)), radiation, lateral boundary conditions (i.e., closed, radiative open or cyclic), cases (isolated convection, tropical or mid-latitude squall lines) and model integration time (e.g., 2.5 to 48 hours). A simple metric, changes in time-integrated precipitation ($dP = 100 \times (P_{dirty} - P_{clean}) / P_{clean}$) as a result of increases in the number concentration of CCN ($dN_0 = N_{dirty} - N_{clean}$), was used to examine the impact of aerosol concentration on surface rainfall (Table 4). Among these modeling studies, the most striking difference is that cumulative precipitation can either increase or decrease in response to higher concentrations of CCN. Phillips et al. [2002], Khain et al. [2004, 2005], Khain and Pokrovsky [2004] and Teller and Levin [2006] changed the number concentrations of CCN gradually and found robust decreases in cumulative precipitation for higher concentrations of CCN. This is completely opposite from the results from Wang [2005], Khain et al. [2005], Lee et al. [2009a], and Fan et al. [2007b].

Tao et al. [2007] used a two-dimensional CRM with detailed spectral-bin microphysics to examine the aerosol impact on a tropical oceanic mesoscale convective system, a summertime mid-latitude continental squall line and a short-lived Florida sea breeze convective storm. Rain suppression in the high CCN concentration runs is evident in all three case studies but only during the first hour of the simulations (Figure 30). Rain reaches the ground early in all the clean cases. This result suggests that microphysical processes dominate the initial stage of cloud development, during which the transition of cloud droplets from condensational to collectional growth is very sensitive to cloud drop size. Compared to the case under polluted conditions, clouds in a clean environment (low CCN concentration) produce fewer but larger cloud droplets, leading to a better chance for raindrop formation from cloud-drop coagulation. This result is in good agreement with many observations [e.g., Rosenfeld, 1999, 2000] and other CRM studies [Khain and Pokrovsky, 2004; Khain et al., 2005; Teller and Levin, 2006; Seifert and Beheng, 2006; van den Heever et al., 2006; Carrio et al., 2007; van den Heever and Cotton, 2007; Lee et al., 2009a], and is in agreement with Twomey’s second indirect effect. During the mature stage of the simulations, the effect of increasing CCN concentration ranges from rain suppression in the mid-latitude continental case to little effect in the Florida sea breeze case to rain enhancement in the Pacific oceanic case. These results suggest that model simulations of the whole life cycle of a convective system are needed in order to assess the impact of aerosols on
precipitation processes associated with MCSs and thunderstorms. These results also show the complexity of aerosol-cloud-precipitation interactions within deep convection.

Khain et al. [2008] and Khain [2009] recently conducted a number of numerical experiments to identify the factors that determine the impact of aerosols on precipitation. A scheme (Figure 31) was proposed to classify aerosol effects on clouds and cloud systems under different environments. It shows that if the ambient relative humidity is high, then the condensation gain (loss) is large and the condensation loss (gain) is low, which could lead to an increase (decrease) in precipitation. Seifert and Beheng [2006], Khan et al. [2008] and van den Heever et al. [2011] have also all showed that the impact of aerosols on precipitation and convection structure depends on cloud type. A decrease in precipitation with an increase in aerosol concentration usually occurs for isolated cumulus clouds and cloud systems that develop within a relatively dry environment and/or within regions of large wind shear or stratocumulus clouds. Fan et al. [2009] also found that the increase in CCN concentration always suppresses convection under strong wind shear conditions but enhances convection under weak wind shear conditions. On the other hand, an increase in precipitation with an increase in aerosol concentration often occurs for clouds forming in a moist environment, such as coastal zones, within cloud ensembles or tropical squall lines. Such a result is also found in other CRM studies [Wang, 2005; Tao et al., 2007 and see Figure 30b; Fan et al., 2007b]. Note that Wang et al. [2005], Tao et al. [2007], Khain et al. [2008], Fan et al. [2007a] and Lee [2011] all showed that aerosols have a larger effect on precipitation under more humid conditions.

Recently, van den Heever et al. [2011] made use of large-scale (10,000km), long-duration (100 days), high resolution (1km) radiative-cloud-equilibrium (RCE) simulations to investigate the impacts of enhanced CCN concentrations on the wide variety of storm types that developed under a wide range of different environmental conditions. They found that high CCN concentrations led to a reduction in surface precipitation from shallow clouds, but an enhancement in the precipitation produced by deep convective clouds, with a mixed response from more moderate convective storms such as congestus. They also noted that the frequency of lighter-precipitation producing systems was decreased under more polluted conditions, while the frequency of heavy precipitation producing systems was enhanced, concluding that polluted conditions result in a greater frequency of more intense rainfall producing systems (Figure 32)
In almost all of the above cases, idealized aerosol concentrations were used in model simulations. Some of the CRM domains were too small to resolve the observed clouds or precipitation systems (the domain size has to be at least twice as large as the simulated features). Furthermore, very few of these CRM studies compared model results with observed cloud structures, organization, and radar reflectivities (Figure 33). Note that model simulations captured quite well storm sizes and structures under different environmental conditions. For example, the leading convection and the extensive trailing stratiform rain area compared well with the radar reflectivity observed during the mature stage of the continental case [Rutledge et al., 1988]. Clean cases (i.e., the control experiments) generally agree better with observations. In terms of radar reflectivity magnitudes, the agreement between simulations and observations is better at lower levels where only liquid phase cloud or rainwater exist. The simulated radar reflectivity tends to be higher at the upper levels and in the anvil area where ice phase particles dominate. This reflects the inadequate description of various ice-phase mechanisms in this calculation, particularly those related to aerosol effects.

5.3.2. Physical Processes behind the CCN Effect on Precipitation

Observational studies suggest that enhanced aerosol concentrations could suppress warm rain processes by producing a narrow drop size spectrum that inhibits collision and coalescence processes [e.g., Squires and Twomey, 1961; Warner and Twomey, 1967; Warner, 1968; Rosenfeld, 1999]. In addition, more aerosols would reduce precipitation and rainfall. This is because more aerosols could produce smaller cloud droplets that results in less efficient collision-coalescence and consequently less rainfall. Many cloud-resolving modeling studies simulated these processes during the early stage of convective systems and for isolated cloud and found that sometimes, more rainfall was simulated under polluted (high CCN concentration) conditions.

Several cloud-resolving modeling studies also examined the physical processes that could lead to aerosol-induced changes in precipitation. In general, three mechanisms were proposed to explain the enhancement of precipitation by changing (increasing or decreasing) the aerosol concentration. The first mechanism is the creation of stronger updrafts/downdrafts resulting from enhanced latent heat release as CCN suppress warm-rain formation and thus retain more liquid water in the cloud to be frozen at upper levels. This effect could be termed as the latent heat – dynamic effect (see Figure 34). Wang [2005] indicated that precipitation increases in tropical deep convection due to higher CCN concentrations through this latent heat effect. Khain et al. [2005] also found that for cases where there was enhanced precipitation with high CCN concentrations, clouds were associated with stronger updrafts/downdrafts as well as stronger
convergence in the boundary layer, which provides a better chance to trigger secondary clouds and prolong the lifetime of the convective system. van den Heever and Cotton [2007], Lee et al. [2008] and Storer et al. [2010] also demonstrated the influences of aerosols on secondary storm development and its effect on increasing precipitation.

The second mechanism is stronger evaporative cooling due to more but smaller raindrops under high CCN concentration conditions. Stronger evaporative cooling could enhance the strength of the near surface cold pool. When the enhanced cold pool interacts with lower level wind shear, convergence could become stronger, producing more vigorous convection that ultimately leads to enhanced surface precipitation. This positive feedback mechanism could be termed as the cool pool effect. It seems to be occurring in the oceanic convective case (Figure 35), in which evaporative cooling in the lower troposphere is more than twice as strong for the polluted scenario compared to the clean scenario. Note that stronger evaporative cooling occurs in the developing stage of convective cloud systems. Lee et al. [2009a] also demonstrated that stronger evaporative cooling occurred under high aerosol concentrations and consequently lead to enhanced surface precipitation.

The third mechanism is the CCN effect on cloud and microphysics. Wang [2005] suggested that with higher CCN concentrations, there is a greater increase in total water content consisting of numerous small liquid particles, leading to more vigorous convection and cold rain processes. As a result, rain production is mainly from ice phase microphysics in the modeled tropical deep convective case for which precipitation is increased due to more CCN (similar processes as the above continental convective case). Cheng et al. [2010] further showed that increasing CCN concentration leads to more cloud drops and cloud ice, but the uncertainty of its effect on surface precipitation comes from various responses of cold rain production (this could be termed as the cold-microphysics effect; see Figure 36). The CCN effect on precipitation in the low-CCN case is dominated by suppressed warm-rain formation (Effect-A) and enhanced riming (Effect-C). On the other hand, in the high-CCN case where changes in snow-rain drop accretion are small, decreases in snow riming (due to reduced collection efficiency, Effect-B) and melting are big enough to dominate the change in cold-rain production, suppressing surface rainfall due to the change in CCN number. The CCN effect on precipitation in the high CCN case is dominated by Effect-A and Effect-B.
However, these three physical mechanisms can affect each other, making it difficult to isolate each mechanism independently. For example, a stronger evaporative cooling (cold pool effect) in the low troposphere for the high CCN-enhanced precipitation case \( (\text{latent heat} – \text{dynamic effect}) \) is evident in Figure 34(a). On the other hand, the stronger cooling could be due to more evaporation associated with a stronger convective downdraft \( (\text{latent heat} – \text{dynamic effect}) \). The microphysics effect can affect both latent heat and the cool pool and vice-versa. Smaller but more cloud drops may enhance the WBF process (i.e., snow deposition growth at the expense of evaporating cloud drops) such that the latent heat release becomes faster. But the reduced riming efficiency has an opposite effect on latent heat release. More numerous cloud drops (due to more CCN) could also lead to more ice nucleation and thus a higher number of snow and graupel particles which, when melted, form more but smaller raindrops that cause stronger evaporative cooling below the cloud base. Convection may be strengthened due to either the latent heat effect or the cold pool effect, and the stronger updraft leads to higher supersaturation and thus enhanced ice deposition nucleation.

5.4. CCN Effect on Convective Precipitation Using Nested Cloud/Regional-Scale Models

Regional-scale models [mainly fifth-generation Pennsylvania State University-National Center for Atmospheric Research (Penn State-NCAR) Mesoscale Model (MM5), Weather Research and Forecasting Model (WRF) and Regional Atmospheric Modeling System (RAMS)] with fine resolution (utilizing interactive nesting techniques) have also been used to study the impact of aerosols on convective precipitation events associated with the Florida sea breeze, the urban heat island effect, tropical cyclones/hurricane, super cells and tornadoes (i.e., \textit{van den Heever et al.}, 2006; \textit{van den Heever and Cotton}, 2007; \textit{Lynn et al.}, 2005a, b; \textit{Zhang et al.}, 2007, 2009; \textit{Li et al.}, 2008a, b; \textit{Lerach et al.}, 2008; \textit{Khain et al.}, 2010; \textit{Storer et al.} 2010). An advantage of this approach is that the model initial and lateral-boundary conditions are provided by large-scale analyses with realistic meteorological fields, and model simulations can be conducted with realistic terrain and land-surface characteristics.

\textit{Lynn et al.} [2005a, b] used spectral-bin microphysics (a simplified version of Khain’s scheme \cite{Khain2004}) and the MM5 to simulate a cloud that approached the west coast of Florida, prior to the sea-breeze development. The use of a continental CCN concentration led to a delay in the growth of rainfall (in agreement with those shown in Section 6.3). The increase in CCN concentration led to convective invigoration and the formation of stronger secondary clouds \cite{Lynn2005b}. Simulations of rain events over the whole peninsula for this day showed significant invigoration of squall lines. There was an increase in precipitation rate and
precipitation amount for a squall line that formed in the vicinity of the east coast of Florida. At the same time, continental CCN concentrations resulted in a 5% reduction in precipitation over the whole computational domain (containing a significant fraction of Florida) versus maritime values.

van den Heever et al. [2006] used the RAMS and a two-moment bulk microphysical scheme [Meyers et al., 1997; Saleeby and Cotton, 2003] to examine the aerosol effect on the formation of a thunderstorm over the peninsula of Florida. Note that the two-moment bulk scheme used in van den Heever et al. [2006] emulates a bin scheme by including explicit activation of aerosols [Saleeby and Cotton, 2004]. Sensitivity tests show that different combinations of CCN, GCCN, and IN result in different amounts and temporal patterns of cloud-water/ice contents and rainfall. Their study showed that a high CCN reduces cumulative precipitation by 22% compared to low CCN. In addition, high amounts of GCCN and IN enhanced surface precipitation for the first 6 hours of integration due to the initial broadening of the cloud droplet spectra. However, the total (12-hour integration) accumulated precipitation was greatest for the clean (low CCN, GCCN, and IN) case. Rapid wet deposition of GCCN during the first 6 hours of integration may be the reason for this. All their experiments involving high CCN resulted in high cloud-water content and weak surface precipitation.

Using a similar modeling configuration, van den Heever and Cotton [2007] examined the sensitivity of urban-induced convective clouds over and downwind of St. Louis, MO. Their results indicate that downwind convergence (dynamic processes) induced by urban land cover appears to be the dominant factor in determining whether or not moist convection actually develops downwind of St. Louis. Once moist convection is initiated, urban-enhanced aerosols play a major role in determining the microphysical and dynamical characteristics of convective storms, particularly when background aerosol concentrations are low (Figure 37). Complicated relationships and feedbacks between microphysical and dynamical processes obscure the general understanding of urban-enhanced aerosol effects on precipitation. Note that Lynn et al. [2005b], van den Heever et al. [2006] and van den Heever and Cotton [2007] explicitly represent mesoscale forcing (i.e., sea-breeze convergence and urban heat island convergence). This is important because cold pools can interact with these kinds of circulation, introducing another level of dynamic complexity. For example, if the cold pool outruns these mesoscale convergence fields, precipitation is reduced whereas when they remain coupled, precipitation is
enhanced (cool pool effect in section 5.3.2). Van den Heever and Cotton [2007] also found that the response of convective rainfall to urban-enhanced aerosols becomes stronger when the background aerosol concentrations are low.

The role that aerosols or dust in the SAL play on tropical storms has recently been studied using regional-scale models [i.e., Zhang et al., 2007, 2009; Cotton et al., 2007]. The impact of dust acting as CCN in the SAL on the evolution of a tropical cyclone (TC) was examined through simulations initialized with an idealized pre-TC mesoscale convective vortex (MCV) using the RAMS. Dust in the SAL can affect the simulated TC intensity by 22 hPa depending on CCN concentrations (Figure 38). High CCN concentrations could weaken the TC intensity. It can also affect eyewall development directly through release of latent heat when activated and subsequent growth of cloud droplets and indirectly through modulating rain band development. Convection in the rain bands was negatively correlated with that in the eyewall in all simulations. The development of rain bands promotes latent heat release away from the eyewall, blocks the surface inflow and enhances cold pools. The convection in the eyewall and rainbands did not show a monotonic relationship to the input CCN due to the non-linear feedback of heating from a myriad of microphysical processes on storm dynamics (Figure 39). The impact of CCN on storm intensity was sensitive to the background GCCN vertical profile and presumably other environmental factors.

The WRF with spectral bin microphysics scheme was used to investigate the potential impact of aerosols on the structure and intensity of Hurricane Katrina as they were ingested into the storm’s circulation during its passage through the Gulf of Mexico in 2005 [Khain et al., 2010]. Continental aerosols invigorated convection largely at the TC periphery, which led to its weakening prior to landfall. The minimum pressure increased by 15 hPa, and the maximum velocity decreased up to 15 m s⁻¹. Aerosols substantially affected the spatial distribution of cloudiness and hydrometeor contents. The results are in good agreement with Zhang et al. [2009]. Figure 40 shows how the TC intensity is weakened due to the aerosol effect. Three processes are responsible for the decrease in intensity: (1) an increase in cloud velocity at the TC periphery, causing mass updraft increases at the periphery with a smaller amount of air mass and water vapor penetrating to the central part of the TC, (2) extra convective heating at the periphery lowering the surface pressure at the TC periphery and decreasing the horizontal pressure gradient, and (3) competition between two zones of convection (at the eye and at the
periphery). Compensatory downdrafts caused by convection at the TC periphery also tend to damp convection in the TC eye. The diagram shown in Figure 40 agrees well with results of Saharan dust effects on TC intensity reported by Zhang et al. [2009]. They found that convection in TC rainbands was negatively correlated with that in the eyewall in all simulations.

Li et al. [2008a] implemented a two-moment bulk microphysical scheme into the WRF model to investigate the effects of aerosols on cloud and precipitation processes. A deep cumulus cloud and precipitation event was simulated and a sensitivity study was carried out using a set of initial aerosol profiles at the surface level to examine the response of precipitation efficiency to the increase in aerosol concentration. Precipitation increased with aerosol concentration from clean to polluted conditions, but was considerably reduced and completely suppressed under extremely polluted conditions (Figure 40). Enhanced precipitation with increasing aerosols at lower CCN is attributable to the suppressed conversion of cloud droplets to raindrops, which causes less efficient warm rain processes but more efficient mixed phase processes. At extremely high CCN, ice production is inhibited because ice nucleation becomes inefficient and anvil formation is hindered due to a large mass loading of small droplets (reduced buoyancy) and less latent heat from droplet freezing. Cloud coverage and core updraft exhibited similar non-monotonic behaviors to precipitation under different aerosol loadings (Figure 40(a) and (d)). Under a similar modeling framework, aerosol effects on precipitation for different regional scale systems were further investigated, including a mesoscale squall line in the southern plains of the United States [Li et al., 2009] and a wintertime storm over the North Pacific [Zhang et al., 2007; Li et al., 2008b]. In these scenarios, precipitation enhancement under polluted conditions was reported along with more efficient mixed phase processes and intensified convection.

Lerach et al. [2008] used the RAMS to study the impact of aerosol indirect effects on the development of a supercell storm and its associated tornadogenesis. Their results indicated that a polluted environment (dust-laden CCN \((2000 \text{ cm}^{-3})\) and GCCN \((0.2 \text{ cm}^{-3})\) concentrations) would provide a favorable environment for tornadogenesis rather than a clean environment (clean continental CCN and GCCN concentrations). Their results indicated that both warm rain and cold rain are reduced in the polluted environment, and a longer-lived supercell is simulated. An EF-1 (surface wind speeds exceeding \(40 \text{ m s}^{-1}\)) tornado was produced due to weak evaporative cooling near updraft and downdraft interfaces close to the rear flank of the supercell. Storer et al. [2010] also made use of RAMS to simulate supercell development under clean and more
polluted conditions. The aim of their study was to assess the relative importance of variations in convective available potential energy (CAPE) and aerosols concentration on the microphysical and dynamical characteristics of supercells (Figure 41). Storms in higher CAPE environments contained more cloud and ice water mass, generated more surface precipitation and had stronger cold pools when compared to those developing in lower CAPE environments. Polluted storms produced higher concentrations of smaller cloud droplets, increased cloud and ice water, and reduced surface precipitation when compared to those storms developing in cleaner conditions. There were also fewer, but larger raindrops in the polluted storms, which, together with the reduced precipitation amounts led to weaker, warmer cold pools. The changes in the cold pool characteristics led to subsequent changes in the secondary convection. When comparing the results from the changes in CAPE and aerosol concentrations they found that the total precipitation produced by storms is primarily driven by CAPE but that the increase in aerosol concentrations produced a decrease in total precipitation by 30-40%. Changes in the cloud water path are found to be much more sensitive to variations in aerosol concentrations than in CAPE, while changes in the ice water path are affected similarly by changes in aerosol and CAPE. Other microphysical parameters such as the mean raindrop diameter demonstrated very little response to changes in CAPE but large responses to changes in aerosol. Finally, they concluded that a number of the aerosol indirect effects observed in their study were modulated by the amount of CAPE in the environment, with the aerosols effects typically being stronger in environments with weaker CAPE. Thus, both aerosol concentrations and environmental conditions need to be considered when assessing aerosol indirect forcing on deep convective storms.

5.5. IN Effect on Precipitation Processes

Only a few CRMs have been used to study the effects of IN on precipitation processes because of the limited understanding of ice formation (see Section 2). In addition, some aerosols can be transported by convective updrafts from the planetary boundary layer (PBL) to the middle and upper troposphere and can be served as IN [Yin et al., 2005]. It would be very difficult to quantify the IN effect on precipitation processes. In addition, most of the previous CRMs were used to examine both IN and CCN effects on precipitation. van den Heever et al. [2006] showed that increasing either the GCCN or IN
concentrations produces the most rainfall at the surface whereas enhanced CCN concentrations reduce surface rainfall. Higher IN concentrations produce ice at warmer temperatures and generate deeper anvils, but simultaneously increasing the concentrations of CCN and GCCN leads to more supercooled liquid water available for freezing and greater ice mixing ratios. Higher concentrations of GCCN and IN result in greater accumulated surface precipitation initially. By the end of the simulation period, however, the accumulated precipitation is the greatest for the case in which the aerosol concentrations are lowest. Their results suggested that such changes in the dynamical and microphysical characteristics of convective storms as a result of the variations in aerosol concentrations have potential climate consequences, both through the cloud radiative effect and the hydrological cycle.

Ekman et al. [2007] studied the sensitivity of a continental storm to changes in IN concentration based on an observed case. They found that the increase in IN concentration, and thus heterogeneous nucleation, would generally result in enhanced updrafts due to latent heat release from the added diffusive growth of increased ice crystals. Such an effect was also identified to enhance homogeneous nucleation, i.e., to make CCN more effective in influencing ice nucleation. Since the dominant mechanism in providing ice particles for cirrus anvils in the modeled case is still homogeneous nucleation, the increase in IN concentration would lead to enlarged anvil coverage and increased precipitation. The finding of Ekman et al. [2007] suggests an interesting link between IN concentration and updraft strength, and consequently, homogeneous nucleation rate, anvil particle concentration, coverage, and radiation effects. It is also in qualitative agreement with observations by Heymsfield et al. [2005].

Fan et al. [2010b] studied the impact of different parameterization schemes on homogeneous nucleation and heterogeneous immersion nucleation for two isolated tropical deep convection cases. Similar to the finding of Ekman et al. [2007], they found that an increase in the immersion nucleation rate would lead to stronger convection, larger anvil coverage, and longer anvil lifetime. Precipitation and ice water path would both be enhanced in either wet or dry environmental conditions. Consequently, the homogeneous nucleation rate would also be enhanced, consistent with Ekman et al. [2007] and Heymsfield et al. [2005]. Note that Ekman et al. [2007] and Fan et al. [2010b] mentioned the substantial uncertainty in modeled results due to various different schemes used to represent nucleation mechanisms, although such uncertainty does not seem to qualitatively affect their major conclusions. Fan et al. [2010a] also compared
CCN and IN effects on convection and precipitation, and found that the CCN effect is dominant. The IN effect does little to convection and precipitation, although ice microphysical properties could change significantly.

Zeng et al. [2009a, b] examined the effect of the ice crystal enhancement (IE) factor (defined as the ratio of the number of ice crystals to ice nuclei) on ice properties associated with tropical and mid-latitude convective systems using a CRM. Their results showed that the IE factor in tropical clouds is about a thousand times larger than that in mid-latitude clouds. This significant difference in the IE between the Tropics and mid-latitudes is consistent with observations of stronger entrainment and detrainment in the Tropics. However, surface precipitation was not affected by the change in IE as shown in Fan et al. [2010a].

6. Large-Scale Modeling of Aerosol-Precipitation Effects

On a global scale, aerosols could affect precipitation through microphysical and dynamical paths. The former path consists of a series of microphysical processes and microphysics-dynamics feedbacks, initiated by the activation of aerosols as CCN or IN. The resultant change in cloud radiation is commonly referred to as the indirect radiative forcing of aerosols [e.g., Ramaswamy et al., 2001]. The dynamical path is implemented through changes in large-scale circulation by persistent aerosol direct forcing [e.g., Wang, 2004] or indirect forcing [Ramaswamy and Chen, 1997; Rotstayn and Lohmann, 2002]. The microphysical path would cause precipitation changes mainly confined to aerosol-laden regions while the dynamical path could alter precipitation remote to these regions.

To simulate the effects of aerosols on precipitation in a large-scale framework, models need to include the feedback of dynamical processes to aerosol forcing. The atmospheric transport model or chemical transport model, a commonly used tool to derive aerosol distribution and radiative forcing driven by prescribed meteorology, would not be useful for this purpose because of its lack of such feedbacks, and for this reason will not be discussed in this paper. In addition, recent advances in computational technology allow current CRMs to run efficiently over domains covering a continental scale. In fact, many details about regional modeling can be found in the previous section. This section, therefore, will concentrate on the efforts using global aerosol-climate models. The term aerosol-climate model will be used in this section and refers to
those general circulation models (GCMs), or global climate models as they are sometimes called, those include an aerosol module or simply include prescribed radiative effects of aerosols.

6.1. Major Methodology

Precipitation is formed through a series of sophisticated processes with feedbacks involving aerosol physics and chemistry, cloud microphysics, and dynamics on different scales. It is located at the bottom of the “food-chain” of aerosol-cloud processes, while serving as the largest sink of atmospheric aerosols. Many of these above processes occur on scales much smaller than the typical grid spacing of current global climate models. The first step toward modeling the effect of aerosols on precipitation using global climate models is thus to have a reasonable representation of aerosol processes in these models.

The study of climate response to aerosol forcing requires long-term integrations using models on one hand equipped with needed details concerning aerosol processes, and on the other hand, having a good computational efficiency. In reality, mainly because of the computational cost, it has never been an easy task to include an interactive aerosol module in climate models to study the change in clouds and precipitation as a climate response to aerosol forcings. Therefore, an optimized balance between aerosol representation and model computational cost needs to be established. So far, the majority of research on modeling the climate impacts of aerosols has been conducted using models with prescribed aerosol profiles. This applies perhaps to all the GCMs that participated in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR4) [Randall et al., 2007].

The evolution of atmospheric aerosols can be represented by several important aerosol properties, such as size distribution, mixing state, and chemical composition. The size of an aerosol with a known chemical composition to a large extent determines both the optical and microphysical character of the particle. In many aerosol-climate models, particularly the earlier ones, only the aerosol mass mixing ratio was predicted. The size of aerosols in these models, when needed, was derived diagnostically using various empirical methods based on the predicted mass mixing ratio. In many such models, the radiative properties of aerosols were derived based on a given geometric size [Haywood and Boucher, 2000; Penner et al., 2001; Forster et al., 2007]. On the other hand, the mixing state of aerosols determines not only the value but also the
sign of the radiative forcing of aerosols. Recent field experiments using single particle analysis have demonstrated the coexistence of external and internal mixtures in the atmosphere and the latter is not always dominant [Hara et al., 2003; Schwarz et al., 2008; Twohy et al., 2008; also see Section 3]. From field experiments, internal mixtures containing an absorbing core (e.g., black carbon, or BC) coated by a scattering material [Martins et al., 1998; Posfai et al., 1999; Naoe and Okada, 2001; Okada et al., 2005; Schwarz et al., 2008] are known to significantly enhance the bulk absorption per mass of aerosols compared to external mixtures of black carbon [Ackerman and Toon, 1981; Horvath, 1993; Chylek et al., 1995; Bond and Bergstrom, 2006; Kim et al., 2008]. This specifically differs from the relatively well-mixed state of organic carbon (OC) and sulfate mixtures [e.g., Okada et al., 2001; Russell et al., 2002]. The majority of current aerosol-climate models include only external or internal mixtures. Among those models that include internal mixtures, a well-mixed assumption is usually adopted so that the radiative effect of aerosols are calculated using a volume-weighted refractive index over all predicted aerosol constituents [e.g., Stier et al., 2005; Kim et al., 2008].

Increasing effort has been made in recent years to improve the representation of aerosols in GCMs (Figure 43). In a few studies where an aerosol model was used to predict number and mass concentrations of aerosol constituents in a series of relatively coarsely defined size bins, or sections (e.g., Jacobson [2001]; Adams and Seinfeld [2002]), the computational cost was still too high to include multiple section series based on the physicochemical mixing of aerosols in multiyear simulations. To reach a good balance between the computational burden and process details in climate impact studies, some of the current aerosol-climate models have adopted the multi-moment approach [e.g., Easter et al., 2004; Vignati et al., 2004; Liu et al., 2005; Kim et al., 2008; Bauer et al., 2008]. This method predicts many aerosol population properties including the number, mass, and size distribution (whether based on a prescribed type of function or not) of many aerosol modes (see Table 5 for a brief description of several selected models and Figure 44 for the model diagram of Kim et al. [2008]). Compared to the earlier mass-only model, or the ‘single-moment’ scheme, the multi-moment approach, when designed appropriately, improves the representation of aerosol size while remaining computationally efficient with the ability to perform multi-decadal or even longer simulations.

6.2. Impact of Aerosols on Precipitation through Microphysical Processes
In global climate models, the representation of the impact of aerosols on cloud microphysics is generally implemented in a highly parameterized fashion due to the coarse spatial resolution of these models. Such a parameterization usually consists of two major steps: (1) the activation/nucleation of cloud droplets or ice crystals and, (2) the “auto-conversion” to convert small cloud droplets or ice crystals (typically from 10 to 100 µm in size) to precipitating hydrometeors, i.e., rain or snow (several hundreds of micrometers to a few millimeters). Note that the latter transition is a parameterization of a series of microphysical processes involving a random collection of small cloud droplets under the influence of up- or downdrafts and turbulent mixing at a subgrid-scale level.

In fact, the majority of global aerosol-climate modeling efforts dealing with the microphysical impact of aerosols has focused on deriving the indirect radiative forcing of aerosols rather than studying the precipitation effect. The former task, in a model without explicitly predicting aerosol number concentration, would be done by using an empirical method to connect aerosol mass with CDNC, auto-conversion rate, and precipitation efficiency to simulate various aerosol effects [Rotstayn, 2000; Kiehl et al., 2000; Menon et al., 2002; Ming et al., 2005; Penner et al., 2006; Jones et al., 2007]. The aerosol-CDNC relationships used by different aerosol-climate models appear to be consistent with each other. The most recent results concerning the aerosol albedo effect derived by models are also close to satellite-based estimations. However, in performing more sophisticated tasks, such as simulating aerosol effects on cloud cover, height, lifetime, and outgoing longwave radiation which involve dynamic feedbacks, these models tend to disagree with each other [Quaas et al., 2009]. To date, the argument about whether the indirect forcing of aerosols should be treated as a forcing or a response still exists, and the estimation of such an effect remains one of the largest uncertainties in climate modeling. For more details about studies of indirect forcing of aerosols, the reader is referred to previous reviews including Haywood and Boucher [2000], Ramaswamy et al. [2001], and Lohmann and Feichter [2005].

Several physics-based parameterizations for simulating the activation of liquid or ice cloud particles have been developed for global models [e.g., Abdul-Razzal and Ghan, 2000; Nenes and Seinfeld, 2003; Kärcher and Lohmann, 2003; Liu and Penner, 2005; Barahona and Nenes, 2007]. These schemes continue to benefit even the most up-to-date aerosol-climate
models with improved aerosol and cloud microphysical representations. Effort was also made to include more physically-based descriptions of cloud microphysical processes beyond aerosol activation in the earlier generation of models. For example, Rasch and Kristjansson [1998] and Kristjansson [2002] introduced parameterizations of several microphysical processes into the Community Climate Model version 3 (CCM3) of the National Center of Atmospheric Research (NCAR), including auto-conversion from cloud water to rain and from ice to snow and the collection of cloud water by rain and by snow. This implementation would enable the model to simulate certain effects of aerosols on cloud droplet size and concentration and, most interestingly, on precipitation onset and quantity.

As a major step forward, a two-moment microphysics scheme for cloud droplets and ice crystals has been developed recently for global climate models [e.g., Lohmann et al., 1999, 2007; Ming et al., 2007; Morrison and Gettelman, 2008; see Table 6]. The evaluation of these relatively comprehensive models is just beginning [Gettelman et al., 2008; Lohmann, 2008]. This type of scheme requires that the model predict both mass and number concentration of a given hydrometeor, i.e., the type of cloud particles categorized using size, phase, aggregate state, and often density. Based on the selected type of probability distribution function (PDF) for the size distribution of a given hydrometeor, and when the predicted number of moments equals the number of undefined parameters of the PDF, all microphysical conversions, along with the bulk motion characteristics of the hydrometeor, can be derived using integrations over size distributions that define various predicted moments. The effort to introduce the two-moment scheme for convective clouds in GCMs has just started. Lohmann [2008] tested a two-moment scheme for convective clouds in the ECHAM general circulation model. Song and Zhang [2011] tested a two-moment scheme including four hydrometeors for convective clouds using a single column climate model. Obviously, such a task is still very challenging due to the sub-grid nature of convection in current climate models.

The attempt to use a two-moment cloud microphysics scheme in GCMs is still in its early stage and current schemes are still less comprehensive than their counterparts developed long ago for cloud-resolving models [e.g., Wang and Chang, 1993]. Nevertheless, by adopting such an approach, modeling aerosol effects on precipitation would become much more physics-based than in the earlier generation of models where only mass mixing ratios of hydrometeors were predicted. Especially when coupled with a size-dependent aerosol module, a two-moment cloud
microphysics scheme provides a connection between the predicted number concentration of CCN by the aerosol module and predicted CDNC by the cloud microphysics module. Along with the two-moment scheme, subgrid-scale variability has also been introduced to further improve the modeled microphysical features in *Morrison and Gettelman* [2008], where such variability of cloud water is represented by a PDF derived from observations and extended to all related microphysical conversions.

Studies have demonstrated that with the two-moment scheme, models tend to produce global total precipitation closer to observations than the previous generation of schemes [e.g., *Lohmann et al.*, 2007; *Gettelman et al.*, 2008]. On the other hand, in representing the formation of raindrops, a process of random nature involving collection among smaller cloud droplets, models with the two-moment scheme still need to include the parameterization of auto-conversion with explicit use of both predicted number concentration and mass mixing ratio of cloud droplets. Such a parameterization thus remains a major tuning job for simulating the effect of aerosols on precipitation. A recent sensitivity study demonstrated that by allowing raindrops to form directly from the activation of GCCN (a hypothesized path), the modeled hydrological cycle would become faster, compared to the case without such a path; this was manifested by faster rainfall and lower atmospheric water vapor content, although the total precipitation amount did not change much [*Posselt and Lohmann*, 2008; also see Figure 45]. Early attempts to bring the two-moment cloud microphysical scheme into the convection parameterization of GCMs have also identified the impact of aerosols on modeled precipitation features. *Lohmann* [2008] found that using the two-moment scheme and incorporating prescribed aerosol emissions could improve the modeled geographical distribution of precipitation changes from the past to the present climate when compared to observations. *Song and Zhang* [2011] found that the two-moment scheme alters the precipitation partition between convective and stratiform portions of modeled clouds, actually leading toward a ratio that is closer to observations than results derived using a simpler microphysical scheme. Both studies have identified a suppression of convective precipitation due to enhanced aerosol concentrations.

Despite the advancement in introducing physics-based aerosol and cloud schemes into GCMs, verification of modeled precipitation changes due to aerosol microphysical effects remains a challenge. This issue inherits the problem in determining the reference state for the pre-industrial climatology of CDNC, a result that is much-needed for estimating the
anthropogenic enhancement to cloud radiative forcing in present days [e.g., Rotstayn and Penner, 2001]. For instance, models often introduce a lower bound to avoid low values of CDNC. This results in an areally uniform distribution of CDNC, which leads to a substantial underestimation of the aerosol effect on cloud albedo [Hoose et al., 2009]. In addition, more advanced schemes for aerosol and cloud microphysics also introduce new parameters. This would potentially add to the difficulty in examining whether the aerosol-cloud-precipitation microphysical effect would be able to explain some known variability in precipitation.

6.3. Impact of Aerosols on Precipitation through Coupling with Large-Scale Dynamics

The direct radiative forcing (DRF) of anthropogenic aerosols involves scattering and absorbing of solar radiation, or absorption/emission in the longwave radiation spectrum. The latter is only effective for large particles such as dust and will not be discussed much here. The optical effects of aerosols exert a negative forcing (cooling) at the Earth’s surface by both scattering and absorbing solar radiation, and a positive forcing (warming) to the atmosphere by absorption. Due to their short lifetime of about one week, aerosols created from anthropogenic activities tend to concentrate over source regions or surrounding areas and hence apply a persistent forcing in these places.

This type of persistent forcing can perturb atmospheric thermodynamic and dynamical processes, and therefore, influence precipitation-producing cloud systems in regions away from aerosol-laden areas. For example, studies using different GCMs all indicate that direct radiative forcing of absorbing black carbon aerosols can lead to a northward shift of precipitation in the ITCZ, specifically over the Pacific Ocean [Wang, 2004; Roberts and Jones, 2004; Chung and Seinfeld, 2005]. Ramaswamy and Chen [1997] and Rotstyn and Lohmann [2002] both found that by including aerosol indirect forcing in their models, the dynamical consequence responding to such a forcing could have a remote impact on ITCZ precipitation, i.e., a shift in an opposite direction to that caused by BC. An active research field aimed at understanding the effects of aerosols on precipitation through coupling with large-scale dynamics has developed. Current results indicate that the precipitation change caused by aerosols through this optical-thermodynamic-dynamical linkage is likely to be significant and could also be nonlinearly related to the aerosol forcing strength [e.g., Wang, 2007, 2009].
Studies of the impact of aerosols on monsoon systems are rapidly growing. Correlations between estimated precipitation and circulation changes with the increasing trend of aerosols have unquestionably fueled research in this direction [Ramanathan et al., 2005; see also Figure 46]. Most of the studies are conducted using 3-dimensional atmospheric GCMs or atmosphere-ocean coupled global circulation models (CGCM), which include either an interactive aerosol module or prescribed aerosol profiles. Paired simulations, which include and exclude aerosol effects, or include reference and altered aerosol profiles, allow comparisons to be made of the climate response to different aerosol forcing assumptions. Aerosol effects would be isolated, barring the assumption that model artifacts in simulating the monsoon system would not be significantly amplified by using different aerosol profiles. Model representations of aerosols and aerosol-climate interactions vary from study to study.

The impact of aerosols, particularly absorbing aerosols, on the Indian summer monsoon has been extensively studied using models [Ramanathan et al., 2001; Chung et al., 2002; Ramanathan et al., 2005; Lau et al., 2006, 2008; Meehl et al., 2008; Randles and Ramaswamy, 2008; Wang et al., 2009b; Collier and Zhang, 2009; Krishnamurty et al., 2009; Manoj et al., 2010]. Analyses looking at total precipitation over India mostly indicate an aerosol-caused decrease in rainfall during the summer monsoon season and an increase during pre-monsoon and onset seasons. Several studies that analyzed precipitation patterns further suggested more complicated features, such as an increase in monsoon rainfall in certain regions, e.g., northwestern India. While qualitative and quantitative details may differ among them, these modeling studies all conclude that BC radiative forcing is large enough to perturb the Indian monsoon system and alter the precipitation pattern and amount in a statistically significant manner (see Table 7).

Various mechanisms, by which the direct radiative forcing of absorbing aerosols (solar heating of the atmosphere and dimming at the surface) modifies the monsoon circulation and rainfall, have been suggested in modeling studies. Wang et al. [2009b] indicated that absorbing aerosols, whether coexisting with scattering aerosols or not, can alter the meridional gradient of moist static energy in the sub-cloud layer over the Indian subcontinent, initiating heating of the air by absorbing aerosols in the PBL and thus changing large-scale atmospheric stability (Figure 47). This would cause a northward extension of the monsoonal circulation and convective precipitation, which coincidentally is in general agreement with observed monsoon precipitation
changes in recent decades, particularly during the onset season (Figure 48). Such a pattern change is also in agreement with results produced in another study using a coupled atmosphere-ocean model although a prescribed aerosol profile was used [Meehl et al., 2008; see Figure 49]. 

Lau and Kim [2006] proposed that the heating of absorbing aerosols on the slope of the Tibetan Plateau could initiate a positive feedback by drawing water convergence from oceans first and then form condensation and thus further heat the air over these elevated places; this is referred to as the “elevated heat pump” effect (EHP; see Figure 50). Ramanathan et al. [2005] suggested that aerosols could cause a decrease in surface evaporation (due to the dimming effect) and in the meridional sea surface temperature gradient. These two factors would affect monsoonal circulation and rainfall strength. The relative importance of these various mechanisms requires further research [Lau et al., 2008].

The impact of aerosols on another monsoon system, the West African Monsoon (WAM), has also been studied. Huang et al. [2009a] compared observational results with a global model simulation which included only direct radiative forcing of black carbon [Wang, 2004]. Both observations and model simulations showed that anomalously high African aerosol loading during boreal cold seasons were associated with significant reductions in cloud amount, cloud-top height, and surface precipitation. This result suggests that the observed precipitation reduction in the WAM region is caused by the radiative effect of absorbing BC. The mechanism for this reduction, however, remains a mystery. In connection to the hypothesis of the aerosol-Indian summer monsoon effect proposed by Wang et al. [2009b], Eltahir and Gong [1996] found a correlation between WAM strength and the subtropical meridional gradient of sub-cloud MSE. So a similar mechanism could also explain the aerosol-WAM effect. Recently, Lau et al. [2009] used a GCM to show that the EHP effect due to Saharan dust and biomass-burning BC has a significant impact on the climate and water cycle of the North Atlantic and the WAM. They found that during the boreal summer, as a result of large-scale atmospheric feedback triggered by absorbing aerosols, rainfall and cloudiness are enhanced over the West Africa/Eastern Atlantic ITCZ and suppressed over the West Atlantic and Caribbean regions. This is related to the elevated dust layer that warms the air over West Africa and the eastern Atlantic and thus causes the air to rise. The response reflects a strengthening of the WAM, manifested in a northward shift of West African precipitation over land, increased low-level westerly flow over West Africa at
the southern edge of the dust layer, and a near-surface westerly jet underneath the dust layer over the Sahara (Figure 51).

The anthropogenic aerosol level in the Southern Hemisphere is lower than in the Northern Hemisphere. Therefore, the impact of local aerosols on the Australian monsoon system is expected to be relatively insignificant. However, since monsoon systems are closely associated with large-scale circulation, aerosol effects in the Northern Hemisphere could influence southern hemispheric circulation and thus precipitation by altering general circulation patterns. Rotstayn et al. [2007] hypothesized that Asian aerosols could lead to an increase in both rainfall and cloudiness over northwestern Australia, which would coincide with the observed rainfall trend in the region since the 1950s. They suggested that this would happen because the meridional temperature gradient over the tropical Indian Ocean would be altered by Asian aerosols, which would further enhance monsoonal circulation toward Australia. A recent analysis of the twentieth century modeling results from 24 models participating in the Coupled Model Intercomparison Project Phase 3 (CMIP3), however, cannot provide support to the above hypothesis [Cai et al., 2010]. Despite the inclusion of either direct or both direct and indirect aerosol forcings in these models, their ensembles did not produce the hypothesized rainfall increase in northwestern Australia.

7. Current and Future Research on Aerosol – Precipitation Interactions

We have described the impact of aerosols on precipitation systems from perspectives of involved microphysical processes, observational evidence, and a range of numerical model simulations. We have also described the discrepancy between results simulated by models, as well as that between simulations and observations. Understanding these discrepancies is a necessary step in further resolving aerosol effects on cloud microphysics, dynamics and precipitation within climate systems.

From all discussions presented in Sections 3 to 6, it is obvious that to isolate the aerosol effect on precipitation from other factors is still a challenge and each of the current attempts to tackle this issue has its own advantages and disadvantages. Large-scale diagnostics of the aerosol effect on precipitation, through CCN, GCCN and IN, are incomplete without considering microphysics (cloud, aerosol, and precipitation), cloud dynamics (life cycle), and the large-scale
background environment (dynamical and thermodynamical forcing, radiative feedbacks, aerosol transport, and the water budget).

Nevertheless, progress has been made in cloud microphysics involving CCN, GCCN, and IN, and their impact on clouds and precipitation ranging from the cloud scale to larger scales. Numerical models have assisted the physical interpretation of mechanisms associated with the aerosol-precipitation interaction. It is still quite challenging to quantify various aerosol effects on precipitation from either modeling or observational studies alone [Khain et al., 2009; Stevens and Feingold, 2009; Koren et al., 2010b]. Combining observations and modeling can enable us to better understand the responses of cloud and precipitation to aerosols and their interactive feedbacks. The following summarizes the remaining scientific challenges and issues regarding aerosol-precipitation interactive processes.

7.1. Microphysics

The fundamentals of the aerosol effect on cloud and precipitation formation are based on the concepts of CN activation into cloud drops and IN nucleation into cloud ice. The former can be described by Köhler curves, presuming that information such as the size distribution and chemical activity coefficient is known. Such information can be derived from measurements, so that aerosol properties can be prescribed accordingly in cloud models. However, prescribing aerosol properties is only suitable for case studies and not for simulations with durations longer than the lifetime of aerosols which is about a few days. More elaborate simulations should be done by coupling the CRM with a detailed aerosol model that has the capability of resolving particle sizes, chemical compositions, and mixing state. The coupling must be done in such a way that the effects of cloud on aerosols, including aerosol scavenging and aerosol-recycling and even cloud chemistry, can be simulated simultaneously along with aerosol effects on cloud microphysical features.

IN nucleation is much less understood than CN activation. The main difficulty comes from the poorly quantified properties of IN including their concentration, variety, and ice nucleating capability. Measurement of these IN properties is still quite lacking, so even empirical descriptions of ice nucleation are tentative for most regions. A prognostic description of IN is possible and should also be incorporated into CRMs with a detailed aerosol model.
Another area that needs further improvement is the cloud microphysical sensitivity to aerosol effects. Many bulk water schemes crudely simplify the collision efficiency between cloud drops and precipitation particles, so the effects of aerosol on cloud drop size and thus collision growth processes cannot be fully resolved.

7.2. Cloud-Scale Modeling

In almost all previous cloud-resolving modeling studies, idealized or composite [e.g., van den Heever et al., 2006] CCN concentrations were used in the simulations. Furthermore, the spatial distribution of CCN was assumed to be uniform, at least horizontally. A non-homogeneous CCN distribution, consistent with non-homogeneous initial meteorological conditions, will be required to assess aerosol-precipitation interactions using regional-scale models in the future. In addition to IN and GCCN, the chemistry of CCN needs to be considered in future modeling of aerosol-precipitation interactions. For example, Fan et al. [2007b], Ekman et al. [2004, 2006] and Lee et al. [2009b] found that aerosol chemical composition and aerosol physics could affect precipitation processes.

Furthermore, almost none of these CRM studies compared model results with observed cloud features such as dynamical and microphysical profiles, cloud extent, evolution of organization, radar reflectivity, and rainfall. This instantly raises the issue of gathering more data, perhaps through major field campaigns, that are suitable to both initialize model simulations (with meteorological and aerosol parameters) and validate model results (i.e., in situ cloud property observations, radar, lidar, and microwave remote sensing). Even though CRM-simulated results have provided valuable quantitative estimates of the indirect effects of aerosols, one needs to realize that these CRMs can only simulate clouds and cloud systems over a relatively small domain, for short periods, and under constraints given by large-scale meteorological conditions. Some CRM domains are too small to resolve whole clouds or precipitation systems (the domain size has to be at least twice as large as the simulated features). Close collaborations between the global and CRM communities are needed in order to explore how to extend the CRM results to a regional and global perspective.

Identifying regime-specific sensitivities using satellite data, in situ observations, and high resolution, fine-scale modeling is a promising approach to improving the representation of clouds in larger-scale climate models [Stevens and Feingold, 2009].

7.3. Large-Scale Modeling
As a critical factor determining climate response and feedback, the role of aerosols in influencing precipitation is undoubtedly a key issue in climate research and prediction. The future direction of global-scale research regarding this should focus on improving the model representation of aerosol-cloud related processes and, perhaps more importantly, the observational constraints for aerosol-climate modeling efforts. An immediate consequence of model advancement in spatial resolution and process representations is the realization that observation data at high resolution is required for model initialization and verification. Clearly, all currently available measurement networks collecting observation data do not meet this requirement. How to provide the constraints for high-resolution global climate models thus needs to be studied.

With the rapid advancement in computational technology, climate models will step into an era of fast development. To better represent the effects of aerosols on clouds and precipitation, size-dependent cloud microphysics schemes along with size- and mixing-dependent aerosol schemes should be adopted in more models and used in long-term climate integrations. One issue that has remained in global climate modeling ever since its earliest days is the parameterization of convection. Because the requirement of a-few-kilometer resolution to resolve convection has been a far stretch for global climate models (and will likely still be the case in the near future), descriptions of convection processes along with aerosol-cloud-precipitation physics in these models have been empirically formulated using parameters resolved at model grid scales. Though still computationally expensive, today’s global cloud-resolving models [i.e., Tomita et al., 2005; Miura et al., 2005; Nasuno et al., 2008] are already being run in an exploratory manner and perhaps will reach the stage of performing long-term integrations with computationally more practical expenses earlier than one would expect. Recently Grabowski and Smolarkiewicz [1999], Khairoutdinov and Randall [2001], and Randall et al. [2003] proposed a multi-scale modeling framework (MMF) that replaces the conventional cloud parameterizations with a CRM in each grid column of a GCM. The MMF can explicitly simulate deep convection, cloudiness and cloud overlap, cloud-radiation interaction, surface fluxes, and surface hydrology at the resolution of a CRM. It also has global coverage and two-way interactions between the CRMs and their parent GCM. The MMF could be a natural extension of current cloud resolving modeling activities. MMFs can also bridge the gap between traditional CRM simulations and non-hydrostatic global cloud-resolving models. The MMF approaches that couple the global aerosol-climate model with a cloud-resolving model with
detailed interactions between clouds, precipitation, and aerosols, on the other hand, could be another solution of less computational demand for addressing this issue.

7.4. Observations

Observations provide both initial conditions and validation data for models. The current generation of cloud-resolving models include reasonable, albeit not complete, microphysical parameterizations, and can simulate the microphysical, dynamical, and chemical evolution during the life cycle of cloud systems. These models can also be used to explicitly calculate interactions between clouds, aerosols, and precipitation. However, all these advancements have raised challenging issues for corresponding measurements. Because of the range of scales involved in modern cloud modeling efforts, space-based remote sensing has served as an ever more necessary part of model validation.

The availability of high-resolution surface precipitation data, compiled from satellite retrievals and surface rain gauge measurements in the recent two decades, has generally improved, though far from resolved, the data shortage situation. Still, how to detect aerosol signals along with clouds and precipitation, specifically how to attribute the signal to microphysical path of aerosols, remains a rarely touched topic and requires well-designed field campaigns in future to address.

As more variables concerning aerosol-cloud-precipitation interactions are available from both satellite (e.g., A-train) sensors that provide global coverage and enhanced special ground-based observations (e.g., ARM) that provide a rich source of information, it is possible to investigate the climatological effects of aerosol-cloud-precipitation interactions beyond case-by-case studies, as demonstrated in Li et al. [2011] and Niu and Li [2011a, b]. If numerical models with adequate physics could reproduce such results obtained from very large ensembles of observations, it would provide solid evidence of the net ACP effect that is currently still missing but of considerable climatological significance.

Enlightened by the role of absorbing aerosols revealed in recent studies, measurements of the absorption strength and the spatial distribution of absorbing aerosols are still rare and should be specifically improved to validate modeling results [Wang et al., 2009a; Chin et al., 2009]. The most urgent need is to specify the vertical profile of aerosol absorption at least over highly
polluted regions, which is a key indicator of the effects of absorbing aerosols in altering atmospheric stability and in governing the interaction and feedback between aerosols and atmospheric dynamics. Another much-needed piece of information is global measurements of particle number concentration of both aerosols and cloud particles. Field experiments have provided and will continue to provide critical observational data in this regard. Current global networks including AErosol ROBotic NETwork (AERONET) and satellites such as Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) and CloudSat provide valuable information about the optical properties of aerosols or its loading in terms of total extinction, but they do not convey any direct information about aerosol number concentration, size distribution and chemical composition, which are key to understanding CCN effects. Out of necessity, AOD has been employed in large-scale modeling as a proxy for CCN because of its availability from satellite and surface networks. The two are indeed grossly correlated [Andreae and Rosenfeld, 2008], but their correlation is subject to a large uncertainty, often off by several factors of magnitude. This is not surprising because AOD is determined primarily by aerosol extinction that is, in turn, influenced by aerosol scattering, and thus proportional to the cross-section; CCN effects are affected by many other factors. Given appropriate composition and adequate super-saturation, aerosols of smaller size can provide more CCN than those of large size with the same bulk optical properties. Current attempts to use available satellite retrievals, such as those from MODIS, to estimate cloud droplet number concentration [e.g., Bennartz, 2007] are only of relative use for certain types of clouds. Next-generation satellites and global networks should be tailored toward extracting CCN as well as aerosol optical properties.

For additional in-depth details, the reader is referred to the following publications that illustrate the complexity of aerosol interactions with convection: Twomey et al. [1984], Albrecht [1989], Rosenfeld [1999, 2000], Khain et al. [2004, 2005, 2008], Cheng et al. [2007], Lynn et al. [2005], Van den Heever et al. [2006], Teller and Levin [2006], van den Heever and Cotton [2007], Wang [2005], Tao et al. [2007] and Levin and Cotton [2009].

Acknowledgments

W.-K. Tao appreciates the inspiring and enthusiastic support by his mentor, Dr. Joanne Simpson, over the past 25 years, and is grateful to Dr. R. Kakar at NASA headquarters for his continuous
support of Goddard Cumulus Ensemble (GCE) model development and applications. This work is mainly supported by the NASA Headquarters Physical Climate Program and the NASA PMM. He also thanks Dr. David Considine for support under the NASA MAP. C. Wang thanks the NSF (AGS-0944121), the Singapore NRF through the Singapore-MIT Alliance for Research and Technology, and the corporate and foundation sponsors of the MIT Joint Program on the Science and Policy of Global Change for supporting his research. Z. Li’s aerosol-related studies have been supported by NASA’s Radiation Science program managed by Dr. H. Maring (NNX08AH71G), and DOE’s Atmospheric System Research program managed by A. Williamson (DEFG0208ER64571), and the Ministry of Science and Technology of China (2006CB403706). C. Zhang thanks the support of the NOAA Office of Global Programs through awards under Cooperative Agreement No. NA17RJ1226 to the Cooperative Institute for Marine and Atmospheric Studies (CIMAS). J.-P. Chen acknowledges the support from projects NSC 98-2111-M-002-001 and NSC 99-2111-M-002-009-MY3.

We like to thank Ms. Maureen Cribb for reading/editing the manuscript and Dr. Xiaowen Li for improving the figures and tables in Section 5. We also like to thank Dr. Takamichi Iguhi for providing Figure 26.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACPI</td>
<td>aerosol-cloud-precipitation interaction</td>
</tr>
<tr>
<td>AERONET</td>
<td>AErosol RObotic NETwork</td>
</tr>
<tr>
<td>AI</td>
<td>aerosol index</td>
</tr>
<tr>
<td>AIE</td>
<td>aerosol indirect effect</td>
</tr>
<tr>
<td>AOD</td>
<td>aerosol optical depth</td>
</tr>
<tr>
<td>ARM</td>
<td>Atmospheric Radiation Measurement</td>
</tr>
<tr>
<td>AVHRR</td>
<td>Advanced Very High Resolution Radiometer</td>
</tr>
<tr>
<td>BC</td>
<td>black carbon</td>
</tr>
<tr>
<td>CALIPSO</td>
<td>Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation</td>
</tr>
<tr>
<td>CCM3</td>
<td>Community Climate Model version 3</td>
</tr>
<tr>
<td>CCN</td>
<td>cloud condensation nuclei</td>
</tr>
<tr>
<td>CDNC</td>
<td>cloud droplet number concentration</td>
</tr>
<tr>
<td>CGCM</td>
<td>coupled global circulation model</td>
</tr>
<tr>
<td>CIP</td>
<td>cloud ice path</td>
</tr>
<tr>
<td>CIR</td>
<td>cloud ice radius</td>
</tr>
<tr>
<td>CLWP</td>
<td>cloud liquid water path</td>
</tr>
<tr>
<td>CN</td>
<td>condensation nuclei</td>
</tr>
<tr>
<td>COD</td>
<td>cloud optical depth</td>
</tr>
<tr>
<td>CRM</td>
<td>cloud resolving model</td>
</tr>
<tr>
<td>CWF</td>
<td>cloud working function</td>
</tr>
<tr>
<td>CWR</td>
<td>cloud water radius</td>
</tr>
<tr>
<td>CWV</td>
<td>column water vapor</td>
</tr>
<tr>
<td>DER</td>
<td>droplet effective radius</td>
</tr>
<tr>
<td>DRF</td>
<td>direct radiative forcing</td>
</tr>
<tr>
<td>EHP</td>
<td>elevated heat pump effect</td>
</tr>
<tr>
<td>ENSO</td>
<td>El Niño / Southern Oscillation</td>
</tr>
<tr>
<td>FMF</td>
<td>fine mode fraction</td>
</tr>
<tr>
<td>GCCN</td>
<td>giant cloud condensation nuclei</td>
</tr>
<tr>
<td>GCM</td>
<td>general circulation model</td>
</tr>
<tr>
<td>IE</td>
<td>ice crystal enhancement</td>
</tr>
<tr>
<td>IN</td>
<td>ice nuclei</td>
</tr>
<tr>
<td>IPCC</td>
<td>Intergovernmental Panel on Climate Change</td>
</tr>
<tr>
<td>ITCZ</td>
<td>Intertropical Convergence Zone</td>
</tr>
<tr>
<td>LBA-SMOCC</td>
<td>Large-Scale Biosphere-Atmosphere Experiment in Amazonia-Smoke, Aerosols, Clouds, Rainfall, and Climate</td>
</tr>
<tr>
<td>LTSS</td>
<td>lower tropospheric static stability</td>
</tr>
<tr>
<td>LWP</td>
<td>liquid water path</td>
</tr>
<tr>
<td>MCS</td>
<td>mesoscale convective systems</td>
</tr>
<tr>
<td>MM5</td>
<td>fifth-generation Pennsylvania State University-National Center for Atmospheric Research (Penn State-NCAR) Mesoscale Model</td>
</tr>
<tr>
<td>MMF</td>
<td>multi-scale modeling framework</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>NAO</td>
<td>North American Oscillation</td>
</tr>
<tr>
<td>NCAR</td>
<td>National Center of Atmospheric Research</td>
</tr>
<tr>
<td>OC</td>
<td>organic carbon</td>
</tr>
<tr>
<td>PBL</td>
<td>planetary boundary layer</td>
</tr>
<tr>
<td>PC</td>
<td>principle component</td>
</tr>
<tr>
<td>PCA</td>
<td>principle component analysis</td>
</tr>
<tr>
<td>PDF</td>
<td>probability distribution function</td>
</tr>
<tr>
<td>PR</td>
<td>Precipitation Radar</td>
</tr>
<tr>
<td>RAMS</td>
<td>Regional Atmospheric Modeling System</td>
</tr>
<tr>
<td>SAL</td>
<td>Saharan Air Layer</td>
</tr>
<tr>
<td>SBM</td>
<td>spectral bin microphysical scheme</td>
</tr>
<tr>
<td>SGP</td>
<td>Southern Great Plains</td>
</tr>
<tr>
<td>TAV</td>
<td>Tropical Atlantic Variability</td>
</tr>
<tr>
<td>TC</td>
<td>tropical cyclone</td>
</tr>
<tr>
<td>TMI</td>
<td>TRMM Microwave Imager</td>
</tr>
<tr>
<td>TRMM</td>
<td>Tropical Rainfall Measuring Mission</td>
</tr>
<tr>
<td>WAM</td>
<td>West African Monsoon</td>
</tr>
<tr>
<td>WBF</td>
<td>Wagner-Bergeron-Findeisen</td>
</tr>
<tr>
<td>WRF</td>
<td>Weather Research and Forecasting Model</td>
</tr>
</tbody>
</table>
References


Andreae, M. O. (2009), Correlation between cloud condensation nuclei concentration and aerosol optical thickness in remote and polluted regions, *Atmos. Chem. Phys.*, 9(2), 543-556.


Fan, J., J. M. Comstock, M. Ovchinnikov, S. A. McFarlane, G. McFarquhar, and G. Allen (2010b), Tropical anvil characteristics and water vapor of the tropical tropopause layer:


Huffman, P. J. (1973), Supersaturation spectra of AgI and natural ice nuclei, *J. Appl. Meteor.,* 12, 1080-1082.

IPCC (2001), Intergovernmental Panel on Climate Change, Climate Change 2001: The physical science basis, edited by J. T. Houghton et al., 881 pp., Cambridge, United Kingdom and New York, NY, USA.

IPCC (2007), Intergovernmental Panel on Climate Change, Climate Change 2007: The physical science basis, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller, 996 pp., Cambridge, United Kingdom and New York, NY, USA.


Ming, Y., V. Ramaswamy, P. A. Ginoux, L. H. Horowitz, and L. M. Russell (2005), Geophysical Fluid Dynamics Laboratory general circulation model investigation of the indirect radiative


National Research Council (NRC) (2005), Radiative forcing of climate change: Expanding the concept and addressing uncertainties, The National Academies Press, Washington, D.C.


Niu, F., and Z. Li (2011b), The systematic variations of cloud top height and precipitation rate with aerosols over the global tropics seen from space, *ACP*, under revision.


Page | 94


