Supporting Information

Vergara-Temprado et al. 10.1073/pnas.1721627115

SI Model Description

The Unified Model (version 10.3) (32) is used in its global setup (GA6, N512 resolution) to provide the initial and boundary conditions for the high-resolution regional simulations. We use two different high-resolution nests: the coarser domain (Fig. 3 *C–D*) is composed of 600 × 600 grid boxes with a grid spacing of 0.07° (rotated grid), and the higher-resolution nest (Figs. 3 *G–H*, 4, and 5) is composed of 500 × 500 grid boxes with a spacing of 0.02°. Both domains have a vertical resolution of 70 height levels quadratically spaced up to 40 km (116-m grid box height at 1 km). We perform the simulations for a total of 18 h, and therefore, we can perform a satellite evaluation with the A-train satellites in all of our simulations. To avoid influences from boundary effects, we eliminate the closest 100 grid boxes to the boundaries in all of the 0.02°-resolution simulations before analyzing them.

Cloud microphysical processes are simulated by using Cloud AeroSol Interactive Microphysics (29, 30, 39). It represents cloud droplets, rain, ice crystals, snow, and graupel and uses a double-moment scheme that predicts both number and mass of each of the hydrometeor types. The hydrometeor size distributions are represented by gamma function with a fixed width.

An aerosol vertical profile obtained from a GLOMAP simulation (40) is used to initiate the aerosol fields and to feed the boundary conditions in the high-resolution domains. The profile is obtained from the mean values during the summer season in a South Atlantic transect (located in 40° to 70° S, 20° W). It is composed of three soluble modes (Aitken, accumulation, and coarse modes) and two insoluble modes (accumulation and coarse) that each follow a log-normal distribution. The two insoluble modes are used to represent the dust concentrations simulated by GLOMAP that are necessary to calculate INP when using the DM15 parameterization. Aerosol particles are subject to horizontal and vertical advection; however, they are not modified by the actions of cloud feedbacks, and therefore, their distribution is similar to the initial profile across the entire domain. We use the activation scheme of Abdul-Razzak and Ghan (41) to calculate the number of activated cloud droplets, which depends on the aerosol distribution and the cloud updraft. The model produces reasonable values of cloud droplet number concentrations compared with satellite observations (see below).

The model used for predicting INP concentrations in this region of the SO (VT17) uses feldspar and marine organic aerosols for predicting the global distribution and temporal evolution of INP following the method shown in the work by Vergara-Temprado et al. (24). Briefly, feldspar particles are emitted as a fraction of the dust in the accumulation and coarse insoluble modes and then moved to the soluble modes by atmospheric aging. The INP contribution from feldspar particles is calculated by assuming that 35% of all feldspar is potassium feldspar and then applying the parameterization shown in the work by Atkinson et al. (42). Marine organic aerosols are emitted as a fraction of submicrometer sea spray particles internally mixed into the accumulation soluble mode. Their contribution to INP concentrations is calculated using the parameterization shown in the work by Wilson et al. (43). The model has been shown to improve the prediction of INP from previous representation. Specifically, it produces a good agreement against observation in marine regions (~72% of data points agree within an order of magnitude) (24). The datasets shown in Fig. 2*A* are presented in the work by Vergara-Temprado et al. (24) and are made up of data from a variety of campaigns in contrasting locations around the world (19, 44–51).

Autoconversion of cloud droplets to rain and droplet accretion follow the work by Khairoutdinov and Kogan (52), and self-collection of rain is parameterized based on the work by Beheng (53). The mass–fall speed relation for graupel follows the work by Locatelli and Hobbs (54), and the graupel density is set at 250 kg m^{-3} (30). The mass–fall speed and mass–diameter relations used for all of the hydrometeors are described in the work by Miltenberger et al. (30). The primary production of ice is defined by the parameterizations described in the text (Fig. 1). The Hallet–Mossop secondary ice production process is represented by producing 350 ice splinters per 1 mg of rimed mass on snow or graupel at a temperature of −5 °C and a linearly decreasing rate to zero at −2.5 °C and −7.5 °C. Freezing of rain drops follows the work of Bigg (55). Switching on and off the Hallet–Mossop process and rain droplet freezing did not play a substantial role in the properties of our studied clouds. Other processes affecting the transfer rates between hydrometers include vapor deposition, evaporation, sublimation, collision coalescence, and sedimentation. The microphysics scheme currently does not include any subgrid treatment of the partitioning between ice and liquid water in mixed-phase grid boxes. The sensitivity to this treatment in high-resolution simulations of the same cloud type has been shown to be much smaller [around 8% increase in reflected SW (17)] than the sensitivities to changes in the representation of INP presented in this paper.

In Fig. 4, we calculated the distribution of INP concentrations at cloud temperature to estimate the concentration of INPs affecting the clouds. We do this by filtering out the grid boxes with a total water mass mixing ratio less than 10^{-6} (residual water amount) (17). We then calculate the distribution of INPs using the different parameterizations in the various simulations combined with the temperature of the grid boxes (in-cloud INP). These values represent the INP concentration affecting the different cloudy grid boxes. From the distribution of INP values calculated, we then obtain the median and the 66% and 95% intervals of the distribution shown in Fig. 4.

Satellite Data and Model Evaluation

The model has been evaluated with several satellite products from the A-train constellation. The simulated values (output every hour) were interpolated to the time when the A train passes through the model domains.

The radiative properties (outgoing SW and longwave radiation) were obtained from the NASA CERES (33) satellite instrument. Data from the Moderate Resolution Imaging Spectroradiometer (MODIS Collection 6, Level 2 data) (56) mounted on the Aqua satellite were used to compare against cloud-top temperatures (CTTs) and cloud-top phase. Observed cloud liquid water path (LWP) was obtained from the Advanced Microwave Scanning Radiometer 2 (AMSR2) (57). For simulated case C2, due to the small scale of the cumulus clouds composing the cloud system, a comparison with MODIS cloud LWP was used instead of AMSR2, as it provides a higher-resolution product being able to resolve the scale of the clouds formed. We note that the observed subdomain mean LWP using the microwave retrieval (AMRS2) for this cloud is 28% lower than the MODIS estimate (0.068 mm for AMSR2 and 0.094 mm with MODIS).

The distributions of CTTs) for the three studied clouds are shown in Fig. S1, *Top*. Overall, the model creates clouds with similar CTTs, although for C2, it seems to miss some warm clouds above 265 K. Changing the INP parameterization does not change CTT substantially, although the M92 parameterization (high INP) tends to produce higher-temperature cloud tops as the higher nucleation events deplete the top (colder) part of the cloud.

The distribution of LWP has a similar behavior as the distributions of reflected SW. The models with low INP representations produce distributions much closer to the satellite observations than the global model and the high INP representation (M92), which produce too few grid boxes with more than 0.1 mm LWP.

The simulated longwave outgoing flux is very close to the satellite observations (Fig. S2) for all of the cases studied, including the global model. Modifying the INP parameterization makes a very small change on the longwave radiative properties compared with the change observed in SW (Fig. 3).

Cloud-top phase from the model was calculated as follows. First, we filter all of the grid boxes with water (ice or liquid) mass mixing ratios less than 10^{-6} to exclude grid boxes with residually small amounts of water. Second, we obtain the cloud-top height for water and ice clouds. If the top of the ice cloud is in a grid box above the top of the liquid cloud, we consider that column to have an ice top, and conversely, we consider columns with a higher liquid cloud top to be liquid-topped. We consider as mixed-phase/uncertain cloud-top phase the columns where the liquid and ice cloud top are in the same grid box. The derived cloud-top phase is then compared against the MODIS optical and IR cloud-top products (Fig. S3). The global model and the M92 simulations produce too little liquid-phase cloud compared with the satellite products. The comparison improves greatly when the other INP parameterizations are used, producing liquid cloud fractions that are much closer than (or in between) the two retrieved estimates.

Satellite cloud droplet number concentrations (CDNCs) are derived from MODIS Collection 6, Level 2 swath data using 1-km (at nadir) resolution pixel-level values of cloud liquid effective radius, cloud optical depth, and CTT data using the method described in the work by Grosvenor and Wood (58), although with some differences related to the use of Collection 6 rather than Collection 5 data. Namely, in Collection 6, pixel-level retrieval confidence quality assurance (QA) flags are no longer used, and also, here we only examine pixel-level data; therefore, the restrictions applied when aggregating to lower resolution are not necessary. Pixels were filtered to include only liquid water pixels that were diagnosed as "confident cloudy" and as not being affected by any thin cirrus on top or shadowing by the MODIS algorithm. Only pixels that had an optical depth larger than five are included, since biases are likely in thinner clouds (59, 60). Furthermore, pixels were required to have a cloud-top height (derived from the 5-km MODIS product) between 0.5 and 3.2 km; the lower limit is imposed to remove pixels that are low-level fog or where the height retrieval is erroneous, and the upper limit restricts the analysis to low-altitude clouds that are most likely to meet the assumptions made for the derivation of CDNC. The maximum solar zenith angle for the swath used in this analysis was 59°; therefore, biases due to high solar zenith angles are unlikely, since these have been shown to begin for angles larger than 65° to 70° (58).

Domain mean CDNC values in cloudy columns with low-level liquid-containing clouds for the different simulations and the satellite-derived values are shown in Fig. S4. The mean simulated values are close to the satellite values for the first and second clouds (C1 and C2). The simulations of C3 produce values that are relatively lower than the satellite values. To test the importance of this bias, we repeated one of the simulations (C3_M92) with a higher aerosol concentration (with around 100 cm^{-3} in the accumulation mode as opposed to around 35 cm^{-3} used previously). In this simulation, CDNC is slightly above the satellite-derived observations; however, the radiative properties of the cloud do not change substantially (Fig. S5).

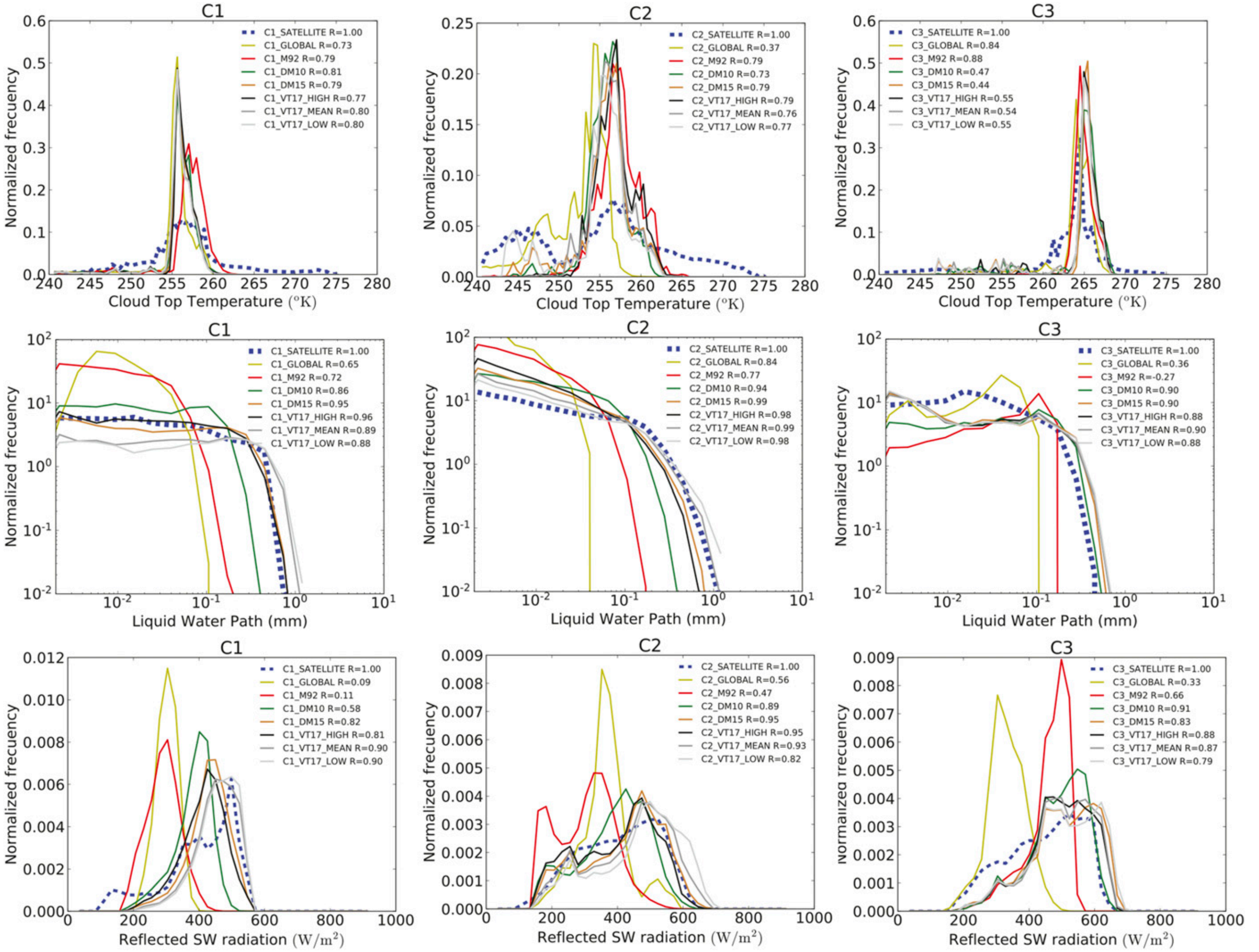

Fig. S1. Frequency distributions. (*Top*) Distribution of frequencies of CTTs for the three studied cloud systems derived from MODIS-retrieved CTT and the model simulations with the different INP parameterizations. (*Middle*) Same but for LWP retrieved with AMSR2 (MODIS for C2). (*Bottom*) Same but for SW retrieved with CERES. The *R* coefficients refer to the correlation between the satellite PDF and the modeled values.

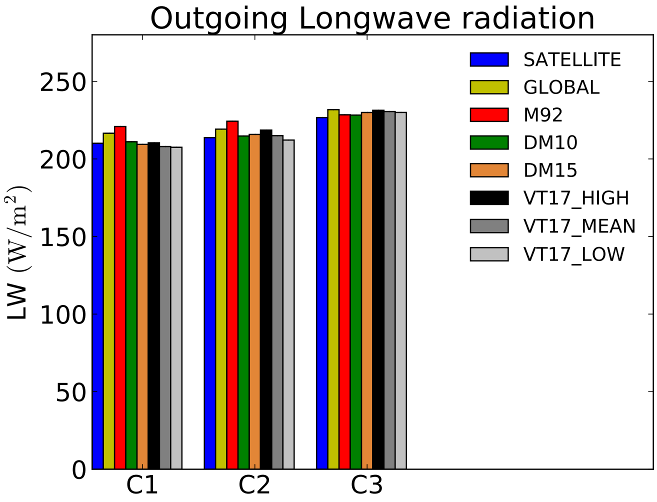

Fig. S2. Bar plot of subdomain-averaged outgoing longwave (LW) radiation (similar to Fig. 2*A* but for LW instead of SW).

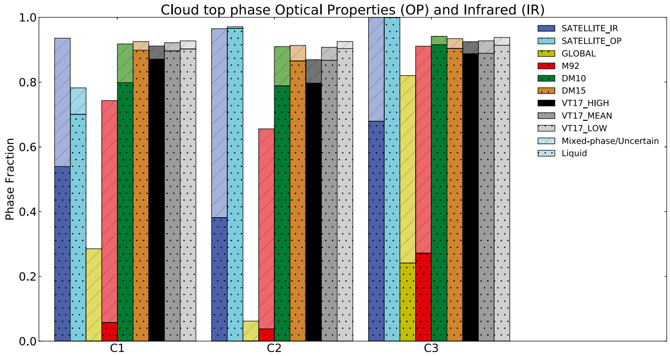

Fig. S3. Cloud-top phase comparison. Runs for the clouds studied with the different INP representations. The satellite columns correspond to the MODIS IR and optical (OP) cloud-top phase product.

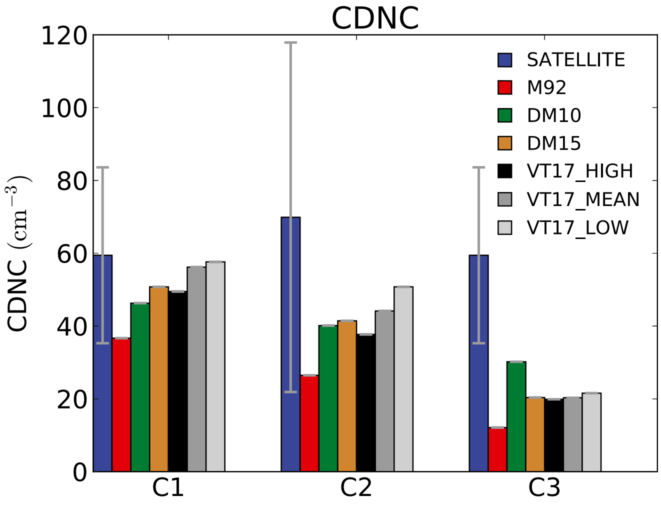

Fig. S4. Satellite-derived CDNC and domain mean simulated CDNC. The error bars represent two SDs of the distribution of satellite-calculated CDNC.

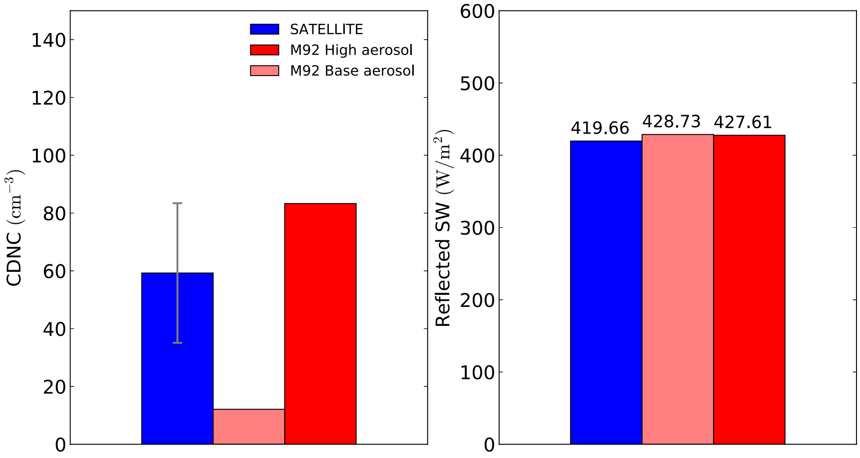

Fig. S5. CDNC sensitivity test. Domain mean cloud droplet number concentration (*Left*) and reflected SW radiation (*Right*) for the simulations using M92 with a high aerosol loading and the standard aerosol profile from GLOMAP used in all of the other simulations. The satellite observations are shown for comparison.

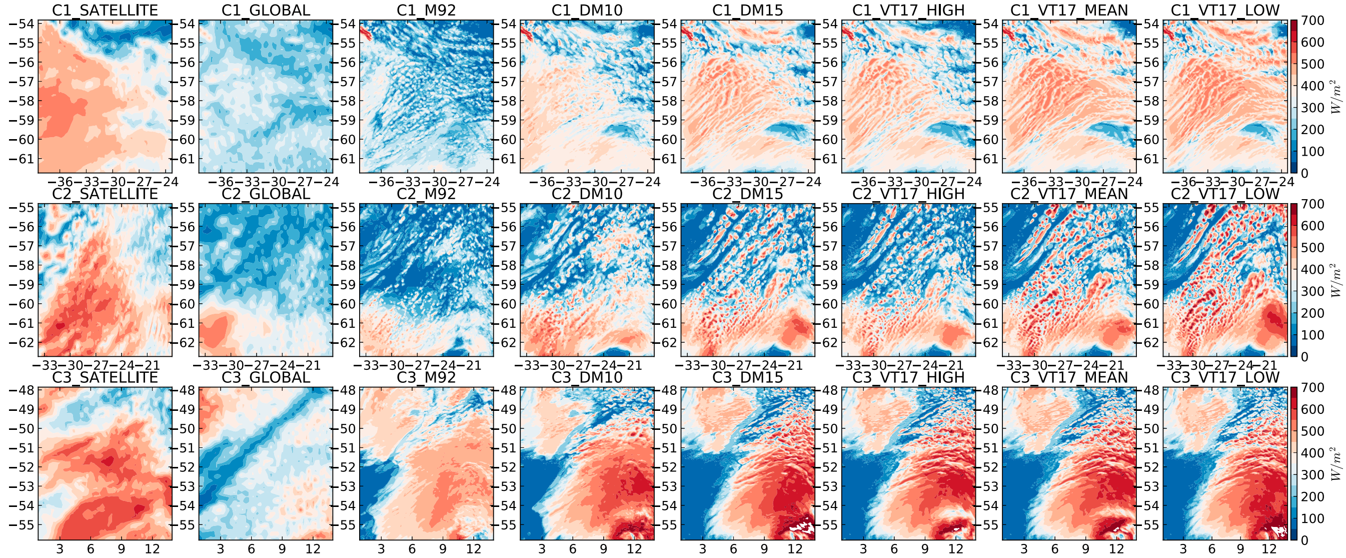

Fig. S6. Subdomain reflected SW radiation for all of the 0.02°-resolution simulations. Each row corresponds to each of the different clouds studied (C1–C3 in *Top*, *Middle*, and *Bottom*, respectively). Column 1 corresponds to the satellite observations, and columns 2–8 correspond to the global model and the 0.02°-resolution simulations with the different INP parameterizations tested (following the naming previously defined in the text).

Table S1. **Time and location of the simulated clouds systems**

Cloud	Center of domain	Date	Time that the A train passes
C1	58° S, 30° W	March 1, 2015	16:30 UTC
C2	59° S, 25° W	January 10, 2015	16:40 UTC
C3	52° S, 8° E	December 9, 2014	13:25 UTC

UTC, Coordinated Universal Time.