

Evaluation of aerosol-cloud interaction in the GISS ModelE using ARM observations

G. de Boer,^{1,2,3} S. E. Bauer,^{4,5} T. Toto,⁶ Surabi Menon,^{3,7} and A. M. Vogelmann⁶

Received 10 October 2012; revised 29 March 2013; accepted 25 April 2013; published 17 June 2013.

[1] Observations from the US Department of Energy's Atmospheric Radiation Measurement (ARM) program are used to evaluate the ability of the NASA GISS ModelE global climate model in reproducing observed interactions between aerosols and clouds. Included in the evaluation are comparisons of basic meteorology and aerosol properties, droplet activation, effective radius parameterizations, and surface-based evaluations of aerosol-cloud interactions (ACI). Differences between the simulated and observed ACI are generally large, but these differences may result partially from vertical distribution of aerosol in the model, rather than the representation of physical processes governing the interactions between aerosols and clouds. Compared to the current observations, the ModelE often features elevated droplet concentrations for a given aerosol concentration, indicating that the activation parameterizations used may be too aggressive. Additionally, parameterizations for effective radius commonly used in models were tested using ARM observations, and there was no clear superior parameterization for the cases reviewed here. This lack of consensus is demonstrated to result in potentially large, statistically significant differences to surface radiative budgets, should one parameterization be chosen over another.

Citation: de Boer, G., S. E. Bauer, T. Toto, S. Menon, and A. M. Vogelmann (2013), Evaluation of aerosol-cloud interaction in the GISS ModelE using ARM observations, *J. Geophys. Res. Atmos.*, 118, 6383–6395, doi:10.1002/jgrd.50460.

1. Introduction

[2] Simulation of the impacts of aerosols on clouds and the resulting change in surface and top of atmosphere radiative budgets continues to be a major source of uncertainty in estimating future climate [Intergovernmental Panel on Climate Change, 2007]. Alteration to the global distribution of aerosol resulting from anthropogenic climate change can impact the properties of clouds through various aerosol indirect effects (AIE). These effects include the first [Twomey, 1977] and second [Albrecht, 1989] aerosol indirect effects, which impact the albedo and precipitation production of clouds, respectively, by altering the number (N_d), and therefore effective size (r_e), of cloud droplets under constant liquid water path (LWP). These changes can result in cloud

thermodynamic and microphysical [e.g., Koren *et al.*, 2005] changes, further impacting cloud properties. Combined with an incomplete and evolving understanding of the interaction pathways between aerosols and clouds, uncertainties in present-day (PD) and pre-industrial (PI) aerosol budgets result in large variability between simulated estimates of the influence of AIE on the earth's radiative budget, with varying studies providing different ranges (e.g., -1.2 to -0.2 W m⁻², Quaas *et al.* [2009]; -1.85 to -0.5 W m⁻², Chen and Penner [2005]).

[3] There have been numerous observational campaigns that aim to quantify the magnitude of AIE. Specifically, the first AIE has been targeted using surface-based [e.g., Feingold *et al.*, 2003; McComiskey *et al.*, 2009], in situ [e.g., Twomey *et al.*, 2005; Berg *et al.*, 2011], and satellite-based [e.g., Rosenfeld and Feingold, 2003; Menon *et al.*, 2008; Quaas *et al.*, 2009] measurements. In general, the first AIE, as defined by the ratio of a change in the natural logarithm of droplet effective radius to a change in the natural logarithm of aerosol amount, has been demonstrated to fall between theoretical limits of 0 and 0.33, with significant variability from one estimate to the next. These differences may result from several competing mechanisms, including differences in instrument sensitivity [Rosenfeld and Feingold, 2003], sample scale [McComiskey and Feingold, 2012], and the region sampled [Sekiguchi *et al.*, 2003]. Additionally, it is possible that uncertainties are the result of an incomplete understanding of the processes involved in the AIE [Shao and Liu, 2009].

¹Cooperative Institute for Research in Environmental Sciences, Boulder, Colorado, USA.

²Physical Sciences Division, NOAA Earth System Research Laboratory, Boulder, Colorado, USA.

³Lawrence Berkeley National Laboratory, Berkeley, California, USA.

⁴The Earth Institute at Columbia University, New York, USA.

⁵NASA Goddard Institute for Space Studies, New York, USA.

⁶Brookhaven National Laboratory, Upton, New York, USA.

⁷ClimateWorks Foundation, San Francisco, California, USA.

Corresponding author: G. de Boer, Cooperative Institute for Research in Environmental Sciences, Boulder, CO, USA. (gijis.deboer@colorado.edu)

©2013. American Geophysical Union. All Rights Reserved.
2169-897X/13/10.1002/jgrd.50460

[4] The last two decades have seen the incorporation of AIE into global climate models (GCMs). Because of the coarse resolution of these models, aerosol impacts on clouds have been parameterized. Often, the number of cloud droplets is simply prescribed to be a function of the number of aerosol particles present. This has been done using several techniques, including logarithmic and exponential fits to measured data [e.g., *Menon and Rotstajn*, 2006]. These parameterizations relate the number of cloud droplets either to aerosol mass concentration [e.g., *Roelofs et al.*, 1998] or, alternatively, to aerosol number concentration [e.g., *Suzuki et al.*, 2004]. *Quaas et al.* [2009] performed an evaluation of aerosol indirect effects in a variety of GCMs, including the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS) ModelE run at $4^\circ \times 5^\circ$ resolution. In that work, the model depicted aerosol indirect effects were compared to those derived from a variety of satellite measurements. Through this technique, a positive correlation between simulated cloud fraction and aerosol optical depth was found, although the exact reasons behind this relationship remain ambiguous. Additionally, that work demonstrated that there are large differences between different GCM aerosol-cloud relationships, and that the simulated relationships are not always of the same sign as those obtained from satellite or ground-based measurements. At the same time, the sign of the relationship between cloud droplet number concentration and aerosol optical depth was demonstrated to be consistent between models, with simulated interactions occurring over oceanic regions closer to observations than those over continents.

[5] The aim of the current study is to evaluate aspects of the first AIE as simulated in the NASA GISS GCM, the ModelE [*Schmidt et al.*, 2006] and to advance our understanding of what improvements may be necessary. This evaluation is conducted using surface and in situ measurements collected at several sites by the United States Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) program [*Mather and Voyles*, 2013]. A description of the ModelE, along with an overview of the different measurement data sets, is provided in sections 2 and 3. Results from the evaluation are provided in section 4, and we conclude with a summary and discussion of results in section 5.

2. Model Description

[6] Simulations were completed using a recent version of the NASA GISS GCM ModelE, developed for the fifth IPCC assessments (CMIP5). The GISS ModelE contributions to the CMIP5 archive are improved over those used for CMIP3 (and described in *Schmidt et al.* [2006] and *Hansen et al.* [2007]) in a number of respects (*Schmidt et al.*, Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, manuscript in preparation, 2013). First, the model has a higher horizontal and vertical resolution (2° lat \times 2.5° longitude, 40 layers). The vertical layers are distributed on a non-uniform grid, with spacing of roughly 25 mb (250 m) from the surface to 850 mb, and roughly 40–50 mb (400–700 m) from 850 to 415 mb. Second, various physics components have been upgraded from the CMIP3 version, namely the convection scheme, stratiform cloud scheme, gravity wave drag, sea ice, and ocean physics.

[7] The GCM is coupled to the Multiconfiguration Aerosol Tracker of Mixing state (MATRIX) [*Bauer et al.*, 2008, 2010]. MATRIX is designed to support model calculations of the direct and indirect effect and permits detailed treatment of aerosol mixing state, size, and aerosol-cloud activation, making it possible to evaluate these quantities against observations. For each aerosol population defined by mixing state and size distribution, the tracked species are number concentration and mass concentration of sulfate, nitrate, ammonium, aerosol water, black carbon, organic carbon, mineral dust, and sea salt. Here we use the aerosol population setup called mechanism 1, given in Table 1 of *Bauer et al.* [2008]. MATRIX dynamics includes nucleation, new particle formation, particle emissions, gas-particle mass transfer, aerosol phase chemistry, condensational growth, coagulation, and cloud activation.

[8] To simulate the indirect effect, we follow a similar treatment as described in *Menon et al.* [2010] that includes several changes to the treatment of cloud drop and ice crystal nucleation following the scheme from *Morrison and Gettelman* [2008]. For cloud droplets, we use a prognostic equation to calculate N_d , based on *Ghan et al.* [1997], given as follows:

$$\frac{d(N_d)}{dt} = S - L_{\text{auto,cont,imm}} \quad (1)$$

where S is the source term, including newly nucleated cloud droplets and L is a loss term accounting for droplet loss through the process of autoconversion, contact nucleation, and via immersion freezing. For stratiform clouds, the source term is obtained from MATRIX using the scheme of *Abdul-Razzak and Ghan* [2000] that is based on Köhler theory for multiple external lognormal modes that are composed of internally mixed soluble and insoluble material.

[9] For this work the model is run continuously from year 2002 to 2009, covering all of the observational campaigns described below. In order to force representative meteorology in the GCM, the horizontal wind components of the model are nudged toward the MERRA reanalysis data set (<http://gmao.gsfc.nasa.gov/merra/>). MERRA winds are available on a 6 hourly time step and are linearly interpolated to the model 30 min time step. The aerosol scheme uses the CMIP5 emissions by *Lamarque et al.* [2010]. This setup has also been used by *Bauer and Menon* [2012], who provide further details.

3. Measurement Description

[10] Observational data sets come from four separate intensive operations periods (IOPs). These IOPs were chosen due to their focus on both aerosol and cloud measurements and due to the variety of regions represented. All of these studies were funded by the US DOE's ARM program. Brief descriptions of each study, as well as an overview of the measurements used are given below and in Table 1. A map showing the various measurement locations is provided in Figure 1.

3.1. The ARM Aerosol IOP

[11] The Aerosol IOP was conducted during May 2003 at the DOE ARM Climate Research Facility (ACRF) Southern Great Plains (SGP) site. This site, extending over parts of the central United States, has a central measurement facility near Lamont, OK (36.61°N , 97.49°W). Although the

Table 1. An Overview of Measurements and Sources Used in This Study^a

Variable	Aerosol IOP	MASRAD	MASE	China AMF	RACORO
T_{sfc}	SMOS	MET	MET	MET	SMOS
P_{sfc}	SMOS	MET	MET	MET	SMOS
RH_{sfc}	SMOS	MET	MET	MET	SMOS
$Wind_{\text{sfc}}$	SMOS	MET	MET	MET	SMOS
CCN_{sfc}	DRI (0.4%)	DMT (0.3%)	DMT (0.3%)	DMT (0.3%)	DMT (0.3%)
LWP	MWR	MWR	MWR	MWR	MWR
AOD	ABE	MFRSR	MFRSR	MFRSR	MFRSR
$R_{e,\text{sfc}}$	n/a	2NFOV	2NFOV	n/a	n/a
N_a	CPC (> 7 nm)	n/a	CPC (> 10 nm)	n/a	CPC (> 10 nm)
N_d	CAS (2.0 – 37.3 μm)	n/a	CAS (2.1 – 41 μm)	n/a	CAS (2.3 – 40.5 μm)
LWC	CAS	n/a	CAS	n/a	CAS
R_e	CAS	n/a	CAS	n/a	CAS
R_v	CAS	n/a	CAS	n/a	CAS

^aPlease note that this table does not include all available measurements from the individual campaigns.

main focus of the experiment was improving understanding of aerosol impacts on radiative transfer, several instruments were included to measure basic cloud properties. Measurements of surface temperature (T_{sfc}), pressure (P_{sfc}), relative humidity (RH_{sfc}), winds (U_{sfc} , and V_{sfc}) and precipitation were made using the ARM Surface Meteorological Observation System (SMOS). Atmospheric liquid water path (LWP) was retrieved from microwave radiometer (MWR) measurements using algorithms described in *Turner et al.* [2007], and aerosol optical depth (AOD) was obtained from the Aerosol Best Estimate (ABE) product. ABE derives AOD from (in order of preference) a multi-filter shadow band radiometer (MFRSR, *Harrison and Michalsky* [1994]), a Raman lidar [*Goldsmith et al.*, 1998], or a combination of sources via algorithms described in *Sivaraman et al.* [2004]. Concentrations of surface cloud condensation nuclei (CCN) were obtained at multiple supersaturations using an instrument from the Desert Research Institute.

[12] As a part of this experiment, the Center for Interdisciplinary Remotely-Piloted Aircraft Studies (CIRPAS) Twin Otter conducted 60.6 flight hours on 15 separate days, profiling aerosol and cloud properties, providing complementary measurements to those collected at the surface. The aircraft's payload included a Condensation Particle Counter (CPC) for counting particles with diameters larger than 7 nm, a Passive Cavity Aerosol Spectrometer Probe (PCASP), providing concentrations for particles between 0.1 and 3.2 μm in diameter, a Forward Scattering Spectrometer Probe (FSSP) for particles and droplets with diameters between 2.4 and 52 μm , and a Cloud and Aerosol Spectrometer (CAS) counting particles between 0.6 and 55 μm . For the current study, we use the CAS bins that cover particles between 2 and 37.3 μm . From the CAS measurements, profiles of cloud liquid water content (LWC), cloud droplet effective radius (R_e), and cloud droplet volumetric radius (R_v) were obtained.

3.2. MASE/MASRAD

[13] The ARM Mobile Facility (AMF) was deployed to Point Reyes, California (38.09°N, 122.96°W) during 2005 for the Marine Stratus Radiation Aerosol and Drizzle (MASRAD) campaign. The MASRAD IOP occurred between mid-March and mid-September of 2005. During the month of July 2005, MASRAD was integrated with the Marine Stratus/Stratocumulus Experiment (MASE,

Lu et al. [2007]), to provide both surface and airborne observations of cloud and aerosol properties. Located on the Pacific Ocean north of San Francisco, Pt. Reyes frequently has marine stratus and stratocumulus clouds. T_{sfc} , P_{sfc} , RH_{sfc} , U_{sfc} , V_{sfc} , and precipitation were obtained from the AMF surface meteorology suite (MET). Surface aerosol concentrations were obtained using the AMF Aerosol Observing System (AOS, *Jefferson* [2011]), which uses a Droplet Measurement Technologies (DMT) CCN counter to derive CCN concentrations, while AOD is obtained from the MFRSR. Finally, cloud optical depth is obtained using a two-channel Narrow Field of View Radiometer (2NFOV, *Chiu et al.* [2006]). This combination of measurements was used by *McComiskey et al.* [2009] to complete a detailed analysis of aerosol-cloud interactions over Point Reyes, investigating the relationships between aerosol (e.g., surface CCN concentration and aerosol light scattering) and cloud (e.g., cloud optical depth, cloud droplet effective radius, and droplet number concentration) properties.

[14] During MASE, the DOE/Pacific Northwest National Laboratory (PNNL) Gulfstream-1 (G-1) aircraft flew legs along the Point Reyes shoreline to collect additional cloud and aerosol information. Included in this set of measurements are the concentrations of particles larger than 3 and 10 nm from two CPCs, particles between 0.016 and 0.444 μm by a Differential Mobility Analyzer (DMA), particles between 0.11 and 2.65 μm by a PCASP, particles between 0.7 and 54 μm by a CAS, and particles between 25 and 1550 μm by a Cloud Imaging Probe (CIP). In the present study, we use observations from the CPC measuring particles larger than 10 nm and from CAS bins covering droplet

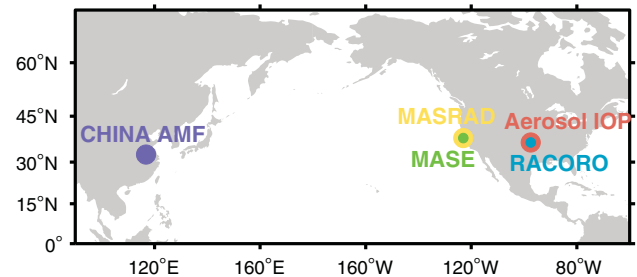


Figure 1. A map illustrating the different measurement locations for each of the five analyzed campaigns.

sizes between 2.1 and 41 μm . Also, during the Aerosol IOP, profiles of cloud LWC, cloud droplet R_e , and cloud droplet R_v were obtained from the CAS probe.

3.3. China AMF Deployment

[15] In collaboration with Chinese partners, the AMF was deployed in Shouxian, China (32.56°N, 116.78°E) between mid-May and late December 2008. Shouxian, located roughly 500 km to the west of Shanghai, is a continental location largely surrounded by farmland. Again, all surface meteorology measurements listed for the above campaigns were available from the MET. Also, during MASRAD, the ARM AOS was deployed to China, measuring a wide variety of aerosol properties including aerosol absorption, concentration, scattering, hygroscopic growth, inorganic composition, and size distribution. In addition, a combination of cloud measurements were obtained, including cloud boundaries, cloud optical depth from a 2NFOV, and LWP from a MWR. Finally, AOD was measured using the MFRSR.

3.4. RACORO

[16] Between January and June of 2009, the ARM Aerial Facility (AAF) completed routine flights over the SGP site during the Routine AAF Clouds with Low Optical Water Depths (CLOWD) Optical Radiative Observations (RACORO) field campaign [Vogelman *et al.*, 2012]. Flights were completed using the CIRPAS Twin Otter equipped with a variety of cloud, aerosol, and radiation sensors. Aimed specifically at clouds with low optical depths, this campaign sought to answer how aerosols impact these thinner clouds. Measurements from the Twin Otter include CAS-derived LWC and cloud drop size distribution, CCN concentration, aerosol size distributions, and basic meteorology and radiation measurements. Additionally, the full suite of surface cloud and aerosol measurements available at SGP are available for much of the campaign. For this campaign, we use CAS measurements from bins covering sizes between 2.3 and 40.5 μm , and CPC aerosol concentrations for particles larger than 10 nm.

4. Model Evaluation

4.1. Notes on Sampling

[17] A major consideration in evaluations such as in this study is how to best analyze available measurements to appropriately represent the scales inherent to the GCM grid box [McComiskey and Feingold, 2012]. This holds true for satellite, in situ and surface-based observations. While in situ and surface-based observations have the potential to capture process-level relationships between aerosol and cloud properties, due to their small spatial sample size, they cannot capture the spatial variability within a GCM grid box without averaging over extended time periods. Satellite observations generally include aggregate scales larger than the spatial variability inherent in cloud and aerosol fields, thereby blurring any relationships that may be hidden within the data set. In this instance, a 2° grid and a surface/aircraft-based data set make this challenging. A simple approach is aggregation (averaging) of the data over time scales that begin to capture the spatial variability in the GCM grid. Naively, it may be assumed that it would be

appropriate to obtain the average over a time period that covers the full scale of the grid box assuming some advective velocity (e.g., 10 m s^{-1}). At 2°, this requires averaging of periods on the order of 6–7 h. Using this technique obscures and would even wash-out relationships of interest in part, due to the changes in aerosol and cloud properties occurring with the diurnal cycle. An alternative approach entails averaging over shorter periods (e.g., 1 h) in order to capture some of the subgrid scale variability in the measurements, while maintaining signals inherent in an evolving atmosphere. This short aggregation timescale is very appropriate for time periods featuring consistent large scale meteorological conditions and a relatively homogeneous surface, but may fail during frontal passages or at coastal sites. While more complex techniques [e.g., McComiskey and Feingold, 2012] may enhance future evaluations, in this work, the latter method (1 h averaging) is employed. Along with this 1 h averaging, vertical sampling from aircraft measurements is averaged to a coarse resolution that is comparable to that of the GCM, without consideration for the number of samples in any grid box at a given time.

4.2. Meteorological Evaluation

[18] Due to the nudged nature of the evaluated runs, it is expected that the general weather conditions simulated for each campaign are similar to those observed. This similarity is important in order to ensure that conditions impacting cloud formation (e.g., temperature and humidity) and aerosol transport (e.g., winds) are generally comparable. Only under similar atmospheric conditions would simulated aerosol-cloud interactions be expected to be comparable to those observed. In order to evaluate simulated meteorology, we compare simulated and observed surface (1 m) air pressure (P_{sfc}), surface (2 m) air temperature (T_{sfc}), surface (2 m) relative humidity (RH), and surface (10 m) wind components (U,V) from each of the campaigns. Distributions of the differences (model minus obs) between simulated and observed values are shown in the form of box plots in Figure 2. For each campaign, distributions of averaged hourly (darker color) and daily (lighter color) differences are illustrated. As might be expected, the averaging reduces overall variability and acts to reduce the differences observed at an hourly interval. It is important to remember that the different campaigns had vastly different numbers of data points on which the statistics are based (in parentheses in Figure 2), with the Aerosol IOP and MASE only representing 1 month each.

[19] Looking first at surface pressure, the influence of the California coastline shows up clearly in the MASRAD and MASE evaluations. Because the model grid box is significantly larger than the site at which observations were taken, the gradient in surface pressure from sea level to more elevated inland locations results in simulated surface pressures that are substantially (20+–30 mb) lower than those that were observed at Pt. Reyes. For the campaigns occurring at SGP and in China (Aerosol IOP, AMF China, and RACORO), the mean error (not shown) generally falls around 0 mb, and overall distributions are similar between ModelE and observations. The majority of cases from these campaigns (majority being defined by the IQR) demonstrate hourly P_{sfc} errors of less than 4 mb.

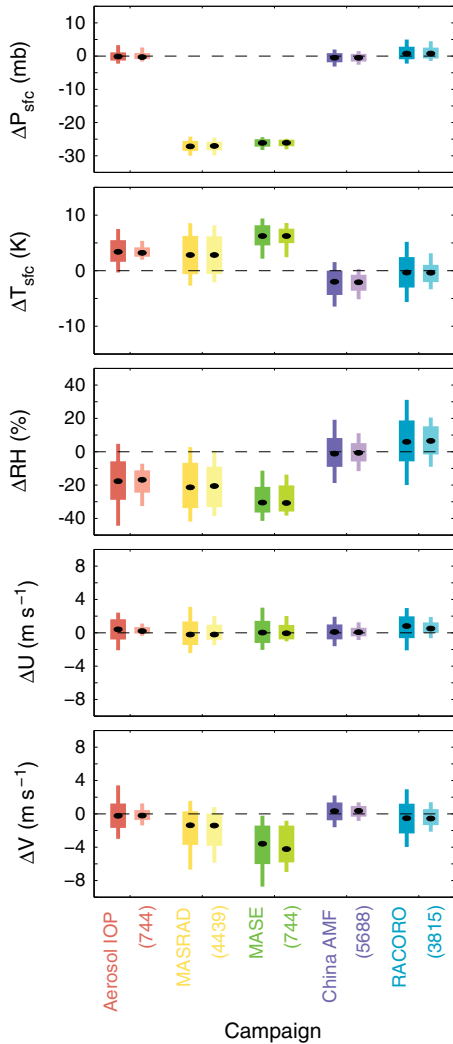


Figure 2. Distributions of differences (model-obs) for hourly (darker) and daily (lighter) average surface pressure (ΔP_{sfc}), surface air temperature (ΔT_{sfc}), surface relative humidity (ΔRH), surface U wind speed (ΔU), and surface V wind speed (ΔV). The box plots used to illustrate the difference distributions include the median (black circle at center of box), interquartile range (IQR: box edges), and the extent of the 10th and 90th percentiles (whiskers) of the data. Numbers in parentheses represent the number of 1 hourly averages included in these distributions.

[20] An evaluation of the simulated surface air temperatures also demonstrates some issues with the ModelE results, as compared to the single site measurements from ARM observations. Comparison to campaign data from the Aerosol IOP, MASRAD, and MASE all demonstrate model warm biases. While this may not be surprising at Pt. Reyes (MASRAD, MASE) due to coastline effects, it is somewhat surprising at SGP (Aerosol IOP). This is particularly true since a similar bias does not appear for the RACORO campaign, also held at SGP. In general, the RACORO comparison fares the best, with mean and median differences of roughly 0 K. China also has smaller differences in T_{sfc} than the other three campaigns, with only a small (2 K) model cold bias. These temperature biases result in corresponding

biases in surface relative humidity, with median differences between measured and observed RH of around 20%–30% for the first three campaigns. Both the AMF China and RACORO campaigns demonstrate small median errors in surface RH.

[21] Finally, an evaluation of the U (east-west) and V (north-south) component wind speeds demonstrates that, despite the nudging, there are some differences in wind properties, primarily in the V component. Median differences in the U component are small, with the majority of differences smaller than 2 m s^{-1} . The V component is a bit of a different story, particularly at Pt. Reyes, with median model errors of around -2 m s^{-1} and -4 m s^{-1} for MASRAD and MASE, respectively. Data points from the other campaigns fall much closer to the zero error line, indicating fairly good agreement between simulated and measured winds.

[22] Given these results, it is possible to assume a similar range of meteorological conditions between the observations and simulations. Therefore, it is not unreasonable to evaluate the simulated interactions between aerosols and clouds using observed atmospheric properties assuming that scaling considerations as discussed above are taken into account.

4.3. Surface-Based Aerosol Evaluation

[23] Figure 3 illustrates the differences between observed and simulated (model minus obs) aerosol optical depth and surface CCN concentration. For the most part, aerosol optical depths were found to be comparable with all campaigns except the China AMF deployment showing mean differences between 0 and 0.25. The China deployment was unique in the high aerosol loadings observed. Generally, these extreme values were not captured by the model at the right time, resulting in the large variability in differences for that campaign, including a mean difference of over 0.5. The coastal Pt. Reyes site was demonstrated to have the closest comparison and the least amount of scatter in the differences between measured and observed values.

[24] Because aerosol concentration measurements are made at various supersaturation levels, care had to be taken to ensure a fair comparison. This included limiting the observational data set to CCN at supersaturations of 0.3% when

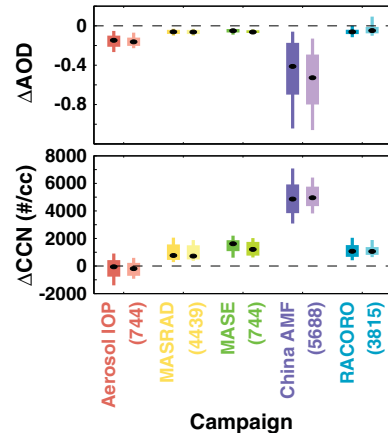


Figure 3. Same as Figure 2, except for (top) aerosol optical depth, AOD and (bottom) surface CCN at 0.3% supersaturation.

possible, since this was the most commonly available value for the different campaigns. Where surface CCN were not available at 0.3%, CCN concentrations at 0.3% were derived by interpolating between the two closest values using a power law of the following form:

$$N_{\text{CCN}} = CS^k \quad (2)$$

where N_{CCN} is the CCN number concentration, S is the supersaturation, and C and k are constants determined from the surrounding CCN data. This method was tested for April 2007 when CCN were available at 0.18%, 0.3%, and 0.43%. When the 0.3% data was withheld from the fitting, the interpolated result at 0.3% agreed with the observed value to within 7%.

[25] Simulated near-surface CCN concentrations were generally different than those observed. With the exception of the Aerosol IOP period, simulated values for all campaigns are higher than those observed. China surface CCN concentrations were generally largely over-predicted by the model. As with the AOD comparison, the Pt. Reyes site features the least amount of variability between model and observed estimates. Of note in these evaluations is that, with the exception of the Aerosol IOP, the sign of the difference between observations and the model for AOD is opposite that of the surface CCN concentration. This demonstrates that while AOD was generally lower in simulations than the observations indicate, this reduced aerosol amount is not present in the form of insufficient CCN at the surface. This implies that the vertical distribution of aerosol concentration in the model appears to be different than in the measured atmosphere, a fact confirmed by CCN profile information available from aircraft during the Aerosol IOP, MASE, and RACORO (not shown). These profiles demonstrate large differences between observed and measured CCN concentrations, particularly in the lower atmosphere (< 800 mb). This inconsistency between surface CCN and AOD biases serves as a reminder of the possible danger of using AOD as a proxy for aerosol concentrations at cloud height. Additionally, because of the discrepancies in aerosol concentrations, assessment of aerosol indirect effects is completed through the evaluation of general relationships between parameters, rather than time-by-time comparison.

4.4. Aerosol-Cloud Interaction

[26] In this section we evaluate the simulation of various interaction pathways between aerosols and clouds in the model. These include aerosol activation to liquid droplets, the parameterization of effective radius and broadening of the droplet size distribution by aerosol effects, and the relationship between surface-based estimates of cloud effective radius and aerosol concentration.

4.4.1. Activation

[27] One of the most fundamental ways in which aerosol particles impact cloud characteristics is through the activation of cloud droplets. As discussed above, it is generally believed that high aerosol concentrations result in a larger number of smaller droplets. Climate models have traditionally handled this activation parameterization via a few different mechanisms [e.g., *Nenes and Seinfeld*, 2003]. The first is through the derivation of empirical relationships linking aerosol number concentration to droplet number concentration. Examples of these types of relationships are available

in *Menon et al.* [2008]. The current version of the GISS ModelE uses this type of relationship for convective clouds. There are different empirical relationships for ocean and land. Over oceans, the relationship is assumed to be

$$N_d = -29.6 + 4.92N_a^{0.694} \quad (3)$$

Over land, the following relationship is used:

$$N_d = 174.8 + 1.51N_a^{0.886} \quad (4)$$

where N_a is the total aerosol concentration, and N_d is the liquid droplet number concentration. As discussed in the model description, these relationships are not applied to stratiform clouds, where Köhler theory is used to calculate droplet activation.

[28] In order to assess the activation parameterizations in ModelE, we compare the relationship between hourly averaged total aerosol number concentration and in-cloud liquid droplet concentration from the observations and the model. For the Aerosol IOP, aerosol concentrations were derived from measurements taken by the CPC ($D > 10$ nm), and cloud droplet concentrations came from the CAS probe ($2.02 \mu\text{m} > D > 43.05 \mu\text{m}$), both sampled on the CIRPAS Twin Otter. For MASE, aerosol concentrations were derived from the CPC probe ($D > 10$ nm) while cloud droplet concentrations again came from the CAS ($2.1 \mu\text{m} > D > 31 \mu\text{m}$), both mounted on the G1 aircraft. Finally, for RACORO, aerosol concentrations were again derived from the CPC ($D > 10$ nm), and cloud droplet concentrations were again taken from the CAS ($2.3 \mu\text{m} > D > 50.1 \mu\text{m}$). While aerosol concentrations relevant to activation could have been used in deriving these relationships, the model parameterizations used for cumulus clouds do not include size cutoffs and are based on the total number of aerosol particles. Therefore, smaller particles (down to 10 nm) were included in these comparisons.

[29] The results of this evaluation are presented in Figure 4. Figure 4 (left) shows a two-dimensional histogram of model aerosol concentration (cm^{-3}) as compared to liquid droplet concentration (cm^{-3}) from the different measurement campaigns. These results are limited to those obtained from the lowest six model levels. Along with these, there are four lines representing different model parameterizations [*Menon et al.*, 2008] (see Table 2). The solid lines represent the relationships introduced in equations 3 and 4, while the dashed lines are similar parameterizations previously used for stratiform clouds. While the stratiform parameterizations are no longer used in ModelE, they are plotted to demonstrate some of the differences we see between locations. Figure 4 (right) includes data points (colored circles) from hourly averaged aircraft measurements for each campaign (as described in the previous paragraph). The measurements from Pt. Reyes (green), for example, appear to follow the stratiform parameterization more closely than measurements from the Aerosol IOP (red) or RACORO (blue). This is not surprising due to the frequent occurrence of stratiform clouds along the California coastline.

[30] One interesting point to note is that the model appears to frequently produce cases for which the number of liquid droplets exceeds the number of aerosol particles within that volume. It is speculated that this could be the result of a number of processes, including advection from other

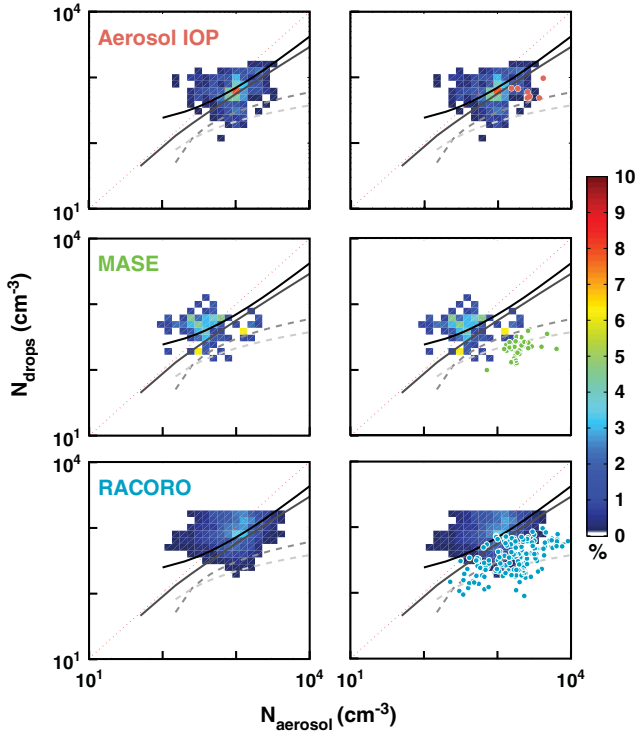


Figure 4. Relationships between in-cloud aerosol concentration and in-cloud droplet concentration for the different campaigns. (left) Two-dimensional histograms of the model values for the time period of the campaigns listed, with the color bar indicating the percentage of cases in each bin. (right) Observational estimates from CPC (aerosol) and CAS (droplets) probes, as illustrated by colored circles. Gray scale curves represent parameterized relationships provided in Table 2. This comparison is only for single-layer, warm, and non-precipitating clouds occurring below 850 mb.

grid boxes, gravitational settling of droplets from higher in the domain, and possibly non-local breakup of droplets. In the current environment, it seems as though horizontal advection of cloud droplets is the most likely cause for the observed behavior (this was also suggested by *Morrison and Gettelman* [2008] as a potential cause for local droplet concentrations exceeding aerosol amounts). While the current figure does include precipitating environments, these cases are limited only to small precipitation rates, thereby limiting the contributions of the breakup and settling processes. Additionally, these high droplet number concentrations are in part caused by an inconsistency between the activation and scavenging calculations, which are calculated separately from one another. Currently, in-cloud removal of aerosol particles in ModelE is being updated to directly take into account droplet number concentration. This should reduce aerosol scavenging, resulting in less cases where droplet concentration exceeds aerosol concentration. Further investigation of this phenomenon is needed but extends beyond the scope of the current effort.

4.4.2. Effective Radius

[31] Cloud droplet effective radius (r_e) is a key property in the calculation of cloud radiative properties [e.g., *Slingo*,

1989]. r_e is defined to be the ratio of the third to the second moment of the droplet size distribution:

$$r_e = \frac{\int_{r=0}^{\infty} N_r r^3 dr}{\int_{r=0}^{\infty} N_r r^2 dr} \quad (5)$$

[32] Parameterization of droplet r_e has been accomplished using several different relationships to cloud properties. *Martin et al.* [1994] and others derived a power-law relationship in the form of

$$r_e = \alpha \left(\frac{\text{LWC}}{N_d} \right)^{\frac{1}{3}} \quad (6)$$

where LWC is the liquid water content in g m^{-3} , N_d is the droplet concentration in cm^{-3} , and α is a prefactor. An initial estimate for α from *Bower and Choulaton* [1992] was 62.04. Later, *Martin et al.* [1994] derived separate values for maritime (66.83) and continental (70.89) clouds. A similar expression was used by *Del Genio et al.* [1996], in the following form:

$$r_e = r_0 \left(\frac{\text{LWC}}{\text{LWC}_0} \right)^{\frac{1}{3}} \quad (7)$$

where LWC is the liquid water content in g m^{-3} , $r_0 = 10 \mu\text{m}$ and $\text{LWC}_0 = 0.25 \text{g m}^{-3}$.

[33] These expressions, however, do not account for the impacts of broadening of the droplet spectrum resulting from processes such as entrainment and mixing, both possibly impacted by aerosol-induced changes to droplet size. In order to accomplish this, α should be a factor of droplet concentration, a change implemented by *Liu and Daum* [2002]. In their parameterization:

$$r_e = \beta r_v \quad (8)$$

where:

$$r_v = \left(\frac{3\text{LWC}}{4N_d\pi\rho_l} \right)^{\frac{1}{3}} \quad (9)$$

and:

$$\beta = \frac{(1 + 2(1 - 0.7\exp(-0.003N_d))^2)^{\frac{2}{3}}}{(1 + (1 - 0.7\exp(-0.003N_d))^2)^{\frac{1}{3}}} \quad (10)$$

[34] In this work, we test these parameterizations against in situ measurements from the three campaigns in which aircraft were used. Instead of using the α values suggested in the literature, we use a range of values (60, 65, 70) surrounding those previously suggested. It should be noted that r_e is calculated using each of the parameterizations from measured properties from the campaigns. In this way, the functional relationship of the parameterization is tested independent of the model's ability to provide accurate input parameters. Applying different parameterizations directly to measured quantities also eliminates the scale issues previously discussed in regard to simulation validation assuming

Table 2. Parameterizations Represented in Figure 4 (From *Menon et al.* [2008])

Parameterization	Domain	Cloud Type
$N_d = -598 + 298 \log_{10}(N_a)$	Land	Stratus
$N_d = -273 + 162 \log_{10}(N_a)$	Ocean	Stratus
$N_d = 174.8 + 1.51N_a^{0.886}$	Land	Cumulus
$N_d = -29.6 + 4.92N_a^{0.694}$	Ocean	Cumulus

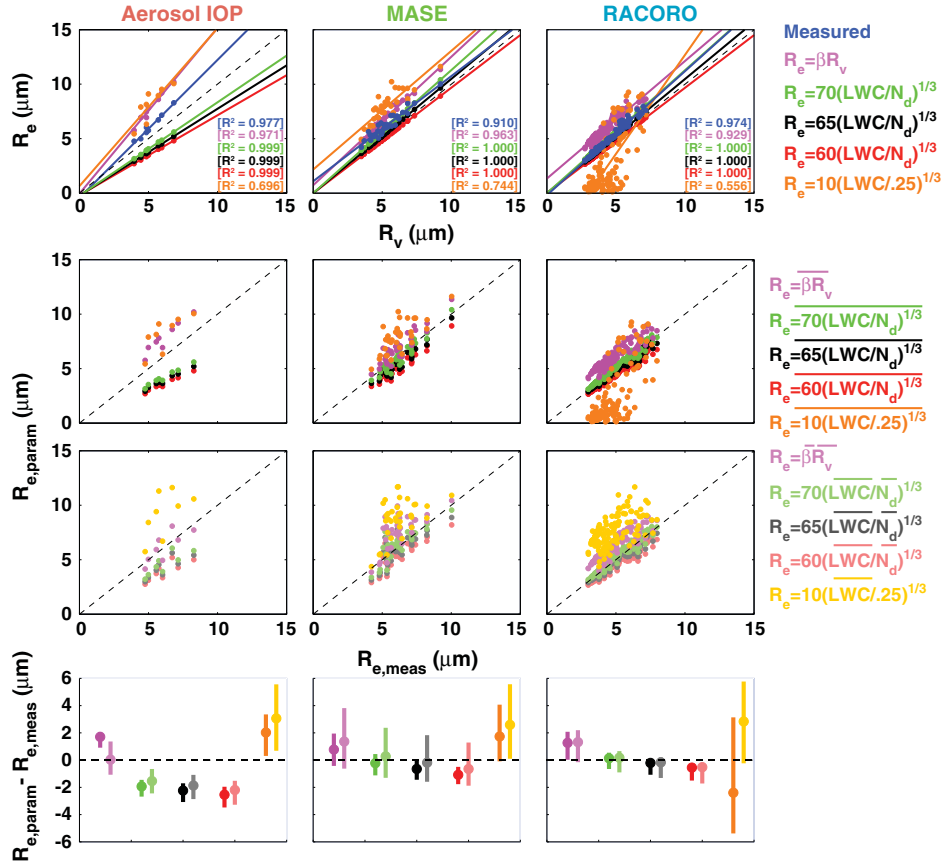


Figure 5. (first panel, from the top) Relationships between cloud droplet effective radius and volumetric radius from aircraft measurements (blue) and various parameterizations, (second and third panels) a comparison of measured and parameterized effective radius for the same grouping of parameterizations for different averaging procedures, and (fourth panel) distributions of errors for the effective radius evaluation. For Figures 5(second)–5(fourth), darker symbols represent one hourly averages of effective radii calculated using instantaneous in situ data, while lighter symbols represent effective radii calculated using one hourly averaged in situ measurements, as indicated by the legend.

that the parameterizations are applied at the local (cloud) scale. The r_e values are then averaged over 1 h periods to allow for comparison to the spatial scale covered by the GCM.

[35] The CAS probe provides size and number concentration estimates, meaning that r_e can be derived from the CAS data as follows:

$$r_e = \frac{\sum_{i=1}^n N_{d,i} \left(\frac{d_{g,i}}{2}\right)^3}{\sum_{i=1}^n N_{d,i} \left(\frac{d_{g,i}}{2}\right)^2} \quad (11)$$

where $N_{d,i}$ is the number measured in bin i , $d_{g,i}$ the measured geometric diameter in bin i , and n is the number of instrument bins. Since LWC can be derived from the CAS particle size distribution, the measured r_e can be compared directly to estimates from the parameterizations discussed above. In addition, the relationship between droplet volumetric radii and both measured and parameterized effective radii is evaluated using these data sets.

[36] Figure 5 illustrates the evaluation of these parameterizations. Figure 5 (first panel, from the top) relates the different r_e parameterizations to the measured volumetric

radius (r_v) for the campaigns featuring aircraft measurements. Polynomial fits, along with their R^2 values are provided. The R^2 values are generally high (most are above 0.9), indicating that the fits are statistically representative of the data presented. Based on this evaluation, it becomes evident that there are sometimes large differences between the different parameterizations, and that there is little evidence from these campaigns to indicate that one parameterization is generally better than others. The blue markers and line (polynomial fit) indicate the relationship between measured r_e and r_v . The slope relating these two properties ranges in value from 0.96 (MASE) to 1.27 (Aerosol IOP), while RACORO falls in between with a slope of 1.13. Generally, equations 7 and 8 predict r_e larger than r_v , and this relationship increases as r_v gets larger (slope > 1). The r_e derived using the various forms of equation 6 produce variable results, with r_e generally smaller than r_v for $\alpha = 60$ (red lines). For other α values, r_e is parameterized to be either smaller or larger than r_v , depending on the campaign.

[37] Figure 5 (second panel) illustrates the relationships between measured and parameterized r_e for each of the three campaigns. Values plotted represent mean values calculated from high temporal resolution (order of seconds) in situ mea-

surements. Generally, r_e derived from equations 7 and 8 exceed measured r_e , while values derived using equation 6 underestimate the measured values. An exception to this is RACORO, where r_e estimates based on equation 7 are generally below measured values, particularly for smaller droplets. Based on the campaigns with larger sample sizes (MASE and RACORO), it seems that r_e from equations 6 with α equal to 70 most closely resemble the measured r_e .

[38] Figure 5 (third panel) illustrates the same relationships as Figure 5 (second panel), except as calculated from 1 h averages of LWC and N_d . This calculation most closely resembles what may be done in a global climate model where LWC and N_d values are intended to be representative of the scale of the grid box. Generally, values calculated using equation 7 are higher than those for the values calculated from high resolution measurements, while estimates from equation 8 are closer to measured quantities. In order to help with this assessment, Figure 5 (fourth panel) illustrates differences between parameterized and measured r_e for all of the data points shown in the middle rows, with circles indicating the mean error, and the lines extending to the minimum and maximum errors. These error bars demonstrate that estimates based on equation 7 are generally the least similar to observations. Beyond that, estimates from equations 6 and 8 both perform well for certain cases, with the best coefficient for α seemingly 70 (green symbols and lines).

4.4.3. Surface-Based First Indirect Effect Evaluation

[39] One of the strongest aspects of measurements made at ARM sites is the high quality of ground-based cloud and aerosol measurements that have been used in previous studies of the first aerosol indirect effect [e.g., *Feingold et al.*, 2003; *McComiskey et al.*, 2009]. These evaluations have been based on the representation of the albedo effect as the slope between the natural log of a cloud property (e.g., optical depth, droplet effective radius, and droplet number concentration) and an aerosol property (e.g., aerosol optical depth and aerosol number concentration). It is important to keep in mind that this relationship only holds for clouds of similar liquid water paths. One potential issue with this approach is that, while the cloud measurements are representative of the cloud in question, there is no guarantee that observed differences in the surface or column aerosol quantity used represent differences in the aerosol properties at the level of the cloud. For example, changes in aerosol optical depth may result from aerosol layers above the boundary layer, and there may be gradients in aerosol number concentration between the surface and cloud level. Assuming a well mixed boundary layer, however, seems that differences in surface aerosol properties should result in different cloud-height aerosol regimes.

[40] In the current study, ground-based retrieval of cloud droplet effective radius is obtained using cloud optical depth and liquid water path following the relationship from *Stephens* [1978]:

$$r_e = 1.5 \frac{\text{LWP}}{\tau_c} \quad (12)$$

where r_e is the column droplet effective radius (μm), LWP is the liquid water path (g m^{-2}) that is available from microwave radiometer (MWR) measurements. τ_c is the cloud optical depth and is available from a two-channel narrow-field-of-view radiometer (2NFOV) only for

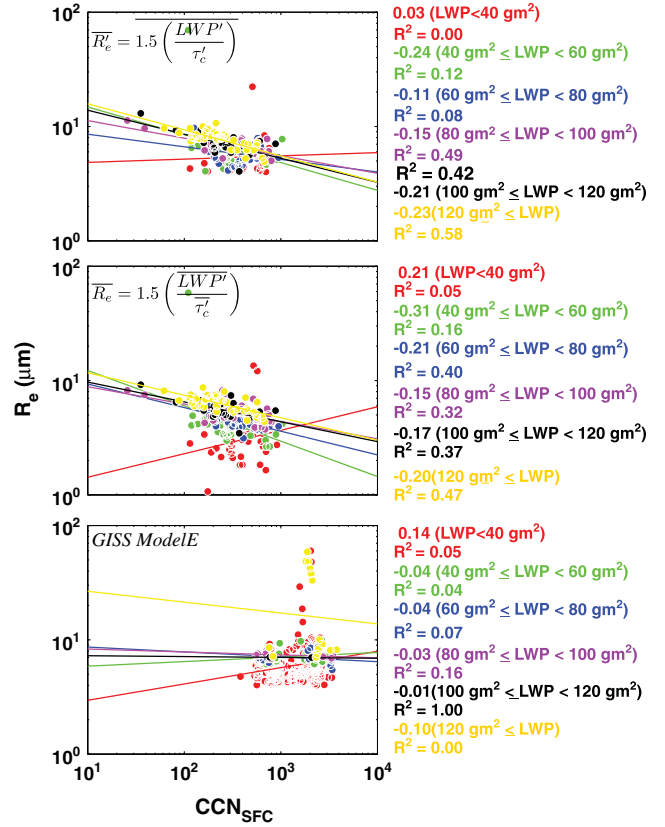


Figure 6. Slopes relating hourly average surface CCN concentration at 0.3% supersaturation and hourly average column cloud droplet effective radius as (top) calculated from instantaneous values of LWP and τ_c , (middle) calculated from hourly average values of LWP and τ_c and (bottom) by the GISS ModelE. Separate slopes have been calculated for various LWP ranges, as indicated by the list on the right.

the MASRAD deployment. For MASRAD, surface CCN concentrations are available from the ground-based aerosol observing system, providing the measurements necessary to derive the Indirect Effect [IE, *Feingold et al.*, 2003], or Aerosol-Cloud Interactions [ACI, *McComiskey et al.*, 2009] as defined by

$$\text{ACI}_r = - \left. \frac{\partial \ln r_e}{\partial \ln \text{CCN}} \right|_{\text{LWP}} \quad (13)$$

[41] To complete the evaluation, cloud cases are grouped into LWP bands of 20 g m^{-2} , and individual hourly averages of CCN concentration (0.3% supersaturation) and cloud column r_e are plotted in log-log space. Slopes are calculated for each LWP band. Observations of ACI are presented in two manners. The first (Figure 6 (top)) illustrates hourly average ACI values, as derived from retrievals of LWP and τ_c taken at the instrument temporal resolution. These values represent the average instantaneous ACI. The second estimate presented (Figure 6 (middle)) represents quantities calculated using the hourly averaged LWP and τ_c values. The latter may be considered to be more representative of the relationships derived in ModelE, due to the fact that model estimates will be derived using values of LWP and τ_c that are supposed to represent the spatial heterogeneity within the model grid box. These two techniques demonstrate qualitatively similar

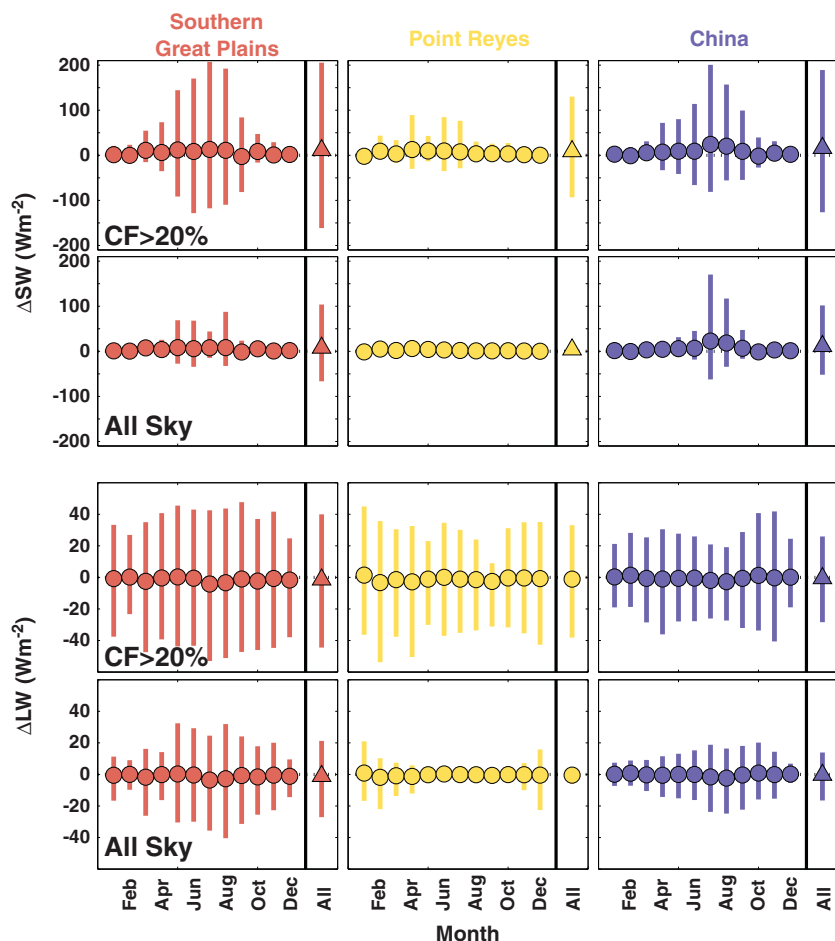


Figure 7. The impact of changes to the effective radius parameterization (DG minus LD) on broadband surface radiative flux densities at the Southern Great Plains (red), Point Reyes (yellow) and China (purple) sites. Included are (top) shortwave and (bottom) longwave evaluations for all-sky conditions and for when both simulations have cloud fractions exceeding 20%. Mean values are represented by triangles (when the difference is significant to 95% level using Wilcoxon rank sum test), or circles (if not significant), while the bars extend between the 10th to 90th percentiles of the data. Please note the different scales used for shortwave and longwave distributions.

results, with all cases featuring a negative slope (positive ACI), except those with LWP less than 40 g m^{-2} . Quantitatively, there are differences in the derived slopes, but results from the two sampling methods agree more closely with one another than either does with model results. All observational ACI values are larger than those derived from the model (Figure 6 (bottom)). With the observations, ModelE does not demonstrate a positive ACI for clouds with low LWP, but unlike the observations, this extends to clouds with LWP up to 60 g m^{-2} . Also, for cloud scenes with higher LWP, the model ACI is generally much lower (smaller slope) than observed. In order to most closely match the observations, the model was sampled only for times featuring single layer clouds and negligible precipitation for the MASRAD period. This distinction (negligible precipitation) is important because precipitation acts to reduce the degree to which a cloud is adiabatic, resulting in the breakdown of clean relationships between aerosol concentration and optical depth. Additionally, precipitation can result in scavenging of aerosol near the surface, resulting in skewed

estimates of ACI. These comparisons also clearly demonstrate the elevated aerosol concentrations found over coastal California within the model, with model values roughly an order of magnitude higher than those observed. It should be noted that the limited MASE data set results in some polynomial fits that are not necessarily representative of the data used to create them. R^2 values for the polynomial fits are provided and are generally very low, indicating that the lines provided to fit the data set may not provide a statistically significant representation of the individual points. Despite this, the sign of the slope is generally clear, except for the model results (Figure 6 (bottom)) and cases with very low LWP ($< 40 \text{ g m}^{-2}$).

4.4.4. Climatological Relevance

[42] As an example of the potential influence of small parameterization changes to simulation of global climate, simulations were completed implementing two separate effective radius parameterizations. The first simulation uses the *Del Genio et al.* [1996] as outlined in equation 7 (hereafter DG). The second simulation completed is identical

to the first, with the exception of the use of a different effective radius parameterization. This second simulation uses the *Liu and Daum* [2002] parameterization discussed above (hereafter LD), which accounts for possible spectral broadening.

[43] Over the same nudged 7 year period (2003–2009), using different effective radius parameterizations significantly impacts the simulation of the surface radiative energy budget. Figure 7 illustrates differences (DG minus LD) in net (top) shortwave and (bottom) longwave radiation in W m^{-2} at the Earth's surface between the two simulations for the three measurement sites. A Wilcoxon rank sum test was used to evaluate the statistical significance of these differences. Mean differences statistically significant to the 95% level are represented using triangles. Mean differences that are not found to be statistically significant are illustrated using circles, while the bars represent the 10th to 90th percentiles of the data sets. Distributions are included for cases where cloud area fraction exceeds 20% in both simulations (top and third rows for shortwave and longwave, respectively), and for all times in the simulations.

[44] The largest changes occur during summer months in the shortwave, with monthly mean changes of 24.0 W m^{-2} for China in July. Summertime extreme cases (within the 10th/90th percentile envelope) reach values around 200 W m^{-2} . Under all-sky conditions, differences are smaller, but not directly comparable because here the frequency of cloud occurrence from one site to the next influences these differences as well as the changes to the clouds themselves. Even here, however, there are differences in the shortwave that can dramatically impact the surface radiation balance. An example of this are the summer months at the China site, where mean differences in shortwave flux density reach as high as 22.7 W m^{-2} . The influence of effective radius parameterization on surface longwave radiation is substantially smaller for the sites evaluated in the current study, but can still not be considered negligible, with extreme values between $20\text{--}40 \text{ W m}^{-2}$. The role of parameterization choice on longwave radiation would be more significant at higher latitudes where the longwave influence of clouds dominates for much of the year due to low sun angles.

[45] In cloudy regions, the differences demonstrated here can have significant consequences on climate. In the polar regions, for example, changes to the surface energy budget could have large impacts on the melt rates of land and sea ice. With the evaluation of effective radius parameterizations (Figure 5) revealing no clear “best” parameterization across the campaigns used in the evaluation, it is somewhat concerning that choosing one versus another results in such large changes.

5. Discussion and Summary

[46] In this study measurements from the US DOE ARM program are used to evaluate the NASA GISS ModelE's representation of processes linked to the interactions between clouds and aerosols. These observations include samples from a variety of geographical locations and weather regimes while the model was run in a nudged mode to constrain the meteorology to be representative of what was observed. Measurements are scaled to capture a portion of

the variability that may be expected across the $2^\circ \times 2.5^\circ$ grid size of the model.

[47] On average, the basic meteorological conditions in the model match those observed. The main exceptions to this are the coastal California measurement locations, where spatial inhomogeneity in surface and elevation make it challenging for the model to reproduce the measurements exactly. This results in biases in surface air temperature, pressure, and relative humidity. Smaller biases are observed at the Southern Great Plains site during the Aerosol IOP. Aerosol properties are less constrained in the model, with large differences in surface CCN concentration and aerosol optical depth. This reflects a general difficulty in using field measurements to evaluate climate models that resolve atmospheric chemistry. Even though the nudging of meteorological fields to reanalysis data sets allow a decent simulation of the observed meteorology, chemical weather can not be reproduced by such simulations. Emission inventories for global climate models only provide decadal mean emission information [e.g., *Lamarque et al.*, 2010], and those are insufficient to provide detailed information at the local level. Therefore, evaluation of systematic model biases is appropriate, while time-by-time evaluations are not.

[48] Evaluation of aerosol-cloud interactions include those of droplet activation, effective radius parameterizations, and the relationship between surface CCN concentrations and the cloud column effective droplet size. Because of the errors in the simulated aerosol concentrations, these evaluations are completed in a way that evaluates the relationships between variables, rather than a direct correlation of them with observations. Looking at droplet activation, for all campaigns with in situ aircraft measurements available, the model tends to activate more droplets than those observed for a given number of aerosols. Contrary to what was observed at the surface, the model appears to produce less aerosol particles at higher altitudes than observed. This discrepancy likely helps to explain the under prediction of AOD by the model, while it simultaneously over predicts surface CCN concentrations. The cause of exaggerated droplet activation is not explored here, but the resulting clouds are likely too reflective for a given amount of aerosol. Increased droplet activation also will work to decrease droplet sizes in the model, ultimately resulting in impacts on precipitation production and other processes related to droplet size.

[49] To evaluate parameterizations of cloud droplet effective radius, measurements from campaigns featuring airborne in situ sampling were used directly. Evaluations demonstrated that there does not appear to be a specific parameterization that clearly outperforms the others for a diverse set of environmental conditions. The parameterization that accounts for spectral broadening outperformed the more primitive parameterizations for the Aerosol IOP, but did not fare as well for MASE and RACORO flights. This inconsistency is troubling when considering the potentially large influence that a choice in this parameterization can have on the surface energy budget. This influence was demonstrated to be as large as 22.7 W m^{-2} , with larger values occurring for individual years and still larger differences when only cloudy conditions were considered.

[50] Finally, when examining the relationship between surface CCN concentration and cloud effective droplet size

(ACI), the model does not simulate as strong a relationship as seen in observations. While observed ACI values generally range between 0.11 and 0.24, depending on the LWP values, the simulated values are generally close to zero, with ACI values between -0.14 and 0.09 . While some of this may be the result of difficulties in accurately sampling the observations to be representative of the climate model scale, the fact that two different sampling methods produce similar results hints at the possibility that the model is simply demonstrating a different sensitivity. Part of this may result from the elevated surface aerosol concentrations in the model, which, as pointed out earlier, do not appear to extend to cloud altitude. This disconnect would imply that despite differences in clouds resulting from aerosol-related processes (e.g., activation), the surface aerosol concentration does not produce an associated change in aerosol at higher altitudes, which would act to change the slope to smaller values as demonstrated.

[51] At this point, the GISS ModelE parameterizations used to relate aerosol properties to clouds struggle to accurately representing measured relationships for the campaigns discussed here. Whether detected differences between simulations and observations are the result of physics, chemistry, or otherwise, remains to be answered. Evaluation of interactions between system components, such as those between aerosols and clouds, can help to illuminate potential weaknesses in the model that may have been hidden in previous evaluations. As an example, while work by Koffi *et al.* [2012] demonstrated promising agreement between ModelE-produced and measured vertical distribution of aerosol, this evaluation used optical properties such as extinction and AOD and did not directly evaluate number density, a parameter critical for cloud processes. The current study demonstrates a potential disconnect between the vertical distribution of aerosol concentration between the model and observations. Continued development of both modeling tools and observational data sets will help to make more thorough comparisons possible. In general, evaluation of these processes in climate models remains at an infant state. Observational campaigns designed with the climate model scale in mind may help to more closely monitor relevant processes without the need for temporal averaging of single point measurements to attempt to account for spatial and temporal inhomogeneity. Additionally, improvements to satellite sensors and retrieval algorithms, in particular, continued work on active remote sensors necessary for high-resolution cloud and aerosol measurements, will allow for significant advancement, but only if those measurements can be collocated with reliable information on atmospheric state variables. Diverse observational records covering extended time periods, such as those recorded by the ARM program, remain critical to evaluation of models due to the large range of environmental conditions covered. At the same time, parallel efforts focusing on the use of satellite-based measurements can complement these localized studies by providing a more global perspective at scales comparable to those of the GCM grid box. With GCM grid-box scales becoming smaller and observational tools advancing, it is the hope of the authors that both surface and space-born observations can continue to expand in space and time and improve in accuracy and detail to best meet the needs of those involved with climate model development.

[52] **Acknowledgments.** This research was supported by the Director, Office of Science, Office of Biological and Environmental Research of the U.S. Department of Energy under Contract DE-AC02-05CH11231 as part of their Climate and Earth System Modeling Program and through the FASTER project. LBNL is managed by the University of California under the same grant. This work was prepared in part at the Cooperative Institute for Research in Environmental Sciences (CIRES) with support in part from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce, under cooperative agreement NA17RJ1229 and other grants. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration or the Department of Commerce. GB was supported in part by the National Science Foundation (ARC-1203902) and US Department of Energy (DE-SC0008794). Computing resources were provided by NASA and the US Department of Energy. A.V. wishes to acknowledge funding from the U.S. DOE (contract DE-AC02-98CH10886). 2NFOV retrievals were generously provided by Christine Chiu, and China AMF data were provided by Maureen Cribb and Zanqing Li. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center. Data were obtained from the Atmospheric Radiation Measurement (ARM) Program sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Climate and Environmental Sciences Division.

References

- Abdul-Razzak, H., and S. Ghan (2000), A parameterization of aerosol activation. Part 2: Multiple aerosol type, *J. Geophys. Res.*, *105*, 6837–6844, doi:10.1029/1999JD901161.
- Albrecht, B. (1989), Aerosols, cloud microphysics, and fractional cloudiness, *Science*, *245*, 1227–1230, doi:10.1126/science.245.4923.1227.
- Bauer, S., and S. Menon (2012), Aerosol direct, indirect, semi-direct and surface albedo effects from sector contributions based on the IPCC AR5 emissions for pre-industrial and present day conditions, *J. Geophys. Res.*, *117*, D01206, doi:10.1029/2011JD016816.
- Bauer, S., D. Wright, D. Koch, E. Lewis, R. McGraw, L.-S. Chang, S. Schwartz, and R. Ruedy (2008), MATRIX (Multiconfiguration Aerosol TRacker of mIXing state): An aerosol microphysical module for global atmospheric models, *Atmos. Chem. Phys.*, *8*, 6003–6035, doi:10.5194/acp-8-6003-2008.
- Bauer, S., S. Menon, D. Koch, T. Bond, and K. Tsigaridis (2010), A global modeling study on carbonaceous aerosol microphysical characteristics and radiative forcing, *Atmos. Chem. Phys.*, *10*, 7439–7456, doi:10.5194/acp-10-7439-2010.
- Berg, L., C. Berkowitz, J. Barnard, G. Senum, and S. Springston (2011), Observations of the first aerosol indirect effect in shallow cumuli, *Geophys. Res. Lett.*, *38*, L03809, doi:10.1029/2010GL046047.
- Bower, K., and T. Choulaton (1992), A parameterization of the effective radius of ice-free clouds for use in global climate models, *Atmos. Res.*, *27*, 305–339, doi:10.1016/0169-8095(92)90038-C.
- Chen, Y., and J. Penner (2005), Uncertainty analysis for estimates of the first indirect aerosol effect, *Atmos. Chem. Phys.*, *5*, 2935–2948, doi:10.5194/acp-5-2935-2005.
- Chiu, J., A. Marshak, Y. Knyazikhin, W. Wiscombe, H. Barker, J. Barnard, and Y. Luo (2006), Remote sensing of cloud properties using ground-based measurements of zenith radiance, *J. Geophys. Res.*, *111*, D16201, doi:10.1029/2005JD006843.
- Del Genio, A., M.-S. Yao, W. Kovari, and K.-W. Lo (1996), A prognostic cloud water parameterization for global climate models, *J. Clim.*, *9*, 270–304, doi:10.1175/1520-0442(1996)009<0270:APCWPF>2.0.CO;2.
- Feingold, G., W. Eberhard, D. Veron, and M. Previdi (2003), First measurements of the Twomey indirect effect using ground-based remote sensors, *Geophys. Res. Lett.*, *30*(6), 1287, doi:10.1029/2002GL016633.
- Ghan, S., L. Leung, R. Easter, and H. Abdul-Razzak (1997), Prediction of droplet number in a general circulation model, *J. Geophys. Res.*, *102*, 21,777–21,794, doi:10.1029/97JD01810.
- Goldsmith, J., F. Blair, S. Bisson, and D. Turner (1998), Turn-key raman lidar for profiling atmospheric water vapor, clouds, and aerosols, *Appl. Opt.*, *37*, 4979–4990, doi:10.1364/AO.37.004979.
- Hansen, J., et al. (2007), Climate simulations for 1880–2003 with GISS modelE, *Clim. Dyn.*, *29*, 661–696, doi:10.1007/s00382-007-0255-8.
- Harrison, L., and J. Michalsky (1994), Objective algorithms for the retrieval of optical depths from ground-based measurements, *J. Appl. Opt.*, *22*, 5126–5132, doi:10.1364/AO.22.005126.
- Intergovernmental Panel on Climate Change (2007), *Climate Change 2007: The Physical Science Basis*. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change, edited by S. Solomon et al., 996 pp., Cambridge Univ. Press, Cambridge, U. K.

- Jefferson, A. (2011), Aerosol Observing System (AOS) handbook, *Tech. Rep. ARM-TR-014*, U.S. Dep. of Energy, Washington, D. C.
- Koffi, B., et al. (2012), Application of the CALIOP layer product to evaluate the vertical distribution of aerosols estimated by global models: AeroCom phase I results, *J. Geophys. Res.*, *117*, D10201, doi:10.1029/2011JD016858.
- Koren, I., Y. Kaufman, D. Rosenfeld, L. Remer, and Y. Rudich (2005), Aerosol invigoration and restructuring of Atlantic convective clouds, *Geophys. Res. Lett.*, *32*, L14828, doi:10.1029/2005GL023187.
- Lamarque, J.-F., et al. (2010), Historical (1850-2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application, *Atmos. Chem. Phys.*, *10*, 7017–7039, doi:10.5194/acp-10-7017-2010.
- Liu, Y., and P. Daum (2002), Indirect warming effect from dispersion forcing, *Nature*, *419*, 580–581, doi:10.1038/419580a.
- Lu, M.-L., W. Conant, H. Jonsson, V. Varutbangkul, R. Flagan, and J. Seinfeld (2007), The Marine Stratus/Stratocumulus Experiment (MASE): Aerosol-cloud relationships in marine stratocumulus, *J. Geophys. Res.*, *112*, D10209, doi:10.1029/2006JD007985.
- Martin, G., D. Johnson, and A. Spice (1994), The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds, *J. Atmos. Sci.*, *51*, 1823–1842, doi:10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2.
- Mather, J., and J. Voyles (2013), The ARM climate research facility: A review of structure and capabilities, *Bull. Am. Meteorol. Soc.*, *94*, 377–392, doi:10.1175/BAMS-D-11-00218.1.
- McComiskey, A., and G. Feingold (2012), The scale problem in quantifying aerosol indirect effects, *Atmos. Chem. Phys.*, *12*, 1031–1049, doi:10.5194/acp-12-1031-2012.
- McComiskey, A., G. Feingold, A. Frisch, D. Turner, M. Miller, J. Chiu, Q. Min, and J. Ogren (2009), An assessment of aerosol-cloud interactions in marine stratus clouds based on surface remote sensing, *J. Geophys. Res.*, *114*, D09203, doi:10.1029/2008JD011006.
- Menon, S., and L. Rotstayn (2006), The radiative influence of aerosol effects on liquid-phase cumulus and stratiform clouds based on sensitivity studies with two climate models, *Clim. Dyn.*, *27*, 345–356, doi:10.1007/s00382-006-0139-3.
- Menon, S., A. Del Genio, Y. Kaufman, R. Bennartz, D. Koch, N. Loeb, and D. Orlikowski (2008), Analyzing signatures of aerosol-cloud interactions from satellite retrievals and the GISS GCM to constrain the aerosol indirect effect, *J. Geophys. Res.*, *113*, D14S22, doi:10.1029/2007JD009442.
- Menon, S., D. Koch, G. Beig, S. Sahu, J. Fasullo, and D. Orlikowski (2010), Black carbon aerosols and the third polar ice cap, *Atmos. Chem. Phys.*, *10*, 4559–4571, doi:10.5194/acp-10-4559-2010.
- Morrison, H., and A. Gettelman (2008), A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model (CAM3). Part I: Description and numerical tests, *J. Clim.*, *21*, 3642–3659, doi:10.1175/2008JCLI12105.1.
- Nenes, A., and J. Seinfeld (2003), Parameterization of cloud droplet formation in global climate models, *J. Geophys. Res.*, *108*(D14), 4415, doi:10.1029/2002JD002911.
- Quaas, J., et al. (2009), Aerosol indirect effects—General circulation model intercomparison and evaluation with satellite data, *Atmos. Chem. Phys.*, *9*, 8697–8717, doi:10.5194/acp-9-8697-2009.
- Roelofs, G., J. Lelieveld, and L. Ganzeveld (1998), Simulation of global sulfate distribution and the influence on effective cloud drop radii with a coupled photochemistry-sulfur cycle model, *Tellus, Ser. B*, *50*, 224–242, doi:10.1034/j.1600-0889.1998.t01-2-00002.x.
- Rosenfeld, D., and G. Feingold (2003), Explanation of discrepancies among satellite observations of the aerosol indirect effects, *Geophys. Res. Lett.*, *30*(14), 1776, doi:10.1029/2003GL017684.
- Schmidt, G., et al. (2006), Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data, *J. Climate*, *19*, 153–192, doi:10.1175/JCLI3612.1.
- Sekiguchi, M., T. Nakajima, K. Suzuki, K. Kawamoto, A. Higurashi, D. Rosenfeld, I. Sano, and S. Mukai (2003), A study of the direct and indirect effects of aerosols using global satellite data sets of aerosol and cloud parameters, *J. Geophys. Res.*, *108*(D22), 4699, doi:10.1029/2002JD003359.
- Shao, H., and G. Liu (2009), A critical examination of the observed first aerosol indirect effect, *J. Atmos. Sci.*, *66*, 1018–1032, doi:10.1175/2008JAS2812.1.
- Sivaraman, C., D. Turner, and C. Flynn (2004), Techniques and methods used to determine the Aerosol Best Estimate value-added product at SGP central facility, paper presented at Fourteenth ARM Science Team Meeting, *Atmos. Radiat. Measure. Clim. Res. Fac.*, Albuquerque, N. M., 22–26 March.
- Slingo, A. (1989), A GCM parameterization for the shortwave radiative properties of water clouds, *J. Atmos. Sci.*, *46*, 1419–1427, doi:10.1175/1520-0469(1989)046<1419:AGPFTS>2.0.CO;2.
- Stephens, G. (1978), Radiation profiles in extended water clouds I: Theory, *J. Atmos. Sci.*, *35*, 2111–2122, doi:10.1175/15200469.
- Suzuki, K., T. Nakajima, A. Numaguti, T. Takemura, K. Kawamoto, and A. Higurashi (2004), A study of the aerosol effect on a cloud field with simultaneous use of GCM modeling and satellite observation, *J. Atmos. Sci.*, *61*, 179–193, doi:10.1175/1520-0469(2004)061<0179:ASOTAE>2.0.CO;2.
- Turner, D., S. Clough, J. Liljegren, E. Clothiaux, K. Cady-Pereira, and K. Gaustad (2007), Retrieving liquid water path and precipitable water vapor from the Atmospheric Radiation Measurement (ARM) microwave radiometers, *IEEE Trans. Geosci. Remote Sens.*, *45*, 3680–3690, doi:10.1109/TGRS.2007.903703.
- Twohy, C., M. Petters, J. Snider, B. Stevens, W. Tahnk, M. Wetzell, L. Russel, and F. Burnet (2005), Evaluation of the aerosol indirect effect in marine stratocumulus clouds: Droplet number, size, liquid water path, and radiative impact, *J. Geophys. Res.*, *110*, D08203, doi:10.1029/2004JD005116.
- Twomey, S. (1977), The influence of pollution on the shortwave albedo of clouds, *J. Atmos. Sci.*, *34*, 1149–1152, doi:10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2.
- Vogelman, A., et al. (2012), RACORO extended-term aircraft observations of boundary layer clouds, *Bull. Am. Meteorol. Soc.*, *93*, 861–878, doi:10.1175/BAMS-D-11-00189.1.