REVIEW

Review of Aerosol–Cloud Interactions: Mechanisms, Significance, and Challenges

JIWEN FAN
Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, Washington

YUAN WANG
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

DANIEL ROSENFELD
Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel

XIAOHONG LIU
Department of Atmospheric Science, University of Wyoming, Laramie, Wyoming

(Manuscript received 26 January 2016, in final form 24 June 2016)

ABSTRACT

Over the past decade, the number of studies that investigate aerosol–cloud interactions has increased considerably. Although tremendous progress has been made to improve the understanding of basic physical mechanisms of aerosol–cloud interactions and reduce their uncertainties in climate forcing, there is still poor understanding of 1) some of the mechanisms that interact with each other over multiple spatial and temporal scales, 2) the feedbacks between microphysical and dynamical processes and between local-scale processes and large-scale circulations, and 3) the significance of cloud–aerosol interactions on weather systems as well as regional and global climate. This review focuses on recent theoretical studies and important mechanisms on aerosol–cloud interactions and discusses the significances of aerosol impacts on radiative forcing and precipitation extremes associated with different cloud systems. The authors summarize the main obstacles preventing the science from making a leap—for example, the lack of concurrent profile measurements of cloud dynamics, microphysics, and aerosols over a wide region on the observation side and the large variability of cloud microphysics parameterizations resulting in a large spread of modeling results on the modeling side. Therefore, large efforts are needed to escalate understanding. Future directions should focus on obtaining concurrent measurements of aerosol properties and cloud microphysical and dynamic properties over a range of temporal and spatial scales collected over typical climate regimes and closure studies, as well as improving understanding and parameterizations of cloud microphysics such as ice nucleation, mixed-phase properties, and hydrometeor size and fall speed.

1. Introduction

Clouds regulate surface precipitation and the atmosphere’s radiative balance, therefore playing a significant role in the climate system. Clouds generally form when air is cooled and becomes supersaturated with respect to water or ice (except funnel clouds). The excess vapor generally cannot form cloud particles spontaneously owing to a high energy barrier, but rather condenses on aerosol particles that serve as either cloud condensation nuclei (CCN) or ice nuclei (IN). Therefore, changing aerosols is bound to impact cloud properties, precipitation, and cloud radiative effects. However, aerosol effects are entangled with dynamic and thermodynamic variables. The various parameters that in combination determine cloud properties include updraft speeds of air that form the
clouds, chemical and physical properties of aerosol particles on which cloud particles nucleate (R. Zhang et al. 2015), and cloud microphysical processes. Our poor ability to disentangle aerosol impacts on cloud radiative forcing from the meteorological effects in observations and poor parameterizations of convection and clouds in numerical simulations especially for large-scale models cause the largest uncertainty in current estimates of climate forcing, which resides in aerosol–cloud interactions (ACI) that are traditionally referred to as aerosol indirect effects (IPCC 2013).

How aerosols affect cloud properties and precipitation through ACI strongly varies among cloud types that are mainly controlled by atmospheric dynamics and thermodynamics. For warm clouds, the “Twomey” effect (i.e., reducing droplet size and increasing reflectance of clouds due to increased droplet number for a constant liquid water path) proposed about four decades ago (Twomey 1977) is relatively well understood. Many different aerosol indirect effects have since been suggested, such as increased cloud lifetime and cloudiness (Albrecht 1989) and suppressed rain (Rosenfeld 1999) that are both controlled by reduced droplet size and narrower droplet spectrum. Recent studies, which will be reviewed in detail in this study, mainly focused on how aerosols change microphysics and dynamic feedbacks in maritime stratocumulus clouds and affect cloud macrophysics such as the transitions between open and closed cells and from shallow to deep clouds.

For deep convective clouds (DCCs) with more complicated dynamics, thermodynamics, and microphysics, aerosol impacts are extremely complex and not as understood as those for shallow clouds. It has been hypothesized that aerosols might suppress warm rain, which allows more cloud water being lifted higher in the atmosphere, where freezing of the larger amount of cloud water releases more latent heat and invigorates convection (Rosenfeld et al. 2008). However, many modeling studies suggested that this thermodynamic invigoration is insignificant or even suppression of convection is seen, especially for clouds with cold cloud base, or strong wind shear, or dry condition (Fan et al. 2009, 2012b, 2013; Li et al. 2008b; Khain et al. 2005, 2008a; Tao et al. 2007; Lebo et al. 2012; Lebo and Seinfeld 2011). On the other hand, numerous observational studies showed the increased cloud-top height and cloud cover with an increase of aerosol loading (e.g., Andreae et al. 2004; Koren et al. 2010; Li et al. 2011; Niu and Li 2012). Fan et al. (2013) revealed a new cloud invigoration mechanism—microphysical invigoration induced by reduced ice particle size and fall velocity as an additional mechanism that explains the commonly observed increased cloud-top height and cloud cover.

Many recent studies also investigated aerosol impacts on clouds by acting as IN especially for dust, mainly through modifying heterogeneous nucleation (e.g., Li and Min 2010; Niemand et al. 2012; Creamean et al. 2013; Fan et al. 2014). IN directly change ice nucleation processes that determine the initial number concentration and size distribution of ice crystals. Here are the ice nucleation processes that are directly connected by IN: the homogeneous freezing of hazy aerosols that occurs spontaneously at temperatures below approximately −38°C and when the relative humidity with respect to ice (RHI) is larger than the threshold values (Ren and Mackenzie 2005) and heterogeneous ice nucleation through different modes: 1) deposition mode in which water vapor deposition on the surface of IN forms ice crystals, 2) immersion/condensation mode in which freezing of solution droplets on the surface of IN immersed within the droplets occurs, and 3) contact mode in which droplets freeze when their surface is in contact with IN from either from inside or outside of droplets. Various insoluble or partially insoluble aerosol particles can act as IN, such as mineral dust, carbonaceous aerosol, biological particles, and volcanic ash, and affect cloud and climate (Hoose and Möhler 2012; Murray et al. 2012; DeMott et al. 2010; Cziczo et al. 2004). New insights gained from these above-mentioned topics will be described in this review. Because of the space limit, we are not able to cover many specific topics associated with ACI such as cloud processing of aerosols and secondary activation.

In addition, the aerosol radiative effect by absorbing or scattering solar radiation, which is referred to as aerosol–radiation interaction (ARI) in the IPCC (2013) report, is also a significant pathway to change ambient meteorology conditions such as temperature and stability, cloud formation, convection, and even large-scale circulation (e.g., Lau et al. 2006; Bollasina et al. 2011; Nabat et al. 2015; Y. Wang et al. 2014c; Sanap and Pandithurai 2015; Fan et al. 2015a). Scattering aerosols generally cool the surface in clear-sky conditions. For strongly absorbing aerosol particles like black carbon (Peng et al. 2016b), they also heat some part of atmosphere depending on the locations (horizontally and vertically) besides cooling the surface, which changes atmospheric stability and even circulation (Y. Wang et al. 2013b), leading to complicated responses of clouds, radiation, and precipitation to aerosol loading (Yang et al. 2013a,b; Fan et al. 2015a; Yang et al. 2016).

Recently, a few review papers related to aerosol–cloud interactions have summarized past efforts that include fundamental theories (Tao et al. 2012), microphysics–dynamics feedback (Altaratz et al. 2014), observations (Rosenfeld et al. 2014a), cloud–resolving modeling (CRM) (Lee et al. 2014), and cloud microphysics parameterizations (Khain et al. 2015). However, some very recent and important findings are not included by those reviews. More importantly, the significance of aerosol
impacts is not systematically summarized, and the current problems and obstacles preventing us from moving forward need to be detailed. Building on those previous review papers, we will summarize recent findings of aerosol–cloud interaction mechanisms for different cloud types (i.e., shallow maritime clouds, deep convective clouds, mixed-phase stratiform clouds, and cirrus clouds) and then discuss the significance of aerosol impacts (i.e., radiative forcing, precipitation, extreme weather, and large-scale circulations), the current challenges in modeling and observations, and research directions needed toward reducing the uncertainties of climate prediction. Those topics have been the subject of the recent Aerosol–Cloud–Climate (ACC) symposiums that have been held at the American Meteorological Society (AMS) annual meeting. A special collection of papers has been created on aerosol–cloud–precipitation interactions in the Journal of the Atmospheric Sciences that are based on presentations from the recent ACC symposiums. This review also serves as an introduction to the special collection.

We will concentrate on ACI with ARI discussed jointly since in the real world ACI and ARI are not separated, and many studies looked at the overall effect without a separation. We discuss the physical mechanisms of ACI for different cloud types first, focusing on recent findings (section 2). Then we provide a review of the most striking aerosol impacts on different cloud systems in terms of precipitation and cloud regional and global radiative forcing in section 3. In the last section, we summarize the issues that prevent us moving forward and suggest future directions.

2. Mechanisms of ACI for different cloud types

One factor contributing to the complexity of ACI is the dominant microphysical and dynamical processes that vary among different cloud types. The dynamic responses to microphysical changes are quite different for different types of clouds; therefore, the dominant ACI mechanisms depend on cloud types as well as the stages of cloud evolution. The following discussion is based on cloud types.

a. Warm clouds—Shallow cumuli and stratocumuli

Much of the previous research on aerosol indirect effects has focused on low-level warm clouds, largely because they strongly reflect solar radiation back to space and cool the surface without impacting outgoing longwave radiation much. This is true especially for marine stratocumulus clouds, which cover roughly one-third of the global oceans (Warren et al. 1988) and act as “air conditioners” to the climate system (Stephens and Slingo 1992). Since warm clouds do not involve mixed-phase and ice phase regimes, they are less complicated microphysically, and thus we have a better understanding of the general aerosol indirect effects compared with deep convective clouds. For warm clouds, we know that aerosols increase cloud albedo (Twomey 1977; Coakley et al. 1987), suppress collision and coalescence processes, and, thus, reduce warm rain (Albrecht 1989; Rosenfeld and Lensky 1998; Rosenfeld 2000), elongate cloud lifetime, and increase cloud cover (Albrecht 1989; Kaufman et al. 2005; Yuan et al. 2011b). IPCC (2007) has included a detailed summary of the main findings in the past about aerosol impacts on warm clouds. Here we focus on the recent findings on the understanding of ACI for warm shallow cumuli and stratocumuli.

1) AEROSOL IMPACT ON FORMATION AND ORGANIZATION OF BOUNDARY LAYER CLOUDS

Warm boundary layer clouds form by convective processes, which are driven by surface heating and/or by cloud-top radiative cooling. Surface heating tends to form cumuliform clouds, whereas cloud-top radiative cooling creates decks of stratus or stratocumulus clouds (Rayleigh 1916; Agee et al. 1973). For shallow cumuli over land, field campaign and modeling studies showed that CCN form more cloud droplets but reduce droplet size (e.g., Gustafson et al. 2008; Xue and Feingold 2006; Shrivastava et al. 2013). Therefore, droplet evaporation is enhanced and the feedback between evaporation–entrainment is enhanced as well (Xue and Feingold 2006). Transitions from shallow cumulus or stratus clouds that are radiatively cooled at their tops to convective clouds that are thermally heated from the land surface are common upon transitioning from early morning stratus clouds to late morning convective clouds. Studies showed that increasing CCN reduced the amount of shallow cumulus clouds by stronger evaporation–entrainment feedback but increased the amount of deep convective clouds (Saleeby et al. 2015), suggesting CCN could enhance the transition from shallow to deep convective clouds through microphysics and dynamics feedbacks.

The causes of such transitions over ocean are not so obvious but still depend on these two factors—cloud-top cooling and surface heating, which are modulated by cloud thickness and height in complicated ways that strongly involve aerosol effects. Rosenfeld et al. (2006) showed that the transition from overcast stratocumulus decks to broken shallow marine convective clouds is associated with a transition from dominant radiative cooling at cloud tops to sensible heating at the surface. Subsequent simulations showed that surface sensible heat flux is essential for maintaining a regime of open cells of marine stratocumulus (Kazil et al. 2014). Satellite observations revealed reduced cloud drop effective radius and increased drop concentrations on the
transitions from open to closed cells (Rosenfeld et al. 2006; Goren and Rosenfeld 2014). This is most evident in ship tracks, which closed the areas of open cells (Goren and Rosenfeld 2012). Satellite observations also revealed an increased occurrence of stratocumulus clouds as aerosol index (AI) increases over ocean, further supporting the hypothesis that aerosols enhance transition of shallow cumulus to stratocumulus (Gryspeerdt et al. 2014). Aerosol impact on cloud organization starts from nucleating a larger number of smaller cloud droplets, which slow drop coalescence and delay or completely suppress warm rain. The dynamic response of such clouds to the changes in rain strongly affects their organization, liquid water path, cloud cover, and hence their radiative effects, which feed back to the cloud organization. A more detailed review of this chain of events is provided next.

2) AEROSOL EFFECTS FOR NONPRECIPITATING CLOUDS

Based on the Twomey effect, a larger number of drop concentrations $N_d$ for the same liquid water path result in correspondingly smaller cloud drop effective radius $r_e$. Because coalescence rate depends on $r_e^{2.8}$ (Freud and Rosenfeld 2012), the collision and coalescence does not lead to much rain formation when $r_e$ is smaller than $\sim$14 $\mu$m, while above this value the coalescence rate increases very fast (Freud and Rosenfeld 2012; Gerber 1996). For nonprecipitating clouds, cloud droplets are usually small and increasing aerosols makes even smaller cloud droplets, which evaporate much faster when mixing with the ambient dry air. This leads to stronger evaporative cooling and mixing at the cloud boundaries and loss of cloud water (Randall 1980a,b). This effect may sometimes overcome the cloud brightening due to the Twomey effect and lead to smaller cloud water content and lower cloud albedo in polluted conditions (Chen et al. 2012). Many previous modeling studies showed similar results (e.g., Xue et al. 2008; Hill et al. 2009).

3) RELATIONSHIPS BETWEEN AEROSOLS, CLOUD DEPTH, AND PRECIPITATION

Marine stratocumulus clouds are typically only a few hundred meters deep. For significant precipitation to occur, cloud drop concentrations $N_d$ should be mostly below $100 \text{ cm}^{-3}$ and $r_e$ at cloud top should exceed 14 $\mu$m. This is an aerosol-limited condition, where CCN concentrations dominate the variability in $N_d$, whereas cloud-base updraft $W_b$ plays a secondary role. Zheng and Rosenfeld (2015) recently suggested that the average $W_b$ of boundary layer convective clouds can be estimated by cloud-base height $H_b$ according to $W_b = 0.9H_b$, where $W_b$ is in meters per second and $H_b$ is in kilometers. As shown and validated observationally by Freud and Rosenfeld (2012), a nearly linear relationship must exist between $N_d$ and the critical cloud depth for rain initiation $D_c$, as depicted by the cloud depth where $r_e$ reaches 14 $\mu$m. Therefore, more CCN are required for suppressing precipitation in deeper clouds. Satellite observations showed that marine stratocumulus precipitates significantly and breaks up when their geometrical depth exceeds $D_c$ (Goren and Rosenfeld 2015).

Interestingly, recent studies consistently found that adding CCN to warm clouds with very low $N_d$ can invigorate them and enhances their vertical development, leading to taller clouds, larger cloud water content, and enhanced rain rates (Yuan et al. 2011b; Christensen and Stephens 2011, 2012; Y.-C. Chen et al. 2015). Yuan et al. (2011b) analyzed satellite data and showed a significant increased cloud amount by volcanic aerosols for trade wind cumuli. Satellite observations from the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument showed that liquid water path and cloud-top heights of ship tracks were higher than the ambient clouds when the ship tracks were imbedded in open cells (Figs. 1a and 1b) that had much smaller $N_d$ than ship track clouds (Y.-C. Chen et al. 2015). These ship track clouds were taller because of enhanced updraft speeds as inferred by the increased divergence rate of their tops (Fig. 1c), as observed by the Multiangle Imaging Spectroradiometer (MISR) on the Terra satellite (Y.-C. Chen et al. 2015). CloudSat spaceborne radar measurements also showed that the invigoration occurred when rain was not suppressed owing to the increased $N_d$ but, rather, enhanced owing to enhanced updrafts and deeper clouds (Christensen and Stephens 2011). This effect was first simulated by Pincus and Baker (1994), who showed that the clouds became higher with added CCN to a background as low as $10 \text{ cm}^{-3}$ owing primarily to enhanced entrainment induced by the faster evaporation of the smaller droplets. Additional causes for invigoration are the faster supersaturation consumption with more cloud drop surface area available for vapor condensation, resulting in reduced vapor supersaturation and greater latent heat release. This mechanism was highlighted as a cause for invigoration of warm cumulus congestus clouds by Koren et al. (2014), who showed a very strong association between aerosol optical depth and cloud depth, fractional coverage, and rain intensity. However, the quantitative strength of the reported relationships is likely exaggerated owing to possible effects of clouds on aerosols (Zhu et al. 2015).

Therefore, it seems clear that open cell clouds (i.e., clouds with partial cover that are composed of small concentrations of large droplet) exhibit a large increase
in cloud water, cloud depth, and precipitation in response to modest increase of aerosols. Additional increase of aerosols eventually suppresses the precipitation and may lead to formation of closed cells (i.e., nearly continuous stratocumulus cloud deck). The closed cells are composed of numerous small droplets and usually precipitate lightly if at all. Adding aerosols to closed cells further suppresses the precipitation by further decreasing droplet size and enhancing entrainment due to strong evaporation (Ackerman et al. 2004). Thus, the direction and magnitude of the precipitation response are strongly influenced by cloud properties and mesoscale stratocumulus cloud regime (Christensen and Stephens 2012), consistent with the conclusion from the previous large-eddy simulations (Wang and Feingold 2009) that dynamic response to aerosol differs between regimes. In any case, rain would be suppressed if cloud-top height is lower than the height where coalescence becomes significant ($D_c$), and aerosols could increase $D_c$ beyond the actual cloud depth (Goren and Rosenfeld 2015). Thus, to gain confidence in simulations of aerosol indirect effects in global climate models, it is important to quantify the distributions of the various marine stratocumulus cloud types and how these might change in the future.

4) AEROSOL-INDUCED TRANSITIONS BETWEEN CLOSED AND OPEN CELLS OF MARINE STRATOCUMULUS

Nonprecipitating clouds that are maintained by radiative cooling at their tops usually occur as solid cloud decks of closed cells. These clouds break up when they start precipitating heavily (few millimeters per day for marine stratocumulus). The precipitation can occur as a result of gradual cleansing of aerosols by cloud/precipitation processing and/or by cloud deepening (Goren and Rosenfeld 2015). The breakup mechanism involves a combination of losing much of the cloud water to precipitation and rain evaporation-cooled downdrafts that form mini gust fronts near the surface and trigger convective clouds when the gust fronts from adjacent cells merge (Rosenfeld et al. 2006; Feingold et al. 2010; Wood et al. 2011). The precipitation keeps scavenging CCN and maintains extremely low concentrations of CCN in a self-perpetuating mechanism. Furthermore, a runaway effect of scavenging can occur, where there are simply no CCN left for allowing effective condensation and formation of clouds, thus causing the collapse of the marine boundary layer (Ackerman et al. 1993). This situation can be reversed, as evident by observations of ship tracks closing open cells over large areas (Goren and Rosenfeld 2012) and by simulations of adding aerosols to open cells (Feingold et al. 2015). The transitions between open and closed cells represent two stable situations with feedback for self-maintenance, representing a bistability of the two states, with an unstable short-lived and narrow transition between them (Baker and Charlson 1990). Upon the transition from open to closed cells, the cloud radiative effect increases, on average, by more than 100 W m$^{-2}$, with only about $1/4$ of the effect contributed by the Twomey effect, $1/3$ by the LWP effect, and the rest by the cloud cover effect (Goren and Rosenfeld 2014). When considering it in the framework of buffering of the primary aerosol effects due to system response (Stevens and Feingold 2009), this represents a 4-times amplification of the Twomey effect. Such effects are limited to marine boundary layer
clouds and have not been documented in clouds over land, where the effect could be less significant as a result of more aerosols in the background condition.

b. Mixed-phase stratiform clouds

Mixed-phase clouds are composed of a mixture of supercooled liquid droplets and ice crystals. It is a dominant cloud type during the colder three-quarters of the year in the Arctic, with liquid on top and ice forming and precipitating beneath (Pinto 1998; Curry et al. 2000). Arctic mixed-phase clouds (AMPC) are often long lived and can persist for several days (Morrison et al. 2011). At lower latitudes, mixed-phase stratiform clouds can occur from deep convection or form from synoptic-scale mid-latitude weather systems (Hogan et al. 2004; Larsson et al. 2006). Phase transformations of water between vapor and liquid and ice particles can occur in the mixed-phase clouds, where ice can grow at the expense of liquid because of the lower saturation vapor pressure over ice compared with that over liquid. This process is known as the Wegener–Bergeron–Findeisen (WBF) mechanism (Wegener 1911; Bergeron 1935; Findeisen 1938). However, the WBF process can only occur under a limited range of conditions because it requires that the vapor pressure exceeds ice saturation but is below liquid saturation (Korolev 2007; Korolev and Mazin 2003). In the updrafts of the mixed-phase clouds, the vapor pressure often exceeds saturation for both liquid and ice (Fan et al. 2011); as such, both droplets and ice particles will grow simultaneously. Based on large-eddy simulations of single-layer and multilayer AMPC, Fan et al. (2011) showed that the WBF process occurs in only about 50% of the mixed-phase regimes, predominantly in downdrafts.

Despite the microphysical instability arising from the WBF process, mixed-phase clouds have a self-maintaining feedback pathway between liquid water, radiation, and turbulence, which explains their persistence (Morrison et al. 2011). Supercooled liquid water leads to strong longwave radiative cooling near the cloud top, which decreases static stability and enhances turbulent updrafts and then condensational growth of droplets (Curry 1986; Solomon et al. 2011). In the Arctic, frequent moisture inversions exist near cloud top owing to large-scale advection. Under such circumstances, the entrainment of air from above the cloud actually moistens the cloud layer and helps to sustain it against the near-continual mass loss resulting from ice precipitation (Solomon et al. 2011).

Both CCN and IN can impact mixed-phase cloud properties and therefore the persistence of mixed-phase clouds. There are not many studies focusing on aerosol impacts on mixed-phase stratiform clouds at mid-latitudes. Fan et al. (2012a) studied how aerosol impacts a mixed-phase stratiform cloud formed after the passage of a midlatitude cold front. They showed that increasing CCN leads to enhanced liquid water content (LWC) due to stronger condensation as a result of enlarged droplet surface area but suppressed precipitation resulting from a much reduced droplet size. Zhang et al. (2012) showed that the occurrences and ice water path in the supercooled stratiform clouds over the “dust belt” on the globe are enhanced compared with the background aerosol conditions, and the enhancements are strongly dependent on the cloud-top temperature, large dust particle concentration, and chemical compositions. For AMPC, plenty of studies on aerosol–cloud interactions have been carried out because of the critical importance of AMPC to the Arctic climate (Lubin and Vogelmann 2006; Garrett and Zhao 2006; Garrett et al. 2009; Prenni et al. 2007; Morrison et al. 2011; Ovchinnikov et al. 2011). For example, the amount of liquid water in AMPC has a large impact on surface radiative fluxes and energy balance, which could affect ice-melting rate (Carrió et al. 2005).

Increased CCN lead to increased cloud droplet concentrations and reduced droplet size in AMPC, which increases longwave radiative emissivity of clouds (Lubin and Vogelmann 2006; Garrett and Zhao 2006; Garrett et al. 2009). The increase in downwelling longwave radiation due to CCN effects can result in surface warming, which may increase surface turbulent fluxes and provide a greater source of moisture, accelerating the positive feedback loop between cloud-top radiative cooling, turbulence, and condensation of droplets (Garrett et al. 2009). Moreover, it is shown that polluted mixed-phase clouds have narrower droplet size distributions and contain one to two orders of magnitude fewer precipitating ice particles than clean clouds at the same temperature, which leads to longer cloud lifetime, greater cloud emissivity, and reduced precipitation (Lance et al. 2011). This result is opposite to the “glaciation indirect effect” caused by increasing IN, which will be discussed below. As for CCN impact on surface precipitation, Borys et al. (2003) showed the smaller droplet size makes riming less efficient and decreases the snow precipitation. However, Lohmann et al. (2003) found whether CCN increase or decrease surface precipitation is crucially dependent on the crystal shape because of very different accretion efficiency of snow crystals with cloud droplets for different ice habits. Earle et al. (2011) analyzed aircraft observations and found that polluted cases were correlated with warmer, geometrically thicker clouds, with higher Nd, LWP, and albedo relative to clean cases. But droplet effective radius was similar, suggesting the complex interactions among environmental conditions, aerosol, and the microphysics and radiative properties of Arctic clouds.
Although IN concentration is typically over five orders of magnitudes lower than CCN concentrations, it is critical for mixed-phase clouds because it enhances WBF and riming processes (Fan et al. 2014). Cloud-resolving model studies have shown that by increasing IN number concentration by 2–3 times, a liquid stratus deck can be transformed into a broken, optically thin ice cloud system (Harrington et al. 1999; Jiang et al. 2000; Prenni et al. 2007; Morrison et al. 2011; Ovchinnikov et al. 2011). Most of these studies suggested that the rapid glaciation is due to the WBF, while Ovchinnikov et al. (2011) suggested that the modified feedback (i.e., less liquid water reduces radiative cooling and slows vertical mixing) also contributes to the rapid diminishing of the mixed-phase clouds under a relatively high IN condition. This glaciation indirect effect leads to ice clouds that have a much lower particle number density and reduced cloud optical depth and settle relatively rapidly (thereby reducing cloud lifetime). Many recent studies on AMPC focused on IN sources, ice nucleation mechanisms, and parameterizations (Jackson et al. 2012; M. Fan 2013; de Boer et al. 2013; Paukert and Hooge 2014; Savre and Ekman 2015). IN recycling was found to be an important IN source that plays an important role in determining liquid and ice partitioning and lifetime of AMPC (Fan et al. 2009; Solomon et al. 2015).

c. Deep convective clouds

DCCs vary in all kinds of forms from tropics to subtropics, from land to ocean, and from islands to mountain ranges (Houze et al. 2015). Aerosol–DCC interactions truly depend on different convective systems. Most of past studies focused on the mechanisms of aerosol impacts on individual convective clouds, which have been intensively reviewed recently (Tao et al. 2012; Altaratz et al. 2014; Rosenfeld et al. 2014a). A series of possible pathways about how aerosols may change Earth’s energy budgets in all their forms through impacting DCCs have been detailed in Rosenfeld et al. (2014a). Therefore, we focus on recent findings that are not included by those previous reviews.

1) CCN EFFECTS

Aerosol impacts on DCCs are extremely complicated owing to complicated dynamic and thermodynamic conditions associated with DCCs and also to the wide span of cloud phases including warm, mixed-phase, and ice clouds. Past studies have revealed that relative humidity (RH), wind shear, and convective available potential energy (CAPE) are the major factors impacting the significance of aerosol impacts on convective intensity, precipitation, and cloud radiative forcing (e.g., Khain et al. 2008b; Khain 2009; Fan et al. 2009, 2012b; Tao et al. 2012; Storer et al. 2010; Storer and van den Heever 2013), because those factors regulate the dominant microphysical processes and the microphysics–dynamics feedbacks. As said above, aerosol effects on DCCs are strongly dependent on different kinds of convective systems. For the systems that are mainly triggered by gust fronts, aerosols may change the intensity and size of gust fronts, which then change the cloud-system organization and impact precipitation and cloud macrophysical properties (Lee 2012; Lee et al. 2014; Morrison 2012; Lebo and Morrison 2014). Supercell systems that are strongly driven by dynamic pressure perturbations are less sensitive to aerosols (Storer et al. 2010; Lebo et al. 2012; Morrison 2012). However, in some situations, adding aerosols leads to larger hydro-meteors, less evaporative cooling, and weaker gust fronts, which do not undercut the updrafts and allows the maintenance of super cells and severe convective storms (Rosenfeld and Bell 2011).

The basic theory proposed for aerosol–DCC interactions is the thermodynamic invigoration through more latent heat release from clean to polluted clouds (Rosenfeld et al. 2008). This thermodynamic effect can be very significant under the conditions of warm-cloud bases (>15°C) where the warm-cloud zone is deeper and the suppression of warm rain by aerosols can be more significant (Li et al. 2011; Rosenfeld et al. 2014a; Fan et al. 2012b) and weak wind shear where convection is nearly vertical and the increase of latent heating can dominate over the increase of evaporative cooling (Fan et al. 2009, 2012b, 2013; Li et al. 2008b; Khain et al. 2005, 2008b; Tao et al. 2007; Lebo et al. 2012). In many cases where strong wind shear exists and/or cloud bases are cold, aerosols could strongly suppress convection and precipitation due to strong evaporative cooling of the small cloud droplets and/or less efficient ice-growing processes (Khain et al. 2008b; Iguchi et al. 2008; Lebo and Seinfeld 2011; Morrison 2012; Fan et al. 2009, 2012b).

Although the thermodynamic invigoration can increase cloud-top height and cloud cover, a recent study (Fan et al. 2013) has found that it is often not the main mechanism leading to the significant increase of cloud cover and cloud-top height (CTH) from the clean to polluted cloud cases on the scale of entire cloud life cycle even for warm-based summer convective clouds. The increased cloud cover and CTH by aerosols were reported consistently by many observational studies (Andreae et al. 2004; Koren et al. 2010; Li et al. 2011; Niu and Li 2012; Storer and van den Heever 2013; Yan et al. 2014; Peng et al. 2016a), in which the thermodynamic invigoration was often hypothesized to explain the results. According to Fan et al. (2013), the simulated
aerosol thermodynamic invigoration contributed up to 27% to the increased cloud cover at the upper levels even for the warm summer convective clouds. The dominant mechanism contributing to the observed increased cloud cover and CTH is a microphysical aerosol effect: the freezing of a larger number of smaller droplets produces more numerous but much smaller ice particles in the stratiform regime of polluted clouds which leads to much reduced fall velocities of ice particles and slows the dissipation of stratiform and anvil clouds significantly (Fig. 2). This microphysical invigoration occurs even when thermodynamic invigoration of convection is absent, which explains the ubiquitously observed increased cloud cover and CTH with increasing aerosol loading. It should be noted that when the thermodynamic invigoration does occur under favorable conditions such as warm cloud base and relatively weak wind shear as seen from cases in the tropical west Pacific (TWP) and southeast China in Fan et al. (2013), it adds to the microphysical invigoration and leads to more significant increase of CTH and cloud fraction. Aerosol impacts on mesoscale convective systems (MCSs) have not been established. However, those results strongly suggest that the stratiform regions of MCSs (which account for a lot of the precipitation and effect on radiative transfer) may be affected by the ingestion of aerosols.

In summary, the reasons for the relatively small contribution of the aerosol thermodynamic invigoration over a long time period and large region could be 1) the increased CTH, cloud cover, and cloud thickness due to microphysical invigoration could produce strong TOA cooling and surface cooling (Fan et al. 2013), which reduces surface temperature and surface sensible and latent heat fluxes and weakens the convection in the polluted case compared with the clean case; and 2) over a regional domain, it is found that large-scale dynamic adjustment buffers the thermodynamic invigoration as well

![Figure 2. Mechanism describing the impact of aerosols on the whole life cycle of deep convective clouds (Fan et al. 2013). In the polluted environment, convective cores detrain larger amounts of cloud hydrometeors of much smaller sizes, leading to larger expansion and much slower dissipation of stratiform/anvil clouds resulting from smaller fall velocities of ice particles due to reduced sizes. Therefore, larger cloud cover, higher CTH, and thicker clouds are observed in the polluted storm.](image-url)
(Morrison and Grabowski 2011; van den Heever et al. 2011).

For DCCs, because of the system complexity and the strong feedback of microphysics to dynamics, they can be very sensitive to small perturbations of dynamic and thermodynamic fields because of rapid, nonlinear growth of the perturbations and solution drift among different realizations of DCCs (Hack and Pedretti 2000; F. Zhang et al. 2007; H. Wang et al. 2012). This can make it difficult to ascertain the robustness of aerosol impacts based on single realizations (Morrison and Grabowski 2011; Morrison 2012). Therefore, various approaches were proposed to reduce the uncertainty caused by natural variability, including an ensemble of simulations (Morrison 2012), using large horizontal domains or multiple-case simulations (Fan et al. 2013), and employing a “piggybacking” approach (Grabowski 2014, 2015). The microphysical piggybacking studies of Grabowski (2014, 2015) suggest that the feedback of cloud microphysics by increasing $N_d$ to dynamics is small in both shallow and deep convective clouds. However, the simple one-moment microphysics parameterizations employed in those studies do not capture the main microphysical processes of cloud–aerosol interactions. This raises another uncertainty issue—the complexity of cloud microphysics parameterizations, which will be reviewed in a separate paper.

2) IN EFFECTS

Besides acting as CCN, aerosols such as dust particles can strongly impact deep convective clouds and mixed-phase clouds by acting as effective IN (e.g., DeMott et al. 2010; Kulkarni et al. 2012). Asian desert dust was found to be good IN (You et al. 2002; Connolly et al. 2009; Field et al. 2012; Niemand et al. 2012; Jiang et al. 2014, 2015). Studies have reported that long-range transport of dust from Asia may have enhanced snow formation in California winter storms (Creamer et al. 2013; Fan et al. 2014). Because of large uncertainty in ice nucleation parameterizations (DeMott et al. 2010) and our limited knowledge on ice formation in DCC, most studies have focused on model sensitivity tests (van den Heever et al. 2006; Connolly et al. 2006; Fan et al. 2010a). Studies showed that updrafts were enhanced as a result of added latent heat release from ice crystal depositional growth by increasing IN for the deposition nucleation mode (Ekman et al. 2007). Enhanced updrafts in turn enhanced homogeneous ice nucleation, increasing anvil cloud coverage and precipitation. The enhanced updrafts and homogeneous ice nucleation by dust IN were also shown in a regional climate study over Asia based on simulations for 2006 and 2010 (Yang et al. 2015). The CRM study by van den Heever et al. (2006) showed the enhanced precipitation by increasing IN as well. However, some CRM studies have shown that IN do not have a significant impact on convective IN (Connolly et al. 2006; Yin and Chen 2007; Fan et al. 2010a) but do significantly impact cloud microphysical properties, leading to an increase in cloud anvil fraction (Fan et al. 2010a).

Some of the important IN effects on DCCs are increasing cloud glaciation temperatures and transforming mixed phase into pure-ice phase, which impacts cloud radiative forcing and precipitation. Glaciation in clouds affected by dust and polluted aerosols was observed to occur at relatively high temperatures near −20°C based on satellite measurements (Rosenfeld et al. 2011, 2014a).

There have been a series of observational studies related to effects of the Saharan air layer (SAL) on tropical thunderstorm complexes. The SAL is an extremely hot, dry, and often dust-laden layer of the atmosphere between 850 and 500 hPa. It has been observed that the SAL interacts with tropical cloud systems and impacts their intensity and evolution (Karyampudi and Carlson 1988; Dunson and Velden 2004; Min et al. 2009; Twyoh 2015). Satellite observations indicated that dust in the SAL reduces convective precipitation but increase stratiform precipitation (Min et al. 2009) and cloud-top temperature (Li and Min 2010). Those studies hypothesized that dust particles in the SAL serve as effective IN, forming large amount of small ice particles and impacting mixed-phase cloud processes (Min et al. 2009; Li and Min 2010). To better understand the mechanisms of how dust aerosols interact with tropical convective clouds, detailed modeling studies are necessary.

IN impacts on regional climate have also been investigated by a few studies. By employing an ice nucleation scheme for heterogeneous ice nucleation that connects ice nucleation rate with dust concentration and surface area in Niemand et al. (2012), Y. Zhang et al. (2015) showed that in northern China where dust is abundant, dust significantly increases ice particle concentrations, ice water path, precipitation, and shortwave and longwave cloud radiative forcing averaged over a 2-yr time period.

d. Cirrus clouds

Cirrus clouds frequently occur in the upper troposphere and play an important role in regulating the radiation budget of the earth–atmosphere system and, thus, impact profiles of atmospheric heating (Yang et al. 2015). Cirrus clouds come in a variety of forms, ranging from optically thick anvil cirrus closely associated with deep convection to optically thin cirrus evolved from detached anvil or generated in situ by the synoptic-scale uplift of a
nucleation, more IN loading will increase cloud formation is dominated by the heterogeneous ice effect (Koop et al. 2000). In contrast, there are still many unknowns about the concentrations and properties of IN from heterogeneous nucleation, their dominant modes of action, and competition between the modes. This uncertainty is due to a lack of reliable instruments that measure the tiny fractions (one of $10^2$–$10^6$) of IN among the total aerosol population in ambient environments (DeMott et al. 2010). Even under controlled laboratory environments, instruments with different designs have large spreads in the measured ice nucleation fractions for the same aerosol samples (Hiranuma et al. 2015). A recent analysis of ice residuals collected in several field campaigns (Cziczo et al. 2013) points to the dominant role of heterogeneous nucleation. However, these campaigns were mainly targeted at the convective anvils or cirrus clouds nearby the convection that created them, where mineral dust and other types of IN can be uplifted to the upper troposphere by convection. Thus, the general conclusion of this study may not be applicable to in situ cirrus or other geographical regions (e.g., mid-latitudes and polar regions).

Aerosol impacts on cirrus clouds are largely determined by the dominant nucleation mechanisms or the balance between homogeneous versus heterogeneous nucleation (Liu et al. 2005; Kärcher et al. 2006; Barahona and Nenes 2009; Gettelman et al. 2012). For example, in a region where cirrus cloud formation is dominated by homogeneous ice nucleation, more IN loading would, by inhibiting homogeneous nucleation (Liu et al. 2012), decrease the number concentration $N_i$ and increase the size (e.g., number-mean diameter $D_i$) of ice crystals. It is equivalent to the “negative” Twomey effect (Kärcher et al. 2006). In a region where cirrus cloud formation is dominated by the heterogeneous ice nucleation, more IN loading will increase $N_i$ and reduce $D_i$ of ice crystals. The abundance of mineral dust from natural sources affects the balance between homogeneous and heterogeneous ice nucleation. Homogeneous nucleation has been suggested to be the most likely mechanism for the cirrus formation at Southern Hemisphere midlatitudes, while heterogeneous nucleation exists in cirrus in parts of the polluted Northern Hemisphere as revealed from the distinct differences in RHI freezing thresholds from in situ measurements of cirrus clouds during the Interhemispheric Differences in Cirrus Properties From Anthropogenic Emissions (INCA) field experiment (Haag et al. 2003). For anvil cirrus formed from the detrainment of deep convection, homogeneous freezing of liquid drops that is independent on RH is a major ice formation pathway (Khain et al. 2005). Heterogeneous ice nucleation at convective area should also impact ice properties of anvil cirrus (Fan et al. 2010b). Anvil cirrus can be impacted by aerosols through cloud microphysics processes in convective clouds (section 2c). It is found that aerosols significantly increase ice particle number but reduce ice particle size and fall velocity in DCCs, leading to an increase of anvil cirrus cloud cover and cloud-top height (Fan et al. 2013).

Besides ice nucleation, many other ice microphysical processes also play important roles in the properties, maintenance, and lifetime of cirrus clouds, including depositional growth/sublimation of ice crystals and sedimentation of ice crystals. A change of ice particle size distribution (PSD) toward small ice crystals (diameter less than 60 $\mu$m) can increase cloud ice amount by 12% and cirrus cloud cover by 5.5% globally by affecting the ice sedimentation rates (Mitchell et al. 2008). The cirrus cloud properties have been reported to vary significantly between earlier and later phases of cirrus clouds (Diao et al. 2013). These large variations in cirrus properties between various phases can potentially bias our interpretation of aerosol microphysical effects on cirrus clouds. The microphysical processes in the later phase could buffer the aerosol effects on ice crystal properties through ice nucleation in the earlier phase of cirrus clouds.

3. Significance of ACI
   a. Radiative forcing

According to the IPCC 2013 (Stocker et al. 2013), the industrial-era (1750–present) effective radiative forcing of aerosol-induced cloud adjustment was reported to be between $-1.33$ and $-0.6$ W m$^{-2}$ with a low confidence level. Such a wide range in the global mean of radiative forcing mainly arises from different representations of buffering mechanisms that result in compensation between distinctive cloud responses to different types of aerosols (Rosenfeld et al. 2014a) as well as various parameterizations of aerosol–cloud interactions with diverse degrees of sophistication in GCMs. As forcing values are dependent of cloud types, domain, and time, we use a table (Table 1) to summarize the forcing values from aerosol–cloud interactions discussed in this section to include detailed information such as cloud type, aerosol type, location, time period, and methods. The
following detailed discussion starts with global studies and then focuses on regional studies. 

Ekman (2014) suggested that the sophisticated parameterization of cloud droplet formation as a function of both aerosol concentration and supersaturation improves the simulation of the historical surface temperature trend. Aerosol indirect effects in GCMs are also influenced by the treatment of rain. For example, the implementation of prognostic equations for raindrops in cloud microphysics of GCMs (Posselt and Lohmann 2009; Gettelman and Morrison 2015) reduces aerosol indirect radiative forcing from 0.5 to 0.9 W m$^{-2}$ owing to the increased accretion rates and shifted distribution of raindrops. The results from the multiscale modeling framework (MMF) in which cumulus parameterization is not needed (M. Wang et al. 2011) showed that simulated change in shortwave cloud forcing from anthropogenic aerosols is $-0.77$ W m$^{-2}$, which is less than half of that ($-1.79$ W m$^{-2}$) calculated by the host CAM5 with traditional cumulus parameterizations. Recent studies (Song and Yum 2012; Min and Zhang 2014) also suggested that the estimates of diurnal mean aerosol radiative forcing also depends on the soundness of simulating of cloud diurnal variations in GCMs.

Global observational estimation and constraint of the aerosol indirect effect (AIE) is mainly derived from satellite and ground-based measurements. Quaas et al. (2009) utilized independent sets of satellite measurements to quantify the relationships between aerosol optical depth, cloud droplet concentration, liquid water path, and TOA radiative fluxes and estimated the global annual mean short-wave aerosol forcing as $-1.5 \pm 0.5$ W m$^{-2}$ inferred from the combination of these predictors for the modeled forcing with the satellite-derived relationships. As a further step, M. Wang et al. (2012) employed the satellite observations to derive the dependence of the probability of precipitation on aerosol

### Table 1. Summary of AIE forcing values from various studies [values are at TOA except Zhao and Garrett (2015), which was estimated based on surface measurements]. The table intends to summarize the studies on the AIE forcing in recent years.

<table>
<thead>
<tr>
<th>AIE forcing (W m$^{-2}$)</th>
<th>Cloud type</th>
<th>Aerosol type</th>
<th>Location</th>
<th>Period</th>
<th>Method/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.77$ [shortwave (SW)]</td>
<td>All</td>
<td>All</td>
<td>Global</td>
<td>Multiyear</td>
<td>MMF (M. Wang et al. 2011)</td>
</tr>
<tr>
<td>$-1.5 \pm 0.5$</td>
<td>All</td>
<td>All</td>
<td>Global</td>
<td>Multiyear</td>
<td>GCM + satellite (Quaas et al. 2009)</td>
</tr>
<tr>
<td>$-0.46$ (nonprecipitating)</td>
<td>Marine warm</td>
<td>All</td>
<td>Global</td>
<td>Multiyear</td>
<td>Satellite (Chen et al. 2014)</td>
</tr>
<tr>
<td>$-0.67$ (precipitating)</td>
<td>Stratocumulus over ship tracks</td>
<td>All</td>
<td>Ocean</td>
<td>Multiyear</td>
<td>Satellite (Goren and Rosenfeld 2012, 2015)</td>
</tr>
<tr>
<td>$0.27 \pm 0.10$</td>
<td>Cirrus</td>
<td>All</td>
<td>Global</td>
<td>Multiyear</td>
<td>GCM (Boucher et al. 2013)</td>
</tr>
<tr>
<td>$-1.14 \pm 0.39$ (SW)</td>
<td>Cirrus</td>
<td>BC</td>
<td>Global</td>
<td>Multiyear</td>
<td>GCM (X. Liu et al. 2009)</td>
</tr>
<tr>
<td>$+1.67 \pm 0.11$ [longwave (LW)]</td>
<td>Cirrus</td>
<td>BC</td>
<td>Global</td>
<td>Multiyear</td>
<td>GCM (Gettelman et al. 2012)</td>
</tr>
<tr>
<td>$+0.06$</td>
<td>Cirrus</td>
<td>Dust</td>
<td>Global</td>
<td>Multiyear</td>
<td>GCM (Liu et al. 2012)</td>
</tr>
<tr>
<td>$-0.47 \pm 0.20$</td>
<td>All</td>
<td>BC</td>
<td>Global</td>
<td>Multiyear</td>
<td>GCM/CTM (Bond et al. 2013)</td>
</tr>
<tr>
<td>$+29.3$</td>
<td>DCC</td>
<td>All</td>
<td>U.S. SGP</td>
<td>10 yr</td>
<td>Surface + satellite (Yan et al. 2014)</td>
</tr>
<tr>
<td>$+2.1 \text{ to } 3.6$</td>
<td>DCC</td>
<td>All</td>
<td>SE China</td>
<td>1 day</td>
<td>CRM (Fan et al. 2012a)</td>
</tr>
<tr>
<td>$-0.3$</td>
<td>DCC</td>
<td>All</td>
<td>Tropical ocean</td>
<td>100 days</td>
<td>CRM (Khairoutdinov and Yang 2013)</td>
</tr>
<tr>
<td>$-1.0$</td>
<td>All</td>
<td>All</td>
<td>NW Pacific</td>
<td>2 months</td>
<td>CRM (Y. Wang et al. 2014c)</td>
</tr>
<tr>
<td>$-1.9 \text{ to } -3.7$</td>
<td>All</td>
<td>All</td>
<td>Three regions</td>
<td>1 month</td>
<td>CRM (Fan et al. 2013)</td>
</tr>
<tr>
<td>Land: $-93.8$ (SW), $+27.2$ (LW)</td>
<td>DCC</td>
<td>All</td>
<td>Tropics</td>
<td>4 yr</td>
<td>Satellite (Peng et al. 2016a)</td>
</tr>
<tr>
<td>Ocean: $-14.2$ (SW), $+4.9$ (LW)</td>
<td>All</td>
<td>All</td>
<td>South Asia</td>
<td>3 months</td>
<td>Field (Ramanathan et al. 2001)</td>
</tr>
<tr>
<td>$-0.5$</td>
<td>All</td>
<td>All</td>
<td>Arctic</td>
<td>Daytime</td>
<td>Surface (Zhao and Garrett 2015)</td>
</tr>
<tr>
<td>$-5.0$ (summer)$^a$</td>
<td>Mixed phased</td>
<td>All</td>
<td>Arctic</td>
<td>4 yr</td>
<td>Surface (Zhao and Garrett 2015)</td>
</tr>
<tr>
<td>$+12.2$ (winter)$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ The forcing was estimated at surface, not at TOA.
concentration and to constrain aerosol lifetime effects in GCMs. By contrasting the satellite-observed cloud radiative effects between closed cells and open cells near the ship tracks, Goren and Rosenfeld (2012, 2015) pointed out that as a result of enhancing cloud albedo the aerosol local radiative effect on maritime stratocumulus clouds can be larger than 100 W m$^{-2}$, while the magnitude depends on the region and season. For global maritime clouds, multiple modern satellite measurements estimated that the intrinsic aerosol–cloud forcing is $-0.46$ and $-0.67$ W m$^{-2}$ for nonprecipitating and precipitating clouds, respectively (Chen et al. 2014).

Global-mean aerosol indirect forcing through cirrus clouds is considered for the first time in the recent IPCC assessment report (Boucher et al. 2013). The total AIE on cirrus clouds was estimated to be $0.27 \pm 0.10$ W m$^{-2}$ from two GCMs (CAM5 and ECHAM5). X. Liu et al. (2009) estimated the changes in cloud forcing from anthropogenic aerosol effects on cirrus clouds and found that anthropogenic soot, which is assumed to be an efficient IN for heterogeneous ice nucleation, changes the shortwave and longwave cloud radiative forcing by $-1.14 \pm 0.39$ (cooling) and $1.67 \pm 0.11$ W m$^{-2}$ (warming), respectively, as a result of an increase in cloud ice number from pre-industry (PI) to present days (PD). Conversely, the estimated soot indirect forcing through cirrus clouds is much lower ($-0.06$ W m$^{-2}$) when the soot nucleation efficiencies are prescribed to be within the range of recent laboratory data (Gettelman et al. 2012). Liu et al. (2012) showed that dust IN also had a significant global impact on net cloud forcing of $-0.40 \pm 0.20$ W m$^{-2}$ when comparing simulations with and without dust IN effects on cirrus clouds and when there is sufficient number of dust IN (200 L$^{-1}$) in the upper troposphere. However, the range of cloud radiative forcing induced by dust IN at TOA can be wide (from $-0.24$ to $-1.59$ W m$^{-2}$) depending on different subgrid temperature formulas, different dust IN efficiencies, and clear-sky longwave feedback due to water vapor (M. Wang et al. 2014).

Absorbing aerosols such as black carbon (BC) particularly complicate the aerosol forcing assessment. Bond et al. (2013) concluded that BC-induced globally averaged cloud forcing is from $-0.47$ to $1.0$ W m$^{-2}$, which is even wider than that of the total aerosol indirect forcings. Globally, the radiative heating induced by absorbing aerosols can modulate the environmental relative humidity and alter cloud cover. GCM simulations by Allen and Sherwood (2010) suggested that the reduced midlevel clouds by BC may contribute to a semidirect forcing of 0.5 W m$^{-2}$, while some other modeling studies reported the negative semidirect forcing of BC due to the loss of high clouds (Penner et al. 2003; Koch and Del Genio 2010).

Regionally, cloud radiative forcings induced by aerosols can be much larger than global means. Fan et al. (2012b) suggested the aerosol invigorated effects on a deep convective cloud system that occur in the late afternoon over southeast China produced up to $+5.6$ W m$^{-2}$ warming at TOA because of the enhanced longwave radiation trapped by the increased high clouds. Similarly, Y. Wang et al. (2014b,c) quantified the aerosol-induced longwave cloud radiative forcing enhancement as being from +0.4 to 1.3 W m$^{-2}$ over the northwest Pacific with invigorated winter storms on the basis of the WRF and MMF model simulations. Those changes in the radiation budget through convective clouds are not considered by conventional GCMs and are missing in the current IPCC assessment report. Averaged over the 1-month period for three different regions (Fan et al. 2013), aerosol indirect effects produce strong surface cooling (from $-5$ to $-8$ W m$^{-2}$, monthly average) owing to the increased cloud-top height, cloud cover, and cloud thickness in the daytime. Over tropical oceanic regions under the radiative–convective equilibrium condition, the aerosol indirect effect is estimated to be only about 0.5 W m$^{-2}$, which is mediated by the interactive sea surface temperature (Khairoutdinov and Yang 2013). Many observations also focused on regional AIE. Over the tropics, Peng et al. (2016a) found significant enhancement in both shortwave and longwave cloud radiative forcing due to aerosol perturbations. Yuan et al. (2011b) used the long-term A-Train satellite observations to show the strong shortwave radiative forcing induced by the volcanic aerosols as a result of aerosol–trade cumulus cloud interactions based on a natural experiment. Using long-term measurements over the Southern Great Plains (SGP) site, Yan et al. (2014) for the first time investigated aerosol mediated cloud radiative forcing from about 300 DCC systems and suggested that the daily mean aerosol-induced cloud radiative forcings are $29.3$ W m$^{-2}$ at the TOA and $22.2$ W m$^{-2}$ at the surface owing to the expansion and thinning of anvils under the influence of aerosols. Over South Asia, the aerosol indirect forcing at the TOA was estimated to be $-5.0$ W m$^{-2}$ from the Indian Ocean Experiment (INDOEX) field campaign (Ramanathan et al. 2001). Over the Arctic, haze aerosols induce $+12.2$ W m$^{-2}$ net surface warming in winter and spring by altering ice clouds, while aerosols produce a net surface cooling of $-11.8$ W m$^{-2}$ during the summer by modifying liquid clouds (Zhao and Garrett 2015).

b. Precipitation

The reduction of drizzle with shallow warm clouds by aerosols due to either suppressed warm-cloud microphysics or aerosol dimming effects has been widely reported (Tao et al. 2012), even though giant CCN from sea salt may contribute to the formation of the
embryonic rain drops (L’Ecuyer et al. 2009; Yuan et al. 2008). The sign and magnitude of the precipitation change under different aerosol concentrations vary significantly. Tao et al. (2012) summarized a wide range of precipitation changes under different CCN concentrations, from $-89\%$ decrease in precipitation with a winter storm by elevating CCN 12 times (Teller and Levin 2006) to $+700\%$ increase in precipitation with a tropical storm by elevating CCN 15 times (Wang 2005). There has been more evidence recently that the total precipitation over a large analysis domain is less sensitive to the aerosol perturbation because of the buffering effects from the cloud microphysics (Li et al. 2008b; Fan et al. 2013), cloud dynamics (Stevens and Feingold 2009), and convective–radiative quasi-equilibrium relationship (Grabowski and Morrison 2011).

Even with the insensitive total precipitation, the probability distribution function of precipitation intensity as well as the local spatial patterns of precipitation can be dramatically modulated by aerosols. By suppressing light precipitation, an increase in CCN concentration may enhance heavy precipitation resulting from enhanced mixed-phase and ice cloud microphysics. Such an aerosol-induced distribution shift of precipitation intensity resulting in more extreme precipitation events has been widely reported from previous observational and modeling studies (e.g., Qian et al. 2009; Li et al. 2011; Y. Wang et al. 2011; Fan et al. 2012a; Tao et al. 2012; Koren et al. 2012; Guo et al. 2014). Even for winter storms, more snow and less light rain were found over the continental United States when CCN concentrations were elevated (Thompson and Eidhammer 2014). Using multiple years of TRMM and MODIS data, Yuan et al. (2011a) showed that the extreme precipitation intensity of tropical convection in the Pacific warm pool region is increased as a result of increase of sulfate concentration. Koren et al. (2012) further argued that the intensification of rain rates by aerosols could be found over both the ocean and land and from tropics to midlatitudes based on the TRMM and MODIS measurements. Fan et al. (2013) showed that monthly CRM simulations for three different regions during the summertime showed the suppression of light rain by increasing CCN consistently, but the enhanced heavy rain in regions of southeastern China and the TWP. In that study, heavy rain was not enhanced over the SGP presumably as a result of cold cloud base and high wind shear.

Anthropogenic aerosols either produced locally or transported from remote areas are linked to the modulation of the regional hydrological cycle. A few studies (e.g., Y. Liu et al. 2009; Gu et al. 2006; Jiang et al. 2013) suggested that the observed tendency of “southern flood and northern drought” over east China during the weakened East Asian summer monsoon (EASM) was mainly due to ARI. Similarly in South Asia, the increased man-made aerosols from India were found to be responsible for the observed reduction of the precipitation of about $-0.95$ mm day$^{-1}$ (50 yr)$^{-1}$ (Bollasina et al. 2011). Some studies also suggest that the absorbing aerosols over the Indio-Gangetic Plain could lead to an earlier onset of South and East Asian monsoons and increased precipitation over the Indian subcontinent (Lau et al. 2006; Lau and Kim 2006). Fan et al. (2015a) proposed a mechanism on how absorbing aerosols redistributed moisture and energy spatially and significantly enhanced the extreme precipitation in a catastrophic flooding event in Southwest China. Over the North Pacific, the long-term measurements from Global Precipitation Climatology Project (GPCP) (Li et al. 2008a) showed a significant trend of increased wintertime precipitation about 1.5 mm yr$^{-1}$ from 1984 to 2005, and the trend was attributed to the aerosol invigoration effect of the pollution outflow from the Asian developing countries (R. Zhang et al. 2007; Y. Wang et al. 2014b,c). Using WRF-Chem simulations, Wu et al. (2013) pointed out that aerosol effects can either enhance or reduce EASM precipitation depending on the relative locations of aerosols and monsoonal clouds in different phases of the summer monsoon. Using the CESM model, Hu and Liu (2013) found that a simulation with anthropogenic aerosols could reproduce the observed trend of spring precipitation decrease over southern China since the 1950s by changing the atmospheric circulation. The effect of greenhouse gases on the spring precipitation reduction is much smaller in that study.

c. Severe weather systems

1) SUPERCELLS, HAILSTORMS, AND TROPICAL CYCLONES

In the recent years, there are an increasing number of studies with their focus on aerosol impacts on severe weather systems. A few observational studies indicated aerosols enhance the occurrence and strength of supercells and hailstorms (Bell et al. 2008; Rosenfeld and Bell 2011; Saide et al. 2015; Yang and Li 2014). However, it is highly debatable in which way and to what extend aerosols can module these severe systems (e.g., Yuter et al. 2013), considering the current observations have difficulties in isolating aerosol effects from any others, such as the perturbing initial conditions in a chaotic system since a small variation in dynamics and thermodynamics may overpower aerosol effects. It is also challenging to tease out ACI from the inherent variability of dynamics and the covariability of aerosol and
dynamic fields in the real atmosphere based on observations only. Therefore, here we mainly discuss the most recent numerical studies that strictly controlled the initial conditions for storm development in their sensitivity experiments. There are many more studies needed for this area.

Supercell thunderstorms and squall lines are common deep convective systems, and they exhibit various sensitivities to CCN perturbations. Previous studies (G. Li et al. 2009; Mansell and Ziegler 2013) simulated the CCN effects on squall lines using CRMs with two-moment bulk cloud microphysics, and they found the aerosol invigoration effect on multicell storms in terms of a 13% increase in domainwide precipitation and a 40% enhancement of the peak updraft velocity from low to moderate CCN concentrations. In contrast, Khain et al. (2009) performed a CRM study of a squall line with the bin cloud microphysics but did not find a significant dependence of accumulated rain and convection strength to aerosols. Lebo and Seinfeld (2011) showed an aerosol-induced increase in updraft velocities and rainfall intensity of a supercell but a decrease in cumulative precipitation due to the larger condensate aloft under polluted conditions.

Numerical studies have shown that aerosols could affect the severity of hailstorms, but the magnitude of influence also depends on the characteristic of the storms. Noppel et al. (2010) used a CRM with a bulk cloud microphysics scheme to simulate a summertime hailstorm that caused severe damage in Germany. Sensitivity tests with different initial CCN concentrations showed that changing CCN characteristics could modify the hailstorm, but they can either enhance or suppress hail amount and hailstone size depending on the specific storm dynamical and microphysical processes. The same hailstorm case was investigated by Khain et al. (2011) using spectral-bin microphysics (SBM), and they showed that altering CCN concentration from 100 to 3000 cm\(^{-3}\) resulted in a 10 times increase in the size and mass concentration of hail, consequently increasing the surface precipitation. With a new version of SBM considering the wet growth of hail, Khain et al. (2016) simulated a hailstone size of 5 cm, in agreement with the observations. Carrió et al. (2014) also examined the responses of hailstorms to enhanced CCN concentration using the RAMS model, and they found CCN effects depend on the cloud-base height of storms. Khain et al. (2015) summarized that different microphysical parameterizations could lead to very different hail properties and aerosol impacts on hailstone amount and size. The hail growth mechanism with SBM is largely the accretion of supercooled water content in the area of cloud updraft, while hail growth is due mostly to freezing raindrops just above the freezing level in the two-moment bulk scheme. Loftus et al. (2014) and Loftus and Cotton (2014) found that a three-moment hail scheme seemed to improve the simulation of hailstone properties significantly compared with the two-moment scheme. As a destructive convection system over land, tornadoes are also found to be impacted by atmospheric aerosols under idealized dynamical conditions. Using RAMS, Lerach et al. (2008) reported that high concentrations of CCN and giant CCN (GCCN) under the polluted condition resulted in reduced warm and cold precipitation and a weak evaporative cooling with a longer-lived supercell, which is favorable for tornado genesis. Similarly, recent real-case numerical experiments by Saide et al. (2015) showed that smoke from Central American fires could enhance the probability of tornado genesis and tornado intensity and longevity through both direct and indirect effects.

The frequency and intensity of tropical cyclones (TCs) are regulated by several environmental factors, such as sea surface temperature (SST), vertical wind shear, vorticity, and humidity of the free troposphere (Emanuel 2013). By reducing the radiation reaching the sea surface and altering latent heat release within tropical cyclones, aerosols could also perturb TC genesis and development, even though aerosol effects should be secondary compared to the determinative dynamic factors. Previous studies show that the intensity of tropical cyclones in numerical models depends strongly on the microphysical schemes used, as reviewed by Khain et al. (2015). Aerosols can modify the thermodynamic and microphysical conditions of tropical cyclones (Rosenfeld et al. 2012; Y. Wang et al. 2014a). Increasing CCN concentrations at the tropical cyclone periphery can lead to weakening of tropical cyclones (H. Zhang et al. 2007, 2009; Rosenfeld et al. 2012; Cotton et al. 2012; Hazra et al. 2013; Lynn et al. 2016), while an increase of CCN at the inner tropical cyclone core can lead to intensification of tropical cyclones (Herbener et al. 2014; Lynn et al. 2016; Khain et al. 2016). Y. Wang et al. (2014a) illustrated the anthropogenic aerosol effect on tropical cyclones through both the radiative and microphysical effects of aerosols and discovered that the combined microphysical and radiative effects of anthropogenic aerosols delayed tropical cyclone development, weakened minimal surface pressure and maximal wind speed near the eyewall, and led to earlier dissipation; however, rainbands were enlarged that increased total precipitation (Fig. 3). Based on a real-case simulation of Hurricane Irene, Khain et al. (2016) found dust aerosols penetrated the eyewall when Irene crossed the band of Saharan aerosols and intensified Irene’s development by altering CCN concentrations.
It is noteworthy that numerous studies showed that Saharan dust exhibit large influences on the genesis and intensification of Atlantic tropical cyclones by modulating cloud hydrometeor contents, diabetic heating distribution, and thermodynamic structure of tropical cyclones, but mainly through dust radiative effects (e.g., Dunion and Velden 2004; Sun et al. 2009; S.-H. Chen et al. 2010, 2015).

2) LIGHTNING

The lightning process depends on the existence of supercooled water, ice crystals, snow, graupel, and hail, and cloud microphysical process in thunderstorms such as diffusional growth may have a large impact on charge separation (Williams et al. 1991). The lightning enhancement by aerosols is supported by various satellite and in situ measurements. The higher frequency of lightning activities over land than that over ocean can be partially caused by the land–sea contrast of aerosol concentrations (Seinfeld and Pandis 2006). Over megacities in southern Brazil, South Korea, and southern China, lightning density were positively correlated with the measured particulate matter (PM) concentrations (Naccarato et al. 2003; Kar et al. 2009; Y. Wang et al. 2011). Lightning enhancement induced by aerosols typically occurs over continental areas. However, Yuan et al. (2011b) proposed that the high aerosol loading from volcanic activity was responsible for the observed anomalously high lightning frequency over the west Pacific Ocean in 2005. The enhancement rate of lightning flashes can be as large as 30 times per unit of aerosol optical depth (Yuan et al. 2012). Combining the long-term measurements of lightning and precipitation from TRMM satellite with surface measurements of visibility, Yang and Li (2014) also revealed the enhanced lightning activity and invigorated thunderstorms in southeast China under polluted conditions. In central China, however, thunderstorm activity has decreased by nearly half in the past 50 years during which aerosol loading has increased dramatically (X. Yang et al. 2013a). The group attributed the different trends to be partially caused different aerosol types: absorbing aerosols that are dominant in central China and suppress convection through aerosol–radiation interaction while sulfates are more significantly present in southeast China.

The cause and effect between aerosols and lightning are difficult to establish merely using observations (Williams et al. 2002) because the electrical parameters are also correlated with other meteorological factors such as buoyancy that generates supercooled water required for lightning. So, doubts are cast on the observational studies showing enhanced lightning by aerosols (Williams and Mareev 2014) including the weekday effects on lightning due to more anthropogenic aerosols revealed by Bell et al. (2009). However, a couple of recent observational studies claimed to avoid the covariability of aerosols and meteorological fields and examined the impact of volcanic aerosols on lightning (Yuan et al. 2011a, 2012). They found a remarkable increase of lightning flash rate due to increased aerosol loading (~30 times or more per unit of aerosol optical depth). Numerical models can better disentangle the convolution between aerosol and meteorology. Using a cloud-resolving version of the WRF Model with an aerosol scheme and a lightning parameterization, Y. Wang et al. (2011) suggested that both total precipitation and lightning potential are enhanced by about
16% and 50%, respectively, owing to elevated aerosol loading over a megacity in south China. Similarly, Mansell and Ziegler (2013) performed cloud-resolving simulations and found that CCN strongly affects the microphysical and electrical evolution of a multicell storm. The parameterization of the ice multiplication process was critical for the CCN–lightning relationship in the cloud microphysics scheme they used. Using WRF Model with spectral-bin microphysics and lightning potential prediction schemes, Lynn et al. (2012) and Khain et al. (2008a, 2010) demonstrated that small continental aerosols fostered the lightning formation at the periphery of tropical cyclones through aerosol thermodynamic invigoration.

d. Large-scale circulations

ACI alters large-scale circulations by affecting the radiation budget and inducing regional energy imbalances, as indicated by many GCM studies. There is a consensus that the indirect forcing of aerosols accounts for more than 70% of total aerosol forcing and is predominant in shaping the meridional energy distributions and modulating the circulation systems on the global scale (Ming and Ramaswamy 2011; Booth et al. 2012; Wang et al. 2015). Coupled GCM simulations with explicit aerosol effects on stratiform clouds (Ming and Ramaswamy 2011) suggested that the large radiative cooling mainly induced by ACI in the Northern Hemisphere weakens the northern branch of the Hadley circulation. The interhemispheric asymmetry in aerosol distribution leads to a northward cross-equatorial energy flux that compensates for the energy deficit in the Northern Hemisphere. Similarly, the recent CMIP5 models predicted that the tropical rainfall pattern mediated by the SST change experienced a southward shift over the ITCZ under the influence of aerosols in the Northern Hemisphere (Xie et al. 2013). A recent modeling study by Wang et al. (2015) examined the response of large-scale circulations to the shift in maximum pollution from the United States and Europe to Asia since the 1970s. A reduced meridional streamfunction and zonal winds over the tropics as well as a poleward shift of the jet stream in the present-day aerosol conditions suggests weakened and expanded tropical circulations under the influence of the altered cloud radiative forcing induced by redistributed aerosols.

Regionally, the major monsoon systems are subject to influences by either anthropogenic pollution or mineral dust. Over South and East Asia, the impacts of anthropogenic aerosols on the long-term variations in atmospheric radiation budget, surface temperatures, and regional circulations associated with the Asian summer monsoon have been studied extensively (e.g., Li et al. 2016, manuscript submitted to Rev. Geophys.; Menon et al. 2002; Lau et al. 2006; Ramanathan and Carmichael 2008; Ganguly et al. 2012). For example, Bollasina et al. (2011) concluded that recent widespread drying in South Asia is an outcome of a slowdown of the tropical meridional overturning circulation, which can be attributed mainly to anthropogenic aerosol emissions. They further suggested that the local pollution is responsible for the earlier onset of the Indian monsoon through the dynamical feedbacks and regional land surface processes in the aerosol–monsoon interaction (Bollasina et al. 2013). Song et al. (2014) teased out the aerosol forcing from greenhouse gas forcing in the recent CMIP5 simulations and revealed that atmospheric aerosols play a pivotal role in driving the weakened low-level monsoon circulation by decreasing the land–sea thermal contrast during summer seasons. A recent review paper summarized that cloud physical properties and precipitation are significantly affected by aerosols in China with aerosols likely suppressing local light and moderate rainfall but intensifying heavy rainfall in southeast coastal regions, and the detailed mechanisms behind this pattern still need further exploration (Wu et al. 2016).

Over the northwest Pacific, there is a unique coupling between extratropical cyclones and Asian pollution outflows during the winter. Through interacting with the maritime clouds and precipitation systems, Asian pollution outflows exert potentially great impacts on regional climate and global circulations (R. Zhang et al. 2007). Y. Wang et al. (2014c) developed a modeling approach to upscale the regionally simulated aerosol forcing to global simulations. By prescribing aerosol-induced cloud radiative forcing anomalies calculated from CRM simulations over the northwest Pacific, the CAM5 simulations predicted the enhanced winter storm intensity due to the Asian pollution outflows, which is consistent with the storm intensity change based on 30 years of reanalysis data. This is consistent with Y. Wang et al. (2014b), who used the MMF model to assess the impact of Asian pollution on the Pacific storm track.

4. Discussion of issues and future research directions

a. Parameterization issues

Although the performance of simulated aerosol fields have improved (Penner et al. 2006), there are still large variations in simulated convective and cloud properties among models, even at the CRM scale (Fridlind et al. 2012; Varble et al. 2011, 2014a,b). Cloud microphysical parameterizations vary from a single moment to bin approaches, producing large differences in simulated convection and clouds even under the same initial dynamic and thermodynamic conditions (X. Li et al. 2009;
Khain et al. 2009, 2015; Fan et al. 2012a; Wang et al. 2013). Khain et al. (2015) described how bin and bulk parameterizations represent the major microphysical processes and their limitations. They concluded that most of bulk schemes are not well configured for studying ACI, since they generally 1) do not include a CCN budget, which can cause unrealistic results for aerosol impact as shown in Fan et al. (2012a); 2) do not explicitly calculate diffusional growth based on supersaturation and droplet sizes — instead, they employ the saturation adjustment approach, which eliminates supersaturation and decreases the sensitivity of bulk schemes to aerosols; 3) employ autoconversion parameterizations that were generally developed under a narrow range of conditions and do not take into account the time evolution of autoconversion to convert cloud water to rainwater; 4) use average fall velocities for collision processes of hydrometeors, which is a big problem for self-collections (because there is no difference of fall velocities for the same hydrometer) and for collisions between different hydrometeor types with the similar average fall velocities; 5) use average fall velocities over the particle size distribution for sedimentation, which does not account for smaller particles that fall slower and larger particles that fall faster; and 6) use two sets of averaged fall velocities, mass-mean and number-mean fall velocities, in two-moment schemes that would result in cloud area with significant mass but negligible number or with significant number but negligible mass as shown in Fan et al. (2015b).

In addition to above limitations with bulk parameterizations, the mean quantities such as mean particle size and fall velocity over a size distribution employed have much less sensitivity to aerosol changes (Fan et al. 2013). Therefore, in many aspects, bulk schemes are not well designed for studying aerosol impacts. For cloud simulations under relatively clean conditions and especially with strong dynamic forcing such as large squall lines (Tao et al. 2016; Morrison et al. 2015), bulk schemes (especially two- or three-moment versions) may perform reasonably well. However, they often give qualitatively different responses to the increase of CCN compared with bin models (e.g., X. Li et al. 2009; Khain et al. 2009, 2015; Fan et al. 2012a, 2013). Besides the different responses to aerosols in convection and precipitation as shown in many previous studies, recent work also revealed that two-moment bulk schemes did not simulate the same aerosol microphysical effects as bin models and observations revealed, such as aerosols significantly increasing the CTH and cloud cover for deep convective clouds (van den Heever et al. 2011; Khairoutdinov and Yang 2013; Fan et al. 2013). The main reason is that the bulk scheme does not simulate the much reduced ice particle size and fall velocity in stratiform/anvil clouds under higher CCN conditions. Many of the above-mentioned limitations can be addressed by improving bulk schemes such as adding CCN budgets, using explicit calculation of diffusion growth, and adding additional moments to better represent hydrometeor size distributions.

Conversely, bin parameterizations have their own problems. They are very computationally expensive; therefore, studies that employ bin-microphysics parameterizations usually use relatively small domains for short time periods. The accuracy of bin schemes is also limited by our theoretical understanding of cloud microphysics. This includes 1) collision–coalescence and collision–conversion processes among hydrometeor particles, especially under a turbulent environment; 2) ice nucleation processes, especially for deep convective clouds; and 3) ice diffusion growth for different ice crystal shapes. There are large uncertainties for many physical parameters in bin models such as condensation coefficient, collision efficiencies, rate of riming, fall velocities of different hydrometeors, and scavenging efficiency of interstitial aerosols. Numerical issues such as spectrum broadening are another concern. As a result of all of these uncertainties, using bin models sometimes may not be beneficial, especially for the cases in which aerosol size distribution and composition are unknown, because the uncertainties due to aerosol properties may be larger than the uncertainties due to the choice of microphysical scheme. Nevertheless, bin parameterizations are physically more realistic. With good constraints obtained from observations of cloud microstructure and aerosols, it is possible to reduce the simulation uncertainty associated with the model treatment of microphysics. Also, as discussed above, bulk schemes are not well configured to respond to aerosols. Even with a specially configured two-moment bulk scheme such as the incorporation of physically based aerosol activation and bin-resolved aerosol population (Lebo and Seinfeld 2011), little change of latent heat is shown from the low-CCN to high-CCN cloud conditions while a bin scheme showed a significant increase of latent heat (Lebo and Seinfeld 2011).

A major issue in current regional and global climate models (RGCMs) is how to accurately parameterize subgrid clouds at various resolutions. Aerosol interactions with those subgrid clouds have not been a consideration for the majority of the parameterizations used in RGCMs. For subgrid parameterizations that consider some types of aerosol impacts, the performance in simulating cloud properties and precipitation was improved significantly (Song and Zhang 2011; Song et al. 2012; Lim et al. 2014; Grell and Freitas 2014; Berg et al. 2015). While there are many issues with subgrid cloud parameterizations need to be resolved, new parameterizations should be scale aware.
and aerosol aware (i.e., consider ACI) to reduce ACI uncertainties of RGCM’s simulations.

b. Measurement issues

The effort toward reducing uncertainties of aerosol–cloud interactions is limited by our physical understanding, which is limited by the available measurement data. Many key measurements for better understanding aerosol, convection, and cloud properties are either lacking or too sparse. We do not have concurrent in-cloud measurements of cloud dynamics, microphysics, and aerosols for deep convective clouds. It has been a challenge to obtain in-cloud vertical velocity for convective clouds. As we know, with reliable convective vertical velocity and CCN data, the aerosol thermodynamic invigoration hypothesis can be directly evaluated with observations. It is very important to directly validate this hypothesis, especially considering that model predictions with different microphysical parameterizations are not consistent. For recent field campaigns, such as the Tropical Warm Pool–International Cloud Experiment (TWP-ICE) and the Midlatitude Continental Convective Clouds Experiment (MC3E), the 3D wind fields within clouds have been retrieved from multi-Doppler radars (Collins et al. 2013; Fan et al. 2015b), which is very useful to better understand cloud dynamics and the feedbacks between microphysics and dynamics as well as evaluate model performance (Fig. 4a). However, because of the lack of corresponding in-cloud microphysical quantities such as mass and number mixing ratios of hydrometers (Fig. 4b), we are not able to fully understand the feedbacks between dynamics and microphysics. Nor are we able to determine whether the larger graupel predicted by the bulk schemes compared with the bin schemes is reasonable or whether the larger graupel is also a reason for the stronger vertical velocities in the middle and upper troposphere.

As discussed in section 2 for aerosol–DCC interactions, the aerosol microphysical effect that led to increased cloud cover and CTH proposed by Fan et al. (2013) is produced only when spectral-bin microphysics is used. Simulations that use bulk schemes do not simulate this effect (Fan et al. 2013; van den Heever et al. 2011; Khairoutdinov and Yang 2013). Validation of this mechanism requires statistical measurements of hydrometeor size and fall velocity at the upper levels of deep convective clouds developed under low- and high-CCN conditions that are not yet available.

To single out aerosol impacts from meteorological factors, measurements of aerosol properties, CCN, cloud droplet size distribution, and updrafts at cloud base are critical. Those data can be obtained through aircraft measurements. However, these types of measurements have been obtained mainly for clouds with weak convection, which is not statistically representative for all cloud types. A new way to disentangle effects of CCN and cloud-base updrafts on convective cloud properties emerged with the advent of the height resolution (375 m) retrievals of cloud properties with the Suomi National Polar-Orbiting Partnership (SNPP) satellite. The retrieved cloud-base height (Zheng et al. 2015; Zhu et al. 2015) and particle size (Rosenfeld et al. 2014b) make it possible to retrieve cloud-base drop concentrations of convective clouds (Fig. 5a) (Rosenfeld et al. 2014c). Cloud-base updrafts can be retrieved based on the cloud-base height (Fig. 5b) (Zheng and Rosenfeld 2015; Zheng et al. 2015). Combining cloud-base updraft and drop concentrations allows retrieving cloud-base supersaturation and CCN concentration (Rosenfeld et al. 2016). Such retrieved data have been validated by aircraft measurements over SGP, and they can be very useful in terms of disentangling effects of CCN and updrafts and obtaining statistical results regionally and even globally.

For MCSs, which are fed by deep layers of air, the biggest problem is that we do not have vertical profiles of aerosols over wide regions. Surface measurements are inadequate to determine the ingestion of aerosols. Aircraft measurements are not frequent enough in time or continuous enough in 3D space to be sufficient in determining how aerosols enter clouds. Satellite measurements of aerosols are too contaminated by clouds and have many other limitations as reviewed by Z. Li et al. (2009). Establishing a platform to measure aerosol properties vertically over a wide region is imperative to help understand aerosol effects on MCSs. For severe storms, lack of robust measurements of storm dynamics and microphysics impedes our further understanding.

Measurements of small-scale (local) cloud particle size distribution are lacking, constituting a big obstacle to improve model representation of cloud microphysics, since hydrometeor size distribution is the backbone to all microphysical processes (Khain et al. 2015). Mean cloud particle size distribution averaged over a large cloud volume is not a very useful property, because it smooths out the differences of size distribution at different locations of clouds. Cloud drop size distributions at cloud bases should be very different from those at the higher altitudes where rain is initiated and those at cloud edges due to mixing processes.

Ice nucleation parameterizations were generally developed from field measurements of wave clouds or stratiform/cirrus clouds and/or laboratory experiments data. The applicability of such parameterizations to various cloud types in real nature has been questionable,
particularly for DCCs. The recent field campaign measurements for cumulonimbus clouds over a tropical region suggested ice formation is mainly secondary, which is different from other types of clouds that primary ice nucleation dominates (Lawson et al. 2015). Observation on ice nucleation should be very dependent on many factors such as temperature, supersaturation, aerosol properties, and even the presence of other hydrometeors. The variability of each factor is so large in real environments; it is therefore very challenging to develop a unified parameterization applicable for all conditions.

c. Future directions

1) MULTISCALE CONCURRENT MEASUREMENTS OF CONVECTION, CLOUD, AND AEROSOL PROPERTIES

Concurrent measurements of CCN and cloud microphysical and dynamic properties over a range of temporal and spatial scales collected over many regions would provide a wealth of data needed make a leap in our understanding in this field. As discussed previously, concurrent measurements of convective intensity and cloud microphysical properties at very small time scales is very important to improve our understanding of the relationships between dynamics and microphysics. The box closure experiments conducted at various climate regions proposed by Rosenfeld et al. (2014a) through integration of satellite, in situ aircraft, and ground-based remote sensing observations provide a strategy to obtain concurrent data. The box closure approach is not only a good way to quantify the changes of matter and energy in all their relevant forms in the climate system due to aerosol perturbation but also provides unprecedented data for evaluation and development of models and parameterizations. However, execution of such a box closure experiment is huge effort and investment and requires a cooperation of highly coordinated groups. The U.S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM)

![Fig. 4](image-url)
program supersite concept provides an applicable approach to obtain such concurrent measurements to improve understanding of ACI processes and quantify aerosol impacts. However, the current supersites of ARM are limited to SGP and North Slope of Alaska (NSA), which may not be ideal locations to study aerosol–convective cloud interactions. The concurrent satellite measurements of CCN, cloud-base updrafts, and microstructure (Rosenfeld et al. 2016) also provide valuable datasets to quantify some of aerosol effects.

2) IMPROVE UNDERSTANDING OF HYDROMETER SIZE DISTRIBUTION AND CONVERSIONS

The majority of aerosol and cloud physical properties are determined by particle size distribution and the various conversions among them. Particle size distribution of aerosols and their activation to become cloud droplets and ice crystals are the starting point of ACI. Therefore, processes that significantly change aerosol sources and sinks need to be understood clearly. The review is not meant to give a detailed review on issues with chemistry and aerosols, but we would point out that the improved understanding in the following areas is imperative to reduce uncertainties of model simulation of aerosol properties and ACI: 1) quantifying emissions especially biogenic, dust, and biomass burning, 2) identifying the chemistry mechanisms for forming new particles especially secondary organic aerosol (SOA) over both land and ocean, 3) understanding how biogenic SOA interacts with anthropogenic aerosols and changes properties to impact CCN, and 4) understanding how naturally emitted dust and biological particles interact with anthropogenic emissions and change their IN ability.

As summarized by Khain et al. (2015), the performances of models in simulating various types of cloud regimes are strongly dependent on cloud microphysical parameterizations. Because of the predominant role that diabatic processes play in the climate system, improving the accuracy of description of cloud microphysics is the ultimate way to reduce uncertainties of ACI in climate projection. To accurately describe cloud microphysical processes, an accurate simulation of size distribution of various types of hydrometeors is required. However, our understanding of particle size distribution for all types of hydrometeors and the conversions between hydrometeors are limited, especially for mixed- and ice-phase clouds. Past observations show that the small-scale (local) particle size distribution have a complicated shape and often cannot be well approximated by gamma or exponential distributions that are assumed commonly in bulk microphysical parameterizations. Many more observations on particle size distribution under different aerosol scenarios are needed to improve microphysical parameterizations. Other microphysical processes with many unknowns include 1) ice particle shapes and how their shapes impact on depositional growth and riming processes; 2) dominant ice nucleation mechanisms for a specific type of clouds, since each type of clouds may have its own dominant ice formation pathway which could lead to very different sensibilities to aerosols (Fan et al. 2014); 3) conversions of liquid drops and ice crystals to snow, graupel, and hail; 4) particle fall velocity; and 5) impacts of cloud electrification on...
hydrometeor orientation and collision efficiency and on NO₂ and the feedback to aerosols and CCN.

In situ aircraft measurements and ground-based radar and lidar observations would be extremely helpful to gain new knowledge on the areas mentioned above. Regarding the hydrometeor habits and conversions, combining the data from polarimetric radars with bin microphysical models is an effective way to make progress (Kumjian et al. 2012; Kumjian et al. 2014). Besides field measurements, laboratory experiments can also be very helpful to improve our understanding and reduce uncertainty. For example, laboratory experiments on ice nucleation (Durant and Shaw 2005; Niemand et al. 2012), secondary ice formation (Hallett and Mossop 1974), efficiency of collisions, and aerosol scavenging and riming (Mitra et al. 1990, 1992) have substantially contributed to our knowledge of microphysical processes. Many laboratory findings have helped us to improve model representations. For example, the theory of wet hail growth developed recently by Phillips et al. (2014, 2015) in the SBM was based on the laboratory measurements conducted by García-García and List (1992).

3) DEVELOPMENT OF CLOUD MICROPHYSICAL PARAMETERIZATIONS

The majority of the current two-moment schemes are not well configured for ACI as discussed in section 4a. Since increasing moments (Milbrandt and Yau 2006; Loftus and Cotton 2014) or using the bin-emulating approach, which uses look-up tables calculated from the bin model approach (Cotton et al. 2003; Saleeby and van den Heever 2013), improved model performances as shown by those studies, it is promising to further develop bulk schemes along these directions, especially for RGCMS where computation cost is an important concern.

As for the further development of bin microphysics, in LES/CRM, we are limited by our knowledge of size distributions, particle shapes and phases, and their conversions. Microphysical processes have too many unknowns, especially for DCCs. Most of the bin models do not include the representation of cloudborne aerosols and only simple treatment of CCN regeneration was applied in some studies (Fan et al. 2009; Lebo and Seinfeld 2011). Moreover, model evaluation is even difficult as a result of lacking measurements of concurrent convection and cloud properties. On this regard, improving understanding and developing observational data for model evaluation and developments that are discussed in sections 4c(1) and 4c(2) is the key. The polarimetric radar measurements are very useful to look at hydrometeor properties for convective clouds, and satellites can measure cloud drop effective radius of the same clouds. For large-scale models the biggest problem with bin microphysics is the computational cost; therefore, development of new numerical algorithms to substantially reduce the computation cost can increase in the likelihood using such schemes in RGCMS. In addition, reducing the number of particle size distributions or number of bin sizes or using the hybrid moment and bin methods can decrease computational costs, but compromise with reduced accuracies to some extent.

4) CONSTRAINING PARAMETERIZATIONS WITH OBSERVATIONS AT VARIOUS SCALES

Generally, parameterizations in models are developed based on theory and/or measurements valid for small scales and/or for specific situations. We suggest caution should be exercised when a specific parameterization, especially the empirical ones, is generalized to other cases. We recommend applying such parameterizations only when relevant measurements are available to validate them. To reduce confusing results on ACI from case studies, we would also suggest that case simulation be validated with aerosol, cloud, and precipitation properties before any impacts of aerosol variability are examined. In a GCM, parameterizations need to be applicable for any part of the globe, for any climate state, and for coarse spatial and temporal resolution. A widely used methodology to adapt parameterizations to changing environments and scales is to adjust (or “tune”) parameters to get a more “realistic” match to available local data (e.g., Rotstayn 2000). A better method would be to use observational data at adequate scales to infer these parameters. In this regard, DOE ARM supersite data would be valuable. The use of information from satellite retrievals is appealing as well, because they provide more complete cloud, aerosol, and radiation measurements at the large horizontal and temporal scales needed to evaluate GCMs (Lohmann et al. 2007). However, the large uncertainty with satellite retrievals is a big obstacle of this method. This situation is changing with the emergence of substantial improvements in satellite measurements (Rosenfeld et al. 2016).

We believe that physical-based parameterizations can be more easily adapted to different cloud and climate regimes relative to empirical parameterizations. However, empirical parameterizations are widely used in our current models including CRMs such as autoconversion, ice nucleation, droplet freezing, and particle size distribution. Replacing empirical parameterizations with
physically based calculations is an overall direction to reduce model uncertainties.

5. Summary

We have reviewed recent theoretical studies and important mechanisms on aerosol–cloud interactions, as well as the significances of aerosol impacts on radiative forcing and precipitation associated with different cloud types. Critical issues and future directions to make a leap in our understanding of this field in both modeling and observations are discussed. Realizing it is hard to revisit all of the important points laid out throughout the paper, below we only summarize some key points in each aspect.

For warm boundary layer clouds, we have known that ACI is a lot more complicated than the Twomey effect and precipitation suppression due to reduced droplet size. Warm-cloud invigoration with increased liquid water content, taller clouds, and more precipitation could occur when adding aerosols to clouds with very low droplet number concentration. The current focus for this type of cloud is on aerosol impacts on cloud organization. Over the land, aerosols may impact the transition of shallow cumuli to deep convective clouds through modifying surface heating and entrainment processes. However, the impact is complicated by land surface processes and land–atmosphere interactions. Over the ocean, aerosols may enhance the transition from open to closed cells, which could increase the cloud radiative effect by more than 100 W m$^{-2}$ over the Southern Ocean, with only about 1/3 of the effect contributed by the Twomey effect, 1/3 by the LWP effect, and the rest by the cloud cover effect. The cloud cover effect is most significant when adding aerosols to clean background clouds. One of the key questions for quantifying ACI globally is how clean it was at the pre-industrial time (Carslaw et al. 2013).

For mixed-phase stratiform clouds, aerosols can affect cloud phase and lifetime significantly by changing the transformations of water between vapor, liquid, and ice particles, particularly by serving as IN. The mixed-phase stratiform cloud properties at the “dust belt” over the globe are impacted by dust particles through IN effects. However, the significance depends on cloud temperature, large-size dust particle concentration, and dust chemical compositions. Over the Arctic, a self-maintaining feedback between supercooled liquid water, radiation, and turbulence that explains their persistence can be broken down owing to increasing IN. However, large uncertainties exist regarding IN sources and ice nucleation mechanisms.

For DCCs, we understand that the thermodynamic invigoration by aerosols due to more latent heat release as a result of freezing extra cloud water can be significant only when environmental conditions favor it—that is, warm-cloud bases, relatively weak wind shear, and high CAPE. The microphysical invigoration (i.e., increased cloud depth and cloud cover) due to much slower fall velocity of ice particles in polluted clouds is significant over a regional domain and the time scale of the entire cloud life cycle. It is shown that increased cloud depth and cloud cover have very significant radiative effect (up to $-13$ W m$^{-2}$ TOA cooling at the daytime), and GCMs could miss a large part of it owing to poor parameterization of cumulus clouds. It is shown that aerosol effects on DCCs are strongly dependent on different types of convective systems due to different dynamics and microphysics–dynamics interactions. However, since DCCs can also be very sensitive to small perturbations of dynamic and thermodynamic fields, aerosol effects are often being questioned, especially considering that observed aerosol and dynamic fields are often covarying. It is not well established how aerosols impact various MCSs, especially in terms of integrating with the dynamics of MCSs, such as the feeding of the deep layer of air and the interactions of cold pool and low-level wind shear.

For cirrus clouds, it is known that aerosol composition is very important in terms of impacting ice nucleation. However, aerosol impacts are largely determined by the dominant nucleation mechanisms or the balance between homogeneous versus heterogeneous nucleation. There is relatively better understanding for homogeneous nucleation, while for heterogeneous nucleation, many unknowns exist about the concentrations and properties of IN, their dominant modes of action, and competition between the modes. Anvil cirrus can be impacted by aerosols through cloud microphysics processes in convective clouds, and we understand that aerosols, by serving as CCN, can significantly increase ice particle number but reduce ice particle size and fall velocity in DCCs, leading to an increase in anvil cirrus cloud cover and cloud-top height. The total AIE on cirrus clouds was estimated to be $0.27 \pm 0.10$ W m$^{-2}$ from two GCMs (CAM5 and ECHAM5), which did not account for AIE on anvil cirrus through cumulus clouds (of about $-0.45$ W m$^{-2}$ at the SGP).

To reach a good understanding of the role of aerosols in our weather and climate systems, large efforts are still needed. The main issues preventing us from making a leap are as follows.

On observations:

- Lack of concurrent profile measurements of cloud dynamics, microphysics, and aerosols at convective cores of DCCs as well as the time evolution of
environment measurements including aerosol properties near the system.

- Lack of statistical measurements of aerosol properties, CCN, cloud droplet size distribution, and updrafts at cloud base, as well as a platform to measure aerosol properties vertically over a wide region to study aerosol–MCS interactions.
- Lack of measurements of small-scale (local) cloud particle size distribution and the spatial and temporal evolutions of size distribution in a cloud. Lack of robust measurements of IN, mixed-phase properties, and hydrometeor size and fall speed for DCCs.
- Lack of long-term concurrent measurements of aerosol properties (size distribution and composition) and meteorological fields or special field experiments to address the covariability of aerosols with dynamics and thermodynamics.

On modeling:

- Poor performances of modeling simulations on updraft intensity and cloud properties even with LES/CRM. The reasons are very complicated. Large-scale forcing and dynamical and cloud microphysical parameterization could contribute.
- Difficulties in singling out true aerosol impacts from natural variability due to sensitivity of DCCs to the initial small perturbation.
- Cloud microphysical parameterization: bulk schemes are not well configured for studying ACI; bin schemes are still too computationally expensive and the accuracy is limited by our understanding of cloud microphysical processes.
- Lack of appropriate ice nucleation parameterizations for DCCs.
- Parameterization of subgrid clouds in GCMs: the scale-aware and aerosol-aware issues.

In line with those significant issues, we propose the following further directions that are key to improve understanding and reduce uncertainties in ACI:

- Establish concurrent measurements of aerosol properties and cloud microphysical and dynamic properties over a range of temporal and spatial scales collected over typical climate regimes, and conduct closure and budget studies. These can be achieved through ARM supersite ideas and box closure experiments. Focus on MCSs owing to their importance in global precipitation and radiation budgets. Long term, such concurrent measurements and the closure studies will also allow us to address the covariability of aerosols with dynamics and thermodynamics. Other approaches to address the covariability include combining with model sensitivity simulations and designing special field campaigns.
- Employ the approaches of 1) ensemble simulations, 2) simulations for a long time period and over a large region, and 3) microphysics piggybacking for modeling studies to isolate aerosol impacts from natural variability.
- Improve physical understanding of hydrometer size distribution, conversions among different types of hydrometers, and the hydrometer size and fall speed. Develop and improve cloud microphysical parameterizations.
- Improve understanding and parameterization of ice nucleation and ice microphysical processes that are the key to accurately simulate cold cloud and DCC properties.
- Constrain parameterizations with observations at various scales. Recommend evaluating model simulations with aerosol, cloud and precipitation properties before any impacts of aerosol variability be examined.

Acknowledgments. This study was supported by the U.S. Department of Energy (DOE) Atmospheric System Research (ASR) Program (Grant 200180). The Pacific Northwest National Laboratory (PNNL) is operated for the DOE by Battelle Memorial Institute under Contract DE-AC06-76RLO1830. Yuan Wang’s contribution to this work was sponsored by NASA ROSES14-ACMAP and was carried at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. X. Liu acknowledges the funding support from the U.S. DOE ASR Program (Grant DE-SC0014239). The authors appreciate Drs. Bob Houze, Jerome Fast, and Steve Ghan at PNNL for their review and helpful comments to improve the paper prior to submission.

REFERENCES


Wegener, A., 1911: Thermodynamik der Atmosphäre. J. A. Barth, 331 pp.


