# Statistical Analyses of Satellite Cloud Object Data from CERES. Part I: Methodology and Preliminary Results of the 1998 El Niño/2000 La Niña

KUAN-MAN XU, TAKMENG WONG, AND BRUCE A. WIELICKI

*NASA Langley Research Center, Hampton, Virginia*

LINDSAY PARKER

*Science Applications International Corporation, Hampton, Virginia*

ZACHARY A. EITZEN

*Colorado State University, Fort Collins, Colorado, and NASA Langley Research Center, Hampton, Virginia*

(Manuscript received 7 June 2004, in final form 19 October 2004)

ABSTRACT

This study presents an objective classification methodology that uses Earth Observing System (EOS) satellite data to classify distinct "cloud objects" defined by cloud-system types, sizes, geographic locations, and matched large-scale environments. This analysis method identifies a cloud object as a contiguous region of the earth with a single dominant cloud-system type. It determines the shape and size of the cloud object from the satellite data and the cloud-system selection criteria. The statistical properties of the identified cloud objects are analyzed in terms of probability density functions (PDFs) based upon the Clouds and the Earth's Radiant Energy System (CERES) Single Satellite Footprint (SSF) data.

Four distinct types of oceanic cloud objects—tropical deep convection, boundary layer cumulus, transition stratocumulus, and solid stratus—are initially identified from the CERES data collected from the Tropical Rainfall Measuring Mission (TRMM) satellite for this study. Preliminary results are presented from the analysis of the grand-mean PDFs of these four distinct types of cloud objects associated with the strong 1997/98 El Niño in March 1998 and the very weak 2000 La Niña in March 2000. A majority of the CERES footprint statistical characteristics of observed tropical deep convection are similar between the two periods in spite of the climatological contrast. There are, however, statistically significant differences in some cloud macrophysical properties such as the cloud-top height and cloud-top pressure and moderately significant differences in outgoing longwave radiation (OLR), cloud-top temperature, and ice diameter. The footprint statistical characteristics of the three observed boundary layer cloud-system types are distinctly different from one another in all cloud microphysical, macrophysical, optical properties, and radiative fluxes. The differences between the two periods are not significant for most cloud microphysical and optical properties and the top-of-the-atmosphere albedo, but are statistically significant for some cloud macrophysical properties and OLR. These characteristics of the grand-mean PDFs of cloud microphysical, macrophysical, and optical properties and radiative fluxes can be usefully compared with cloud model simulations.

Furthermore, the proportion of different boundary layer cloud types is changed between the two periods in spite of small differences in their grand-mean statistical properties. An increase of the stratus population and a decrease of the cumulus population are evident in the El Niño period compared to the very weak La Niña period. The number of the largest tropical convective cloud objects is larger during the El Niño period, but the total number of tropical convective cloud objects is approximately the same in the two periods.

*Corresponding author address:* Dr. Kuan-Man Xu, Climate Science Branch, NASA Langley Research Center, Mail Stop 420, Hampton, VA 23681.
E-mail: Kuan-Man.Xu@nasa.gov

© 2005 American Meteorological Society

## 1. Introduction

The most recent Intergovernmental Panel on Climate Change (Houghton et al. 2001) report states that "[a]s has been the case since the first IPCC Assessment Report in 1990, probably the greatest uncertainty in future projections of climate arises from clouds and their interactions with radiation." This statement suggests that the roles of clouds in the climate system are very complicated, and rapid progress in understanding their roles has not been made. To accurately predict the cloud feedbacks in the climate system, advanced understanding is needed for two of the elements in the feedback loops: the ability of cloud models/parameterizations to obtain cloud properties from changing atmospheric states and the ability to use these cloud properties to predict the radiative energy fluxes that in turn heat or cool the atmosphere.

Because a wide range of temporal and spatial scales are associated with clouds and cloud processes, it is not straightforward to represent them in models (Fig. 1). Each cloud model can only be expected to cover a narrow range of these scales, depending upon the main purpose of that particular model. For example, the smaller-scale processes such as turbulence and microphysics have to be parameterized, as opposed to explicitly resolved, in cloud-resolving models (CRMs), which only explicitly resolve cloud-scale dynamics and mesoscale processes. In climate models, a simple cloud parameterization is traditionally used to represent the gross effects of clouds and cloud processes on the large-scale circulations. This approach has many difficulties, which will be discussed later. A recently proposed multiscale modeling framework (MMF) approach replaces most of the physical parameterizations in a large grid cell of GCM with a high-resolution CRM (Grabowski 2001; Khairoutdinov and Randall 2001). This approach allows for an explicit simulation of many cloud processes, including convection, overlapping clouds in both the radiative and microphysical senses, and convectively generated gravity waves, as pointed out by Randall et al. (2003). Preliminary tests of the MMF approach have indicated that a GCM with an embedded CRM can produce the Madden–Julian oscillation as well as higher-frequency tropical waves in a much more realistic manner than a GCM with a traditional cloud parameterization (Grabowski 2003; Randall et al. 2003).

One major difficulty in representing clouds and cloud processes with the traditional approach is the extreme nonlinearity of many cloud processes in addition to the subgrid variabilities: particle condensation is triggered as a step function at 100% relative humidity; microphysical phase change occurs as a nonlinear function of ambient temperature and moisture; and finally there is a nonlinear relationship between cloud physical properties and radiative heating and cooling. Due to inadequate understanding of cloud processes, some of these nonlinearities are typically represented by "tunable" parameters in cloud parameterizations of GCMs that are used in future projections of climate. A recent comparison of GCMs with long-term radiation budget mea-

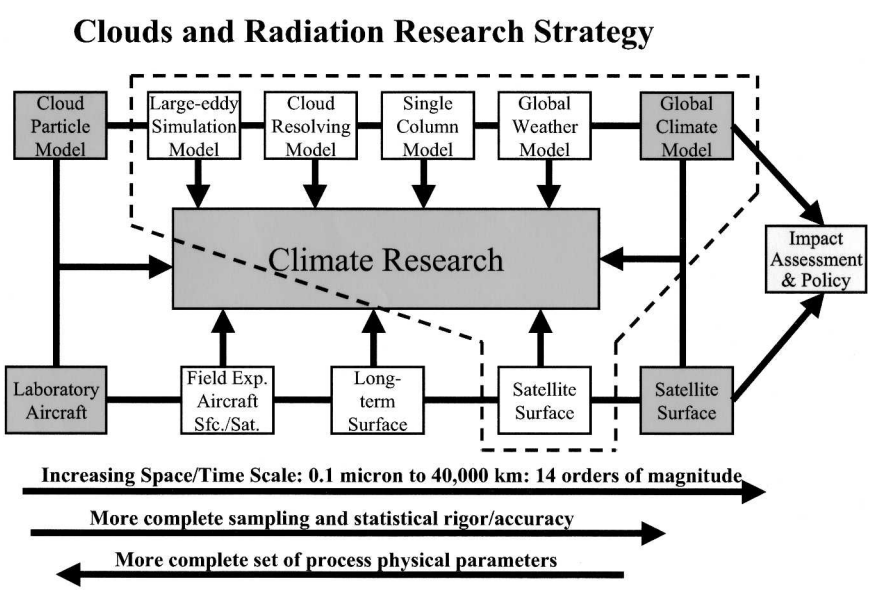


FIG. 1. Research strategy for the role of clouds in climate (after Wielicki et al. 1995). Current research strategy is shown by the dashed line.

surements indicates that some observational periods are poorly simulated by GCMs (Wielicki et al. 2002). The problems with these simulations are mainly due to the use of tunable parameters and our inadequate knowledge of cloud processes. The use of the MMF approach eliminates the need for many of these tunable parameters because cloud dynamical and microphysical processes are operating on their native scales in CRMs. Thus, it may be able to do a better job at simulating the long-term record of the radiation budget.

Major efforts to date in attacking the cloud–radiative feedback problem have naturally focused on vertical "pairings" at a given time and space scale as shown in Fig. 1. Regional field experiments such as the Global Atmospheric Research Program (GARP) Atlantic Tropical Experiment (GATE: Kuettner and Parker 1976), and the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE: Webster and Lukas 1992) can only provide a few detailed case studies. Because of the nonlinearities discussed above, this approach is challenged by the existence of only a limited number of cases that must then be extrapolated to global conditions. The ongoing Department of Energy's Atmospheric Radiation Measurement (ARM) Program (Stokes and Schwartz 1994; Ackerman and Stokes 2003) gives us a much longer record by taking measurements of vertical columns over three sites for the past decade. This gives more cases per year, but is limited to three sites on the globe. On the opposite end of the temporal and spatial range, monthly mean global and regional satellite and surface data have been used to verify GCM simulations. However, these monthly composites are best suited to studying linear cloud–radiative feedbacks, not nonlinear ones. Satellite data have also been composited over short periods according to dynamical regimes to evaluate performance of GCMs (e.g., Klein and Jakob 1999; Tselioudis et al. 2000; Webb et al. 2001; Tselioudis and Jakob 2002), but the quality of the satellite data used in these studies is not as high as that used in this study. A new approach is advocated in this study, which compares numerical weather prediction (NWP) cloud model performance and process-resolving models, such as large-eddy simulation (LES) models and CRMs, with matched satellite, surface, and atmospheric state data.

A new and innovative method of satellite data analysis is proposed to systematically evaluate and improve cloud models and cloud parameterizations. Specifically, this new technique classifies satellite data into distinct cloud objects, which are contiguous areas of cloud systems within a satellite swath that satisfy physical criteria of a cloud-system type (e.g., trade/shallow cumulus, solid stratus, and deep convection). These observed cloud objects are then matched with nearly simultaneous atmospheric state data from the European Centre for Medium-Range Weather Forecasts (ECMWF). The atmospheric data could also be used as inputs for cloud model simulations. The cloud model results are then compared with the statistical characteristics of satellite footprint data of cloud objects. Instead of using the traditional gridded-mean comparison, the approach advocated here advances cloud model evaluation by using large, statistically robust ensembles of matched atmospheric states and satellite cloud object data, with an emphasis on evaluating the higher-order distributions of finescale characteristics of cloud objects between satellite observations and cloud models. This dataset can also used to evaluate the MMF approach with the frequency of occurrence for each cloud-system type in various regions from a GCM simulation in addition to the higher-order distributions of finescale characteristics of cloud objects. Errors in both frequencies of occurrence and statistical distributions of cloud physical properties will be identified to further improve the MMF approach through an iterative process.

The main goal of this study is to present a methodology for generating the cloud object data product for four oceanic cloud-system types: deep convection, trade/shallow cumulus, transition stratocumulus, and solid stratus. Preliminary results are shown to illustrate the strengths of this satellite data analysis approach for two short periods of satellite observations. Section 2 presents the data and methodology. Results for tropical convective systems and boundary layer cloud systems are presented in sections 3 and 4, respectively. Summary and conclusions are given in section 5.

## 2. Data and methodology

### *a. Input data*

The basic satellite measurements used to produce the cloud object dataset are the level 2 footprint data product from the Clouds and the Earth's Radiant Energy System (CERES: Wielicki et al. 1996) investigation, the Single Satellite Footprint (SSF) top-of-the-atmosphere (TOA)/Surface Fluxes and Clouds Edition 2B product. The SSF product contains a wide variety of individual parameters including geometric properties of the relationship between the satellite and the earth, temporally and spatially coincident imager-based radiances, cloud properties, and aerosols (ocean only for TRMM), observations of broadband reflected shortwave (SW), emitted longwave (LW) and window (WN) radiances, and top-of-atmosphere fluxes of SW and LW from the CERES instrument. Each CERES SSF file contains

one hour of footprint data located in colatitude and longitude at a surface reference level. A list of parameters in the cloud object data product is given in Table 1.

The CERES SSF broadband radiative fluxes in the cloud object data are produced using the new generation of angular distribution models (ADMs) derived from the Tropical Rainfall Measuring Mission (TRMM) CERES broadband radiance observations (Loeb et al. 2003). These improved ADMs significantly reduce both root-mean-square (rms) and bias TOA broadband flux errors for all scenes by a factor of 3 to 4 when compared with those of the previous generation Earth Radiation Budget Experiment (ERBE) data (Suttles et al. 1988, 1989). For example, the new CERES ADMs have 200 SW scene types and several hundred LW scene types (compared to only 12 scene types from ERBE) and include stratification by both cloud fraction and optical depth. These added scene types can produce highly accurate instantaneous TOA fluxes from directional radiance measurements by increasing the ADMs' sensitivity to parameters that strongly influence anisotropy. Loeb et al. (2003) estimated that 1° regional instantaneous TOA fluxes are accurate to within 10 W $m^{-2}$ in the SW and 3.5 W $m^{-2}$ in the LW with little dependence on cloud phase, cloud optical depth, or infrared cloud emissivity.

The CERES SSF data also include cloud microphysical, macrophysical, and optical properties. The SSF combines instantaneous CERES broadband radiation data with scene information from a higher-resolution imager (Fig. 2), which is the Visible/Infrared Scanner (VIRS) on the TRMM satellite and the Moderate Resolution Imaging Spectroradiometer (MODIS) on the *Terra* satellite. Scene identification (type and clear or cloudy) and cloud properties (cloud effective height, temperature, pressure, particle types, and equivalent diameters) are produced using a remote sensing technique that uses VIRS radiance and sounding information from ECMWF. These data are energy-weighted over the larger CERES footprints according to the CERES instrument energy point–spread function. Details of the retrieval methods are described in the cloud retrieval section of the CERES algorithm theoretical basis document (Minnis et al. 1997).

TABLE 1. A list of parameters in the CERES SSF data product used in the production of the cloud object dataset.

| Parameter type | Name of parameters |
|---|---|
| Cloud macrophysics | Cloud fraction<br>Effective cloud height<br>Effective cloud pressure<br>Effective cloud radiative temperature |
| Cloud microphysics | Cloud liquid water path<br>Cloud ice water path<br>Cloud water droplet radius<br>Cloud ice particle diameter |
| Cloud optics | Visible cloud optical depth<br>Cloud infrared emissivity |
| Radiation | TOA broadband albedo<br>TOA reflected broadband shortwave radiative flux<br>TOA outgoing broadband longwave radiative flux |

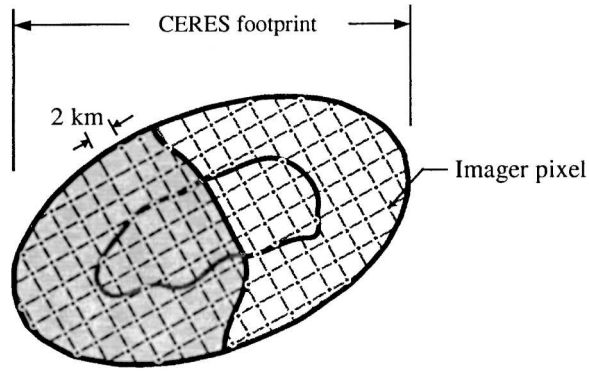


FIG. 2. Relationship between a CERES footprint and imager pixels. The imager pixel data are energy-weighted over the larger CERES footprint according to the CERES instrument energy point–spread function (Minnis et al. 1997).

### b. Cloud object method

Because of the large volume of SSF data (6 GB per day), the first major task for this analysis is to reduce the data volume that one has to work with by selecting only the portion of the original CERES SSF data that contain the selected cloud-system types. A basic assumption is that clouds are not random, white-noise phenomena. It is posited that individual clouds are elements in larger cloud systems. A contiguous region that is composed of individual CERES footprints that satisfy a physically based selection criterion or criteria for a given cloud-system type (see section 2d) is called a cloud object. Thus, a cloud object is a part of a larger cloud system because of the limited width of the satellite swath and the use of specific cloud-system selection criteria. The limited width of the satellite swath can truncate a whole cloud system. The selection criteria can break a large cloud system into several smaller cloud objects.

The second major task that one faces at the outset is to identify individual cloud objects within a swath of data contained in each SSF file. This is accomplished using a "region growing" strategy based on imager-

TABLE 2. Cloud-system type selection criteria used in this study.

| Cloud-system type | Cloud-top height | Cloud optical depth | Cloud fraction | Latitude band |
|---|---|---|---|---|
| Tropical deep convection | >10 km | >10 | 1.0 | 25°S–25°N |
| Trade/shallow cumulus | <3 km | — | 0.1–0.4 | 40°S–40°N |
| Transition stratocumulus | <3 km | — | 0.4–0.99 | 40°S–40°N |
| Solid stratus | <3 km | — | 0.99–1.0 | 40°S–40°N |

derived cloud properties to identify the cloud objects within a single satellite swath (Wielicki and Welch 1986). For all CERES fields of view (FOV) in a 700-km-wide TRMM satellite orbit swath, each CERES FOV that meets the selection criteria in Table 2 is marked as that cloud type. These "seed points" are grown using the algorithm described in Wielicki and Welch (1986). Only fields of view that are adjacent and that meet the selection criteria of a single cloud type can be joined in a cloud object. By adjacent, we mean CERES FOVs that are next to each other along the scanning direction, or perpendicular to it. These two directions would constitute orthogonal $x$ and $y$ axes if the satellite ground track is plotted as the $y$ axis. The region-growing method allows holes to occur within cloud objects. The shape of cloud objects can be irregular. Cloud objects are uniquely determined when they share no adjacent CERES fields of view. Cloud objects that grow to an equivalent diameter of less than 100 km (deep convection) or 75 km (boundary layer clouds) are ignored in the present analysis.

The CERES FOVs are roughly elliptical in shape (Fig. 2): the long axis is in the scanning direction, and the short axis is in the satellite ground track direction. An image of the data along the ground track looks like a simple television raster image and, like most imagers, the optimal sampling design is to allow some FOV overlap. The actual field of view of the CERES radiometer is a fixed scanning angular view toward the earth. Since the distance to the earth increases away from the satellite point, the size of the area viewed on the earth with each CERES FOV increases with viewing angle (Wielicki et al. 1996). CERES FOVs used in this study of TRMM data are limited to those with matching VIRS cloud imager data. The VIRS imager scan angle limit of 49° viewing zenith angle limits the amount of FOV growth to modest levels and sets a 700-km swath width from the TRMM 350-km orbit altitude. For the current study, a single average CERES FOV area of 100 km$^2$ is used. This representative field of view area is determined by dividing the total number of CERES FOVs in a single orbit by the total area of the earth covered in a single TRMM orbit (with 700-km swath). As a result, the cloud object equivalent diameter will be slightly smaller if the cloud object is near the satellite ground track, slightly larger if the cloud object is near the edge of the swath, compared to that calculated from using the assumption of a fixed FOV area. Obviously, it will have almost no error for objects large enough to cover most of the swath, or at intermediate positions. Since the largest error due to this effect is only ±30% in cloud object diameter and it constitutes noise rather than bias, we chose a single fixed CERES FOV area when calculating cloud object equivalent diameters. On average, this will cause a one-sigma noise in cloud object diameter of roughly 20%, which is insignificant in our statistical analysis.

A unique feature of the cloud object analysis is that cloud objects are identified independently of where they are located relative to latitude and longitude coordinate grids. In other words, this approach shifts toward a Lagrangian view as opposed to an Eulerian framework, although such a view is just a snapshot of the cloud systems. Temporal evolution of the cloud objects cannot be examined due to the infrequent overpass of the EOS satellites at a given location (Wielicki et al. 1996) unless geostationary satellite data are incorporated into this analysis.

### c. *Cloud object data product*

There are three types of data in the cloud object data product: 1) cloud object footprint data, 2) cloud object statistical data, and 3) cloud object histogram data. All of these data are available on the Web (http://cloud-object.larc.nasa.gov/) for the climate research community. The cloud object footprint data contain SSF cloud footprint data, with parameters listed in Table 1, that satisfy the selection criteria to be given in section 2d. The cloud object statistical data contain normal statistics for each of the individual cloud objects. The cloud object histogram data contain files that provide the characteristics of the spatial variabilities for cloud objects. The footprint data are placed into bins of equal widths for different parameters. An example is given in Fig. 3. Figure 3a shows the spatial distribution of OLR fluxes for a tropical convective cloud object observed on 17 March 1998, along with the boundaries of the TRMM satellite swath, while Fig. 3b shows the probability density of OLR fluxes, which is the frequency of occurrence in a bin interval normalized by the bin

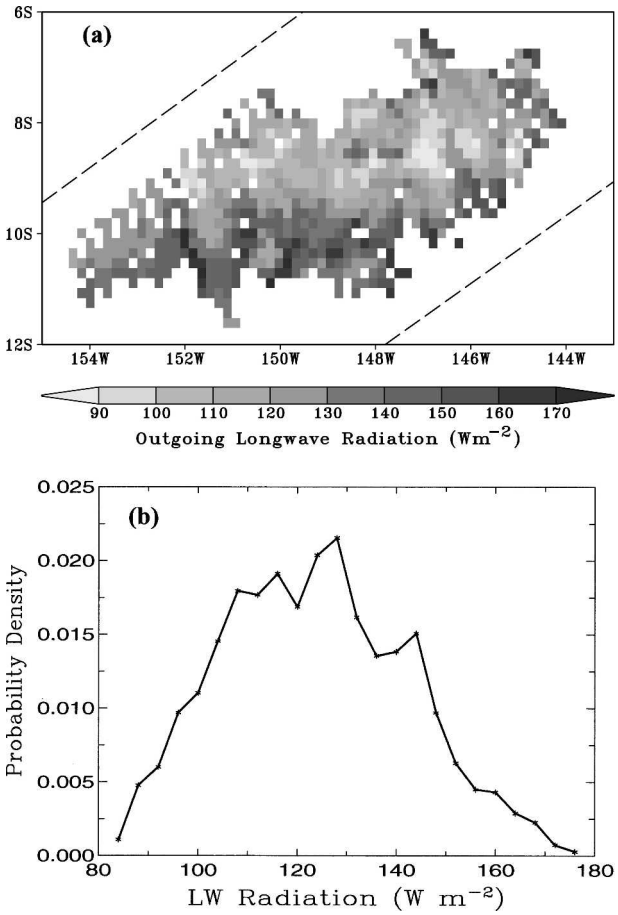


FIG. 3. (a) Horizontal distribution of outgoing longwave radiative fluxes and (b) the associated probability density function for a 17 Mar 1998 tropical deep convective cloud object in the tropical Pacific Ocean. The dashed lines in (a) show the boundaries of the TRMM satellite swath.

width. The bin width used here is 4 W m$^{-2}$. Figure 3b illustrates the usefulness of the probability density function (PDF) representation of the spatial variabilities for a cloud object. The OLR has a quasi-Gaussian distribution for this cloud object. Distributions of other parameters can be made in a similar manner.

The individual cloud object histogram files can be used to produce the overall statistics for a large ensemble of cloud objects associated with a given geographic region, cloud object size, or cloud-system type. The statistics for these ensembles are termed "grand-mean" statistics in this study. The grand-mean statistics can also be produced according to fields obtained from the matched atmospheric states such as the low-level stratification and SST. These grand-mean PDFs provide a representative description of cloud systems under specific climate conditions. The main reason for presenting the grand-mean statistics, instead of presenting the statistics for individual cloud objects, is that PDFs vary significantly from one individual cloud object to another due to short-term variabilities in the cloud systems. Statistically meaningful PDFs for climate analysis can only be achieved for a large number of cloud objects. Some examples will be shown to support this point in section 3a.

### *d. Cloud-system type selection criteria*

Four distinct types of single-layer, oceanic cloud systems are initially studied in this investigation. These types are tropical deep convection, boundary layer solid stratus, transition stratocumulus, and trade/shallow cumulus. For tropical deep convection, we are only interested in the convective towers and their associated thick upper-tropospheric anvils. Four criteria must be simultaneously satisfied for this cloud-system type. First, the footprints must have 100% cloud fraction, that is, overcast conditions. Second, a threshold value of 10 for cloud optical depth is used to eliminate thin anvil clouds. Third, the cloud-top heights must be greater than 10 km. The threshold values on cloud optical depth and cloud-top height are loosely based upon the International Satellite Cloud Climatology Project (ISCCP) definition of deep convection (Rossow and Schiffer 1991) and have slightly narrower ranges than the convective deep cloud regime identified by Jakob and Tselioudis (2003). Last, the cloud footprints must be located over the Pacific Ocean between 25°S and 25°N so that only tropical convective systems are included. In the near future, this analysis will be extended to include cloud objects in the entire Tropics.

For each of the boundary layer cloud-system types, the cloud-top height must be less than 3 km. A higher threshold on the cloud-top height would include many frontal clouds and ice-phase clouds, which would greatly change the characteristics of the PDFs because of the large differences in cloud optical properties between liquid-phase and ice-phase clouds. For the three boundary layer cloud-system types, different thresholds on cloud fraction are applied: solid stratus, 0.99–1.00; transition stratocumulus, 0.40–0.99; and trade/shallow cumulus, 0.10 to 0.40. The lower limit of cloud fraction on the trade/shallow cumulus cloud system is designed to consistently remove the uncertainties in imager-based measurements of cloud fraction and other cloud characteristics. The cloud fraction ranges were chosen based upon the determination of large systematic variations in cloud optical depth distributions with cloud fractions that are resolved by high-resolution Landsat data (Barker et al. 1996; Barker and Wielicki 1997). Boundary layer cloud objects are identified in the entire latitude band from 40°S to 40°N to include tropical

and subtropical cloud systems. A summary of the selection criteria is provided in Table 2.

### e. Matching the cloud objects with ECMWF meteorological data

To facilitate comparison with cloud model simulations and to understand the sensitivity of cloud objects to atmospheric states, the instantaneous satellite cloud object data are matched with ECMWF meteorological fields at the same location within ±3 h of the satellite overpass. The ECMWF data are available every 6 h. All two-dimensional and three-dimensional variables are provided on grids with a horizontal resolution of about $0.56° \times 0.56°$ in the Tropics. A rectangular area is chosen to match the ECMWF data with the observed cloud object, which usually has an irregular shape. The exact size of the rectangular area is dependent upon the shape and size of the cloud object because the rectangular area must include all footprints of the cloud objects. When a large number of cloud objects becomes available, only those cloud objects that occurred near the daily ECMWF analysis times (0000, 0600, 1200, and 1800 UTC) will be selected. This tight match in time will be helpful to maximize the ability to study the cause and effect more directly between the atmospheric states and cloud objects. The matched data will allow classification of cloud objects using either satellite data or atmospheric state data.

### f. Periods of analysis

Two months of the CERES data collected from the TRMM satellite have been analyzed for this study: March 1998 and March 2000. Data from other periods will be analyzed in Part II of this series of papers. March 1998 was during the peak of the 1997/98 El Niño event; March 2000 represents a very weak La Niña event. The TRMM Precipitation Radar (PR) and TRMM Microwave Imager (TMI) data have been analyzed to characterize the precipitation distributions of observed cloud objects for March 1998 (Eitzen and Xu 2005).

### g. Method for comparing two PDFs

Statistical tests are used to detect statistically significant differences between two grand-mean PDFs. These are done in the following ways. First, the differences in PDFs are measured by a root-mean-square method for two PDFs of the same parameter, which is called the Euclidean distance or L2. This PDF distance measure is defined as

$$\mathrm{L2} = \left\{ \sum_{i=1}^{N} [f(x_i)\Delta x_i - g(x_i)\Delta x_i]^2 \right\}^{1/2},$$

where $f$ and $g$ are two PDFs, with a total of $N$ bins where the $i$th bin is located at $x_i$. The bin width is denoted by $\Delta x_i$. As mentioned in section 2b, the frequency of occurrence is normalized by the bin width. That is, $f$ and $g$ satisfy $\Sigma_{i=1}^{N} f(x_i)\Delta x_i = \Sigma_{i=1}^{N} g(x_i)\Delta x_i = 1$. The bin width $\Delta x_i$ is uniform for the PDFs examined here. The maximum possible value of L2 is $\sqrt{2}$, which occurs if two single-point PDFs are not collocated. The minimum value of this measure is zero, which indicates no difference between the PDFs. This approach takes consideration of the difference for the entire range of PDF bins. Another PDF distance measure, called the Hellinger distance or $H$, is designed to give more weight to small differences in the bins with small probability densities (Matusita 1955). This is defined as

$$H = \left\{ \sum_{i=1}^{N} [\sqrt{f(x_i)\Delta x_i} - \sqrt{g(x_i)\Delta x_i}]^2 \right\}^{1/2}.$$

As will be shown in section 3, the values of $H$ are usually higher than those of L2. However, the maximum possible value of $H$ is also $\sqrt{2}$. Both PDF distance measures will be used in section 3, but L2 is preferred because it gives an equal weight for the differences at all bin intervals.

Second, the bootstrap method (Efron and Tibshirani 1993) is used to determine whether the difference between two grand-mean PDFs is statistically significant. A statistically significant difference between two grand-mean PDFs means that the individual cloud objects forming the two grand-mean PDFs came from two different populations. Cloud objects, but not their individual footprints, are assumed to be independent from each other. The null hypothesis is that all cloud objects came from one population. The probability, calculated taking the null hypothesis to be true, that we would observe a statistic value as extreme or more extreme than the one we did observe is the significance level. The test statistic chosen in this study is either L2 or $H$. The bootstrap method is based upon the fact that the sampling distribution of the test statistic, a relative frequency distribution of the test statistic, is a good estimate of the population distribution. To approximate the population distribution of the test statistic, one needs to draw a large number of resamples randomly with replacement from a population, calculate the value of the test statistic, and repeat $B$ times.

Specifically, in this study, the two populations of $m$ and $n$ cloud objects are first combined into one population. Then, two sets of cloud objects of sizes of $m$ and $n$ are resampled randomly from the population, and the values of the distance measures between the two bootstrapped sets are calculated. Any cloud object in the

TABLE 3. Number of tropical deep convective systems according to different size classes and geographic regions for the Mar 1998 and Mar 2000 periods.

| Equivalent diameter of cloud objects (km) | Mar 1998 (strong El Niño) | | | | Mar 2000 (near normal) | | | |
|---|---|---|---|---|---|---|---|---|
| | Western | Central | Eastern | Total | Western | Central | Eastern | Total |
| 100–150 | 45 | 45 | 36 | 126 | 73 | 37 | 32 | 142 |
| 150–300 | 53 | 52 | 31 | 136 | 68 | 51 | 25 | 144 |
| >300 | 31 | 26 | 11 | 68 | 33 | 10 | 11 | 54 |
| Total (>100 km) | 129 | 123 | 78 | 330 | 174 | 98 | 68 | 340 |

population can be sampled once, more than once, or not at all at any given time. This procedure is repeated $B$ ($B$ is chosen to be 5000) times to generate a statistical distribution of the test statistic (L2 or $H$). The bootstrapped distance value is compared to the value from the true arrangement of cloud objects, that is, two separate populations. If the bootstrapped value is greater than the true value between two populations in less than 5% of a total calculation of $B$ times, the two populations are deemed to be statistically different. That is, the null hypothesis is rejected at the 5% significance level.

## 3. Results for tropical deep convective systems

### a. Comparison of frequency of cloud objects

The number of tropical deep convective systems identified from the CERES SSF data is listed in Table 3, according to the range of equivalent diameters, for the March 1998 and March 2000 periods. This study only considers those cloud objects with equivalent diameters greater than 100 km because these large systems are more likely to be reproduced by NWP models than the smaller ones and tend to last for a longer period of time. Both of these characteristics increase the probability of a match of the assimilated meteorological data with satellite cloud objects. In this study, three size classes are considered. They are in the ranges of 100–150 km, 150–300 km, and greater than 300 km. This analysis also separates the tropical cloud systems into three longitudinal regions: western (130°E to date line), central (date line to 130°W), and eastern (130° to 80°W) Pacific.

For all cloud objects with sizes greater than 100 km, the regional differences are very pronounced between the two periods. The western Pacific had the most cloud objects in March 2000 because the eastern Pacific had cold SSTs that inhibited convection during the very weak La Niña. During the 1997/98 El Niño, however, more cloud systems occurred over the central and eastern Pacific because of the higher SSTs there. This is especially evident for the largest size class of cloud objects in the central Pacific and the two smaller size classes of cloud objects in the eastern Pacific. Table 3 also shows that there are more cloud objects in the largest size class in March 1998 than in March 2000. These results may be due to the intensification of the Hadley circulation and the disappearance of the Walker circulation during the intense 1997/98 El Niño (e.g., Cess et al. 2001; Wang et al. 2003).

### b. Examples of individual cloud objects

Figure 4 shows PDFs of TOA albedo and OLR for a few individual cloud objects. Five of the cloud objects shown in Fig. 4 have 4000–5000 footprints; two of them have about 3000 footprints, and one has about 7500 footprints (labeled as 0302). The equivalent diameters of these cloud objects range from 600 to 1000 km with typical sizes of about 800 km. There are large variabilities among the cloud objects in both the TOA albedo and cloud optical depth in terms of the peak values and the shapes of the curves. These variabilities are primarily associated with different large-scale dynamics or varying dynamical states among the cloud objects. For an identical dynamical state, the PDFs for several cloud objects may be slightly different. The reason for these differences is called "weather noise." So, the variabilities among the cloud objects are associated with both dynamical processes, which act at short time scales, and weather noise.

When a large number of cloud objects are combined to produce the grand-mean PDFs, variabilities associated with short-term dynamical processes are smoothed out. The resulting grand-mean PDFs provide a more representative description of cloud systems under specific climate conditions than any individual cloud object does. These examples have illustrated the importance of using grand-mean statistics of footprint characteristics in examining climate sensitivity and evaluating cloud model simulations using the combined data prod-

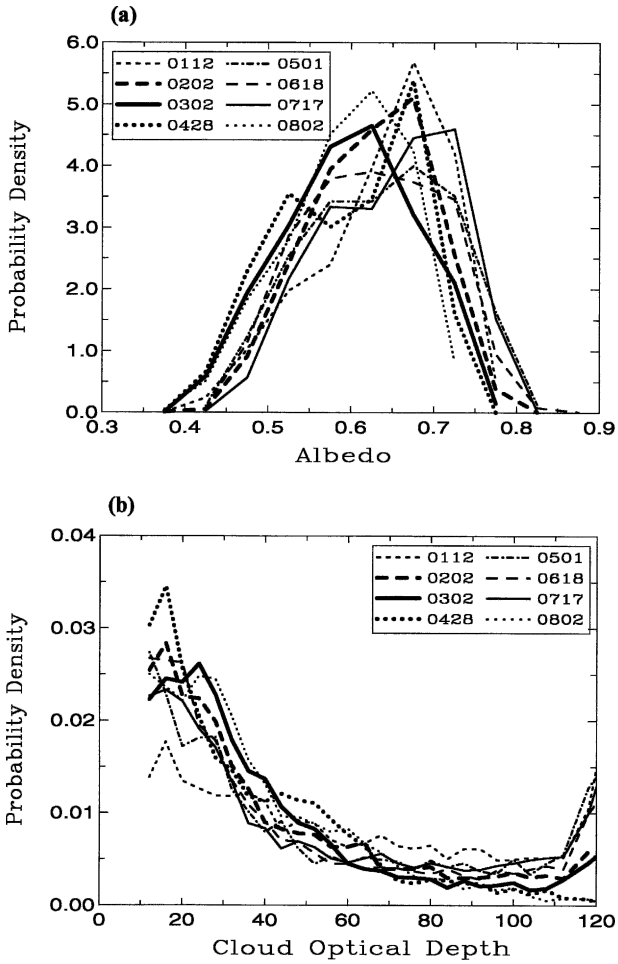


FIG. 4. Probability density functions of (a) TOA albedo and (b) cloud optical depth for eight selected cloud objects of tropical deep convective cloud-system type observed during the 1997/98 El Niño period. The legend (mmdd) denotes the month and day that each cloud object occurred.

uct of cloud objects with matched atmospheric state data. The former topic will be investigated further in Part II of this series of papers, while the latter is discussed in Eitzen and Xu (2005).

### c. *Comparing March 1998 and March 2000 PDFs*

In this section, selected results are shown to illustrate the CERES footprint statistical characteristics of cloud systems in terms of PDFs for all cloud objects of one size class. Because the largest size class (equivalent diameter greater than 300 km) has the highest number of footprints, its PDFs are smoother in comparison with those of the smaller size classes. This cloud-system class also dominates the total number of footprints used in the PDF analysis for all combined size classes. Thus, its PDFs are rather similar to those for all combined size classes of cloud objects. There are 68 such cloud objects in March 1998 and 54 in March 2000 (Table 3). Unlike the PDFs of individual cloud objects shown earlier, the differences in the grand-mean PDFs are rather small for most parameters between the March 1998 and March 2000 periods. The statistical significance test described in section 2g is used to assign a probability that the two PDFs come from the same population ($p$ value). The smaller the probability, the more likely the two PDFs come from different populations. Table 4 shows the statistical significance levels for several parameters examined in this section, using either L2 or $H$ as the test statistic.

Figure 5a shows the PDFs of SST associated with the largest cloud objects for the two periods. During March 2000 more cloud objects occur over SSTs between 299 and 301.5 K. The peak distribution shifted to higher SST values (SSTs greater than 302 K) in March 1998 during the 1997/98 El Niño event. This shows a fundamental difference between a climatologically abnormal El Niño period and a very weak La Niña period because more cloud objects occur over warmer SSTs during El Niño (Table 3). The statistical significance testing confirms this; that is, there is a 5.9% probability for L2 and <1% probability for $H$ that these two cloud object populations have similar SST distributions (Table 4).

The most significant difference between the two ob-

TABLE 4. Selected measures of the differences in PDFs between the Mar 1998 and Mar 2000 periods and the associated significant levels.

| Parameters | Euclidean distance (L2) | Significance for L2 | Hellinger distance ($H$) | Significance for $H$ |
|---|---|---|---|---|
| Sea surface temperature | 0.1841 | 5.9% | 0.3465 | <1% |
| Cloud-top height | 0.0806 | <1% | 0.1832 | <1% |
| Cloud-top pressure | 0.0678 | <1% | 0.1452 | 1.1% |
| OLR | 0.0338 | 7.9% | 0.1267 | 1.9% |
| TOA albedo | 0.0115 | 96% | 0.0268 | 89% |
| Cloud-top temperature | 0.0378 | 7.9% | 0.1337 | 1.1% |
| Cloud optical depth | 0.0106 | 77% | 0.0222 | 8.4% |
| Ice water path | 0.0233 | 15% | 0.0444 | 45% |
| Ice diameter | 0.0512 | 7.8% | 0.1771 | 16% |

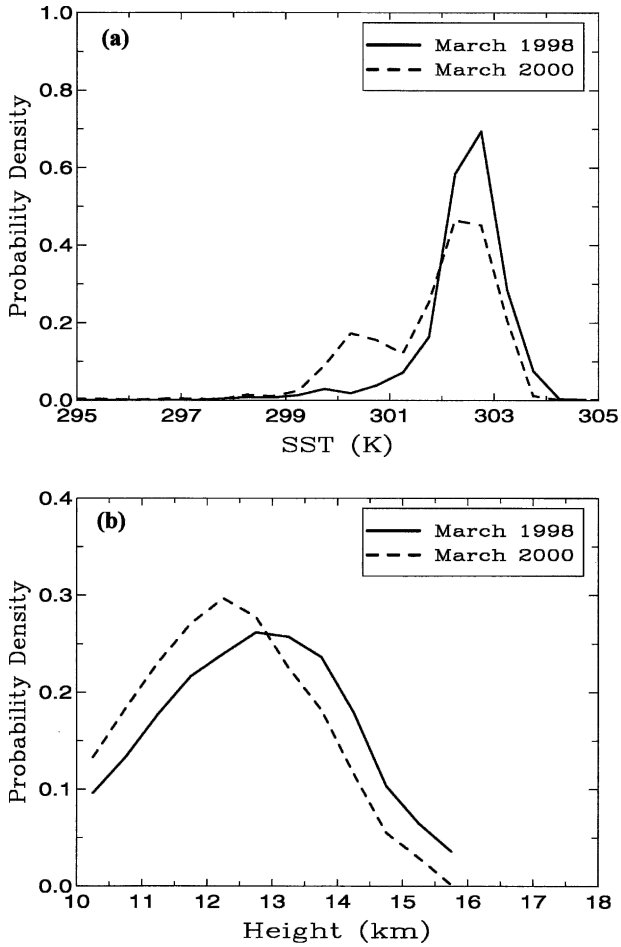


FIG. 5. Probability density functions of (a) sea surface temperature and (b) cloud-top height for tropical deep convective cloud objects during the Mar 1998 and Mar 2000 periods.

servation periods appears in the PDFs of cloud-top heights (Fig. 5b). Cloud tops are approximately 0.5–1.0 km higher in March 1998 than in March 2000. The cloud-top height corresponding to the maximum probability density is 12.75 km in March 1998, compared to 12 km in March 2000. The probability density is lower by approximately 0.05 per kilometer for cloud-top heights less than 12.5 km in March 1998 than in March 2000. The opposite is true for cloud-top heights greater than 13.0 km. This result is related to the significant changes in the stratification of the atmosphere that occurred during El Niño, particularly in the central and eastern Pacific, compared to a climatologically near-normal year. Statistical significance testing shows that the two PDFs have less than 1% probability of being similar using either L2 or $H$ as the test statistic.

In spite of the significant differences in the cloud-top heights, there are no significant differences in the TOA SW albedo, and the PDFs of the LW radiative fluxes are somewhat different between the two periods (Fig. 6). This is because optically thick cloud footprints are selected for tropical convective cloud objects. The sensitivity of the attenuation of solar radiation for optically thick clouds to the cloud-top height is rather weak. On the other hand, the TOA OLR is closely associated with the cloud-top temperature for this type of cloud object (Fig. 7a). The differences in the cloud-top temperature below 210 K are smaller than might be expected given the large differences in cloud-top heights between the two periods (Fig. 5b) if one were to assume that the stratification of the atmosphere remains the same between the two periods. This assumption is probably not true, as explained below.

Both cloud-top temperature and OLR are moderately likely to be different between these two periods, which have statistical significance for Hellinger distances at less than the 2% level but at the 7.9% level for L2. For high OLRs that correspond to those of middle to upper tropospheric anvils, the cloud-top temperature and OLR show some significant differences between

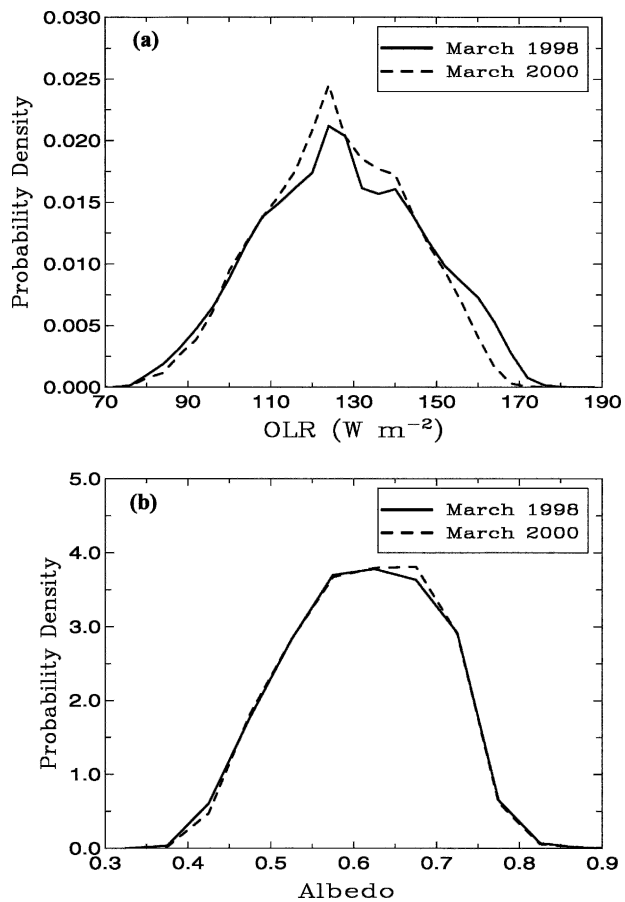


FIG. 6. Same as in Fig. 5, but for (a) OLR flux and (b) the TOA albedo.

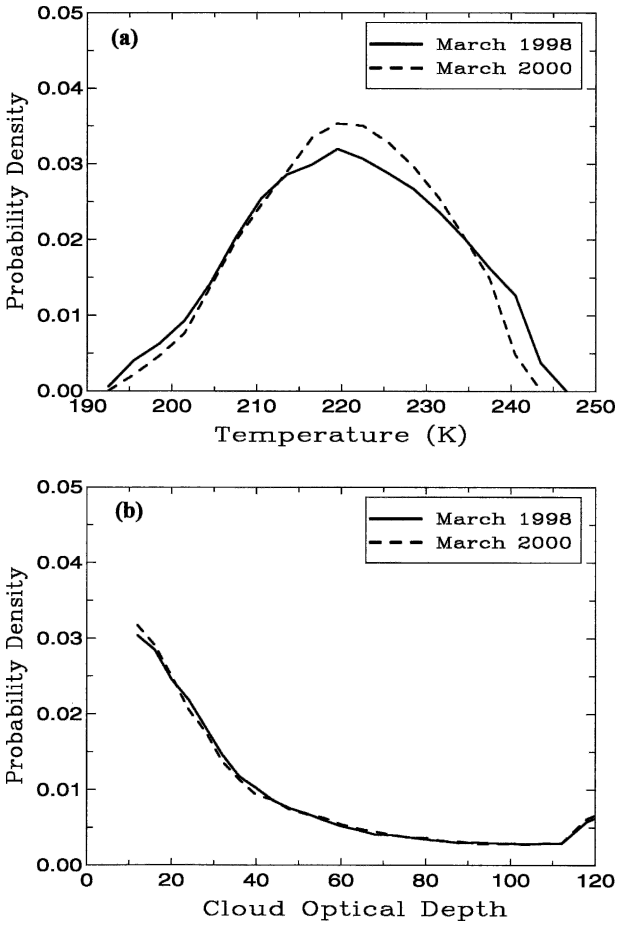


FIG. 7. Same as in Fig. 5, but for (a) cloud-top temperature and (b) cloud optical depth.

these two periods, but for very low OLRs, for example, less than 100 W $m^{-2}$, there is virtually no difference between these two periods. This result is probably associated with the change of the stratification of the atmosphere. During El Niño, tropical deep convection is more intense in the central and eastern Pacific. The lapse rates in the convective regions are thus less steep because more convective available potential energy (CAPE) is consumed by cumulus convection. As a result, the higher cloud-top heights are not necessarily associated with colder cloud-top temperatures in March 1998, compared to March 2000. Instead, the cloud-top temperatures do not change much between these two periods, especially for OLRs less than 100 W $m^{-2}$. This result supports the fixed anvil temperature (FAT) hypothesis of Hartmann and Larson (2002) observationally, who proposed that the emission temperature of tropical convective anvils would not change during climate change. The FAT hypothesis was previously only supported by modeling results.

In spite of some significant differences in the cloud-top height and pressure and moderately significant differences in the OLR and cloud-top temperature, the TOA albedo, cloud microphysical, and optical properties of cloud objects do not show many differences between these two periods (Figs. 6b, 7b, 8a, and 8b). The cloud optical depths show exponential distributions in both periods while the ice water paths show lognormal distributions. The differences for these parameters are not statistically significant between the two periods (Table 4). The ice diameter has a statistical significance at the 7.8% level for L2 and at the 16% level for $H$. This means that there are moderately significant differences in the ice diameters between these two periods.

In summary, a majority of the CERES footprint statistical characteristics of observed tropical convective systems are similar between the March 1998 and March 2000 periods, in spite of the climatological contrast (strong El Niño versus very weak La Niña). There are, however, statistically significant differences in some cloud macrophysical properties such as cloud-top height and cloud-top pressure. The statistical signifi-

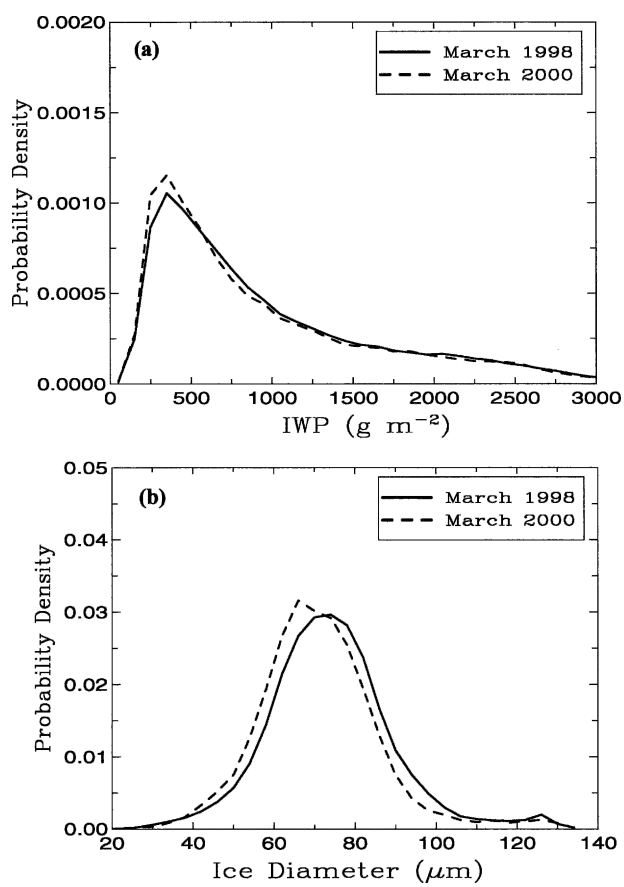


FIG. 8. Same as in Fig. 5, but for (a) ice water path and (b) ice diameter.

TABLE 5. Number of boundary layer cloud-system types identified from the SSF data for the region between 40°S and 40°N for the Mar 1998 and Mar 2000 periods.

| Equivalent diameter of cloud objects (km) | Mar 1998 (strong El Niño) | | | | Mar 2000 (near normal) | | | |
|---|---|---|---|---|---|---|---|---|
| | Cumulus | Transition | Solid stratus | Total | Cumulus | Transition | Solid stratus | Total |
| 75–150 | 3233 | 8565 | 3328 | 15 126 | 3968 | 8102 | 2731 | 14 801 |
| 150–300 | 227 | 1737 | 1071 | 3035 | 380 | 1579 | 883 | 2842 |
| >300 | 9 | 323 | 490 | 822 | 28 | 302 | 406 | 732 |
| Total (>75 km) | 3469 | 10 625 | 4889 | 18 983 | 4376 | 9983 | 4020 | 18 275 |

cance levels are between 5% and 10% for OLR, cloud-top temperature, and ice diameter for the L2 measure, indicating moderately significant differences of these parameters between the two periods. A majority of the analyzed parameters have nearly Gaussian distributions except for cloud optical depth and ice water path, which are distributed exponentially and lognormally, respectively. These statistical properties will be contrasted with those of the boundary layer cloud systems in the next section.

## 4. Results for boundary layer cloud types

### *a. Comparison of frequency of cloud objects*

The total number of boundary layer cloud objects is nearly identical (within 3%) between the March 1998 and March 2000 periods. The total number includes all cloud objects that occur between 40°S and 40°N and have equivalent diameters greater than 75 km. The proportion of different boundary layer cloud-system types is significantly changed as a result of the shifted circulation patterns related to El Niño in March 1998, that is, a large increase in solid stratus type (22%) and a large decrease in cumulus type (−21%), and a slight increase in stratocumulus (6%). The increases/decreases occur over all three size classes, as shown in Table 5. The disappearance of the Walker circulation in the Pacific during the 1997/98 El Niño may have impacted the partitioning of the cloud-type population since the large-scale subsidence was significantly stronger in some boundary layer cloud–dominated regions.

One of the regions dominated by boundary layer cloud systems is the southeast Pacific, which is defined as the region between 40°S and the equator and between 150°W and the coast of South America in this study. In this region, the impact of the 1997/98 El Niño on the boundary layer cloud population is largely similar to its impact on the entire 40°S–40°N latitude belt (Table 6). The number of the solid stratus population increases slightly (5%) but that of the cumulus population decreases significantly (>40%). The stratocumulus population only slightly decreases. The increase of the solid stratus population is mainly for the largest cloud-system size. The smaller size classes have fewer cloud objects during March 1998 (Table 6). The increase of the stratus cloud population is related to the increase of large-scale subsidence associated with the descent branch of the Hadley circulation as the Walker circulation shifted eastward and disappeared during the 1997/98 El Niño. The northeast Pacific region (Table 7), which extends from 40°N to the equator and from 150°W to the coast, experienced large increases in the solid stratus (56%) and stratocumulus (24%) population, but no change in the cumulus population during the 1997/98 El Niño. The increases in the stratus and stratocumulus population occurred mainly in the relatively warmer SSTs (>301 K). Higher SSTs do not usually favor the occurrence of solid stratus and stratocumulus clouds unless there is a significant increase in the

TABLE 6. As in Table 5 except for the southeast Pacific region, which is defined as 150°W to the coast of South America and between 40°S and the equator.

| Equivalent diameter of cloud objects (km) | Mar 1998 (strong El Niño) | | | | Mar 2000 (near normal) | | | |
|---|---|---|---|---|---|---|---|---|
| | Cumulus | Transition | Solid stratus | Total | Cumulus | Transition | Solid stratus | Total |
| 75–100 | 238 | 749 | 329 | 1316 | 425 | 830 | 262 | 1517 |
| 100–150 | 105 | 510 | 225 | 840 | 204 | 506 | 252 | 962 |
| 150–300 | 20 | 216 | 178 | 414 | 56 | 197 | 186 | 439 |
| >300 | 0 | 24 | 90 | 114 | 1 | 20 | 61 | 82 |
| Total (>75 km) | 363 | 1499 | 822 | 2684 | 686 | 1553 | 761 | 3000 |

TABLE 7. As in Table 5 except for the northeast Pacific region, which is defined as 150°W to the coast of North America and between 40°N and the equator.

| Equivalent diameter of cloud objects (km) | Mar 1998 (strong El Niño) | | | | Mar 2000 (near normal) | | | |
|---|---|---|---|---|---|---|---|---|
| | Cumulus | Transition | Solid stratus | Total | Cumulus | Transition | Solid stratus | Total |
| 75–100 | 182 | 414 | 163 | 759 | 172 | 291 | 101 | 564 |
| 100–150 | 42 | 211 | 84 | 337 | 51 | 221 | 67 | 339 |
| 150–300 | 20 | 137 | 81 | 238 | 20 | 123 | 43 | 186 |
| >300 | 0 | 51 | 46 | 97 | 1 | 21 | 29 | 51 |
| Total (>75 km) | 244 | 813 | 374 | 1431 | 244 | 656 | 240 | 1140 |

strength of subsidence associated with the Hadley circulation. An increase in the strength of the subsidence likely explains the change of the proportion of different boundary layer cloud types during the El Niño.

### *b. PDFs for the southeast Pacific region*

As mentioned in section 3, the PDFs for the boundary layer cloud objects are distinctly different from those of the tropical deep convection. Are there any similarities between them? A few examples for the 100–150-km size class will be shown in this section to contrast the differences and point out some similarities among the boundary layer cloud objects and between tropical deep convection and boundary layer cloud systems.

Before cloud microphysical, macrophysical, and radiative properties of boundary layer clouds are discussed it should be pointed out that boundary layer clouds occur over a much wider SST range (Fig. 9a) than tropical deep convective cloud objects do (Fig. 5a). Tropical deep convection occurs only over warm SSTs (299–304 K: Fig. 5a) and in the ascent branches of the Walker and Hadley circulations. But boundary layer clouds usually occur in the descent branches of these circulations over much larger areas and wider SST ranges (290–303 K: Fig. 9a). The wide SST ranges are related to the much wider descent regions of the Hadley circulation. This distinction determines the differences between tropical deep convective and boundary layer clouds, as will be shown later. Other factors such as the inversion strength and the magnitude of the large-scale subsidence above the boundary layer determine which type of boundary layer cloud object occurs in a certain area. The differences in some cloud macrophysical properties, as shown later, between the March 1998 and March 2000 periods are also related to differences in the SST distributions. Despite the noisy features in the PDFs shown in Fig. 9a, it can be seen that the SST distribution is relatively flat for each specific boundary layer cloud-system type in March 1998 with a range that is wider by about 2 K than in March 2000. The PDFs have peaks at specific values of SSTs in March 2000. The peak is located at a few degrees higher for cumulus clouds than for solid stratus clouds. This feature of the SST distributions can strongly impact the OLR fluxes and some cloud macrophysical properties, as shown later.

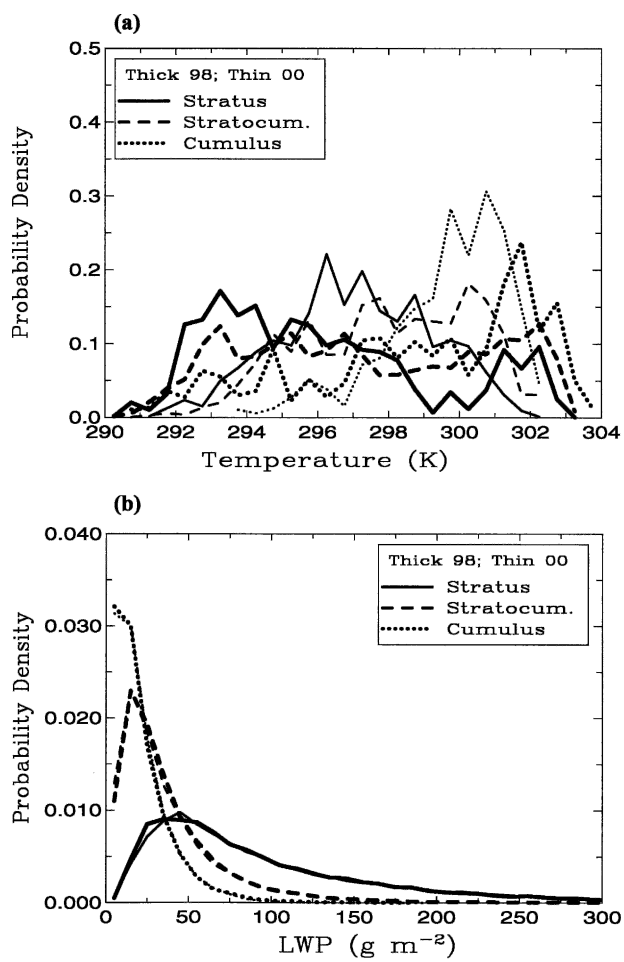


FIG. 9. Probability density functions of (a) sea surface temperature and (b) liquid water path for boundary layer cloud-system types during the Mar 1998 (thick curves) and Mar 2000 (thin curves) periods.

The liquid water path (LWP) PDF does not show many differences between the March 1998 and March 2000 periods (Fig. 9b). However, the PDFs differ greatly among the three cloud-system types. The LWP PDFs are nearly exponential for cumulus cloud objects, but lognormal for stratocumulus and solid stratus cloud objects. The PDFs for solid stratus cloud objects have much longer tails than those of stratocumulus cloud objects. Similar PDFs were identified by earlier studies using high-resolution Landsat data (e.g., Wielicki and Parker 1992). The most remarkable result of the earlier observations is that a clear Gaussian-like behavior in LWP only occurs for nearly overcast ($C > 95\%$) or overcast cloud conditions. For almost all broken cloud cases, similar to the cumulus cloud object in this study, the PDF peaks at the smallest measurable LWP. Barker et al. (1996) and Barker and Wielicki (1997) showed that these frequency distributions could be approximated very accurately by using a Gamma function requiring only the mean LWP, $\langle \mathrm{LWP} \rangle$, and the standard deviation of LWP, ($\sigma$), to predict the Gamma distribution. The key Gamma function parameter is $\nu = (\langle \mathrm{LWP} \rangle / \sigma)^2$: for $\nu \gg 1$ the distribution is essentially Gaussian-like, while for $\nu < 1$ the distribution becomes increasingly skewed with peak values at zero LWP. The agreement with the prior high-resolution observations is very encouraging and indicates that the boundary layer cloud object types can be studied using the CERES measurements.

The cloud optical depth PDFs are rather similar to those of LWP (Fig. 10a). This is not surprising because cloud optical depth is related to LWP by the equivalent droplet radius, which has a relatively narrow Gaussian-like distribution (not shown). The differences among the three cloud-system types are as significant as those of LWP. For the TOA albedo, all PDFs can essentially be described by the Gamma distribution. The distribution of albedo is nearly Gaussian for the solid stratus cloud objects (Fig. 10b), whose PDF is similar to those of tropical deep convection except that it has a smaller median value (Fig. 6b). The albedos corresponding to the peak probability densities are very different among the three boundary layer cloud-system types, with the lowest for cumulus and the highest for solid stratus because of the significant differences in both cloud fraction and cloud optical depth in the CERES footprints for these two types of boundary layer cloud objects.

Significance tests of the PDF differences between the two periods show no significant differences for all cloud microphysical (except for droplet radius for stratus; not shown) and optical properties, as well as the TOA albedo (Table 8). This is an important new finding. As the large-scale circulation patterns were changed during the 1997/98 El Niño, these properties of each specific cloud-system type were not significantly changed in a large geographic region despite the significant differences in SSTs. The change of large-scale circulation patterns is, thus, more likely to cause the change of the proportion of different boundary layer cloud types, as well as their macrophysical properties discussed below.

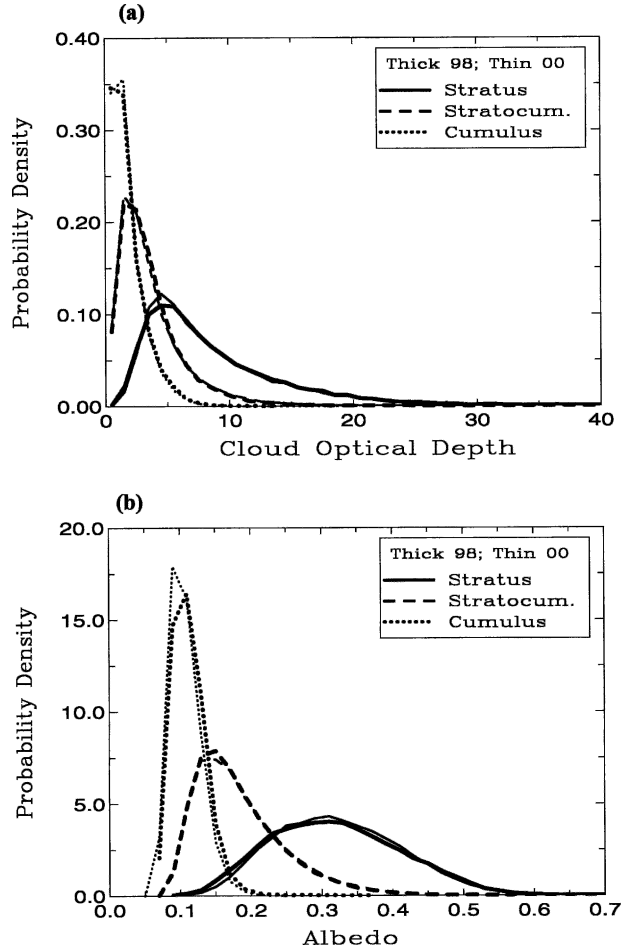


FIG. 10. Same as in Fig. 9, but for (a) cloud optical depth and (b) TOA albedo.

It is obvious that the OLR PDFs are significantly different for each specific cloud-system type between the two periods (Fig. 11a, Table 8). The PDFs have flatter distributions in March 1998 than in March 2000, although all of them are basically described by the Gamma distributions with different modes. The Gaussian-like behavior for solid stratus and stratocumulus cloud types is also similar to that of tropical deep convection (Fig. 6a), as discussed in section 3. The flatter distribution of OLRs in March 2000 is associated with the characteristics of the SST PDFs from the same period because these boundary layer cloud types are strongly influenced by the underlying surface condi-

TABLE 8. The L2 distances for measuring the differences between two PDFs of selected parameters between the Mar 1998 and Mar 2000 periods for all three types of boundary layer cloud objects. The associated significance levels are also shown.

| | L2 distance | | | Significance level for L2 | | |
|---|---|---|---|---|---|---|
| | Cumulus | Stratocumulus | Stratus | Cumulus | Stratocumulus | Stratus |
| Sea surface temperature | 0.2109 | 0.1522 | 0.1866 | <1% | <1% | <1% |
| Cloud-top height | 0.0428 | 0.0558 | 0.1198 | 20% | <1% | <1% |
| Cloud-top pressure | 0.0154 | 0.0224 | 0.0514 | 78% | <1% | <1% |
| Cloud-top temperature | 0.1849 | 0.0705 | 0.1890 | <1% | <1% | <1% |
| OLR | 0.1426 | 0.0301 | 0.0839 | <1% | 1.1% | <1% |
| TOA albedo | 0.0851 | 0.0124 | 0.0193 | 6.4% | 31% | 15% |
| Cloud optical depth | 0.0176 | 0.0191 | 0.0176 | 86% | 17% | 28% |
| Liquid water path | 0.0123 | 0.0247 | 0.0182 | 93% | 5.5% | 30% |
| Water droplet radius | 0.0415 | 0.0166 | 0.0460 | 10% | 34% | <1% |
| Emissivity | 0.0165 | 0.0123 | 0.0218 | 32% | 5.3% | 23% |

tions (Fig. 9a). This dependence of the OLR PDF on the SST could be removed by examining the LW cloud radiative forcing, which is defined as the difference between the cloudy and clear sky OLRs. Clear sky OLRs would have to be calculated from a radiative transfer code.

Cloud macrophysical properties such as cloud-top height (Fig. 11b), temperature, and pressure (not shown) are significantly different for each specific cloud object type between the two periods except for the cloud-top height and pressure of boundary layer cumulus cloud objects (Table 8). Because the lapse rates of stratus and stratocumulus clouds are very close to moist adiabatic values, the distributions of cloud-top height, pressure, and temperature are expected to be strongly dependent upon the distribution of SSTs. Boundary layer cumulus clouds behave differently because the atmospheric stabilities or lapse rates are more variable than those of stratus and stratocumulus clouds. They may not be strongly dictated by the distribution of SSTs. It is noted that the macrophysical properties of these boundary layer clouds are less strongly impacted by the underlying surface conditions as the size of cloud object increases, especially for stratus and stratocumulus cloud objects (not shown). This is because the larger cloud objects are better organized and thus more strongly influenced by the large-scale subsidence than by the SST.

In summary, the CERES footprint statistical characteristics of satellite-observed boundary layer cloud objects are distinctly different in all cloud microphysical, macrophysical, and cloud optical properties and radiative fluxes among the three cloud-system types. In spite of these differences, some of these properties have distributions with similar shapes to those of tropical deep convection but with different modes, for example, OLR, cloud droplet equivalent radius (compared to equivalent ice diameter of deep convection), TOA albedo of stratus cloud objects, cloud optical depth of cumulus cloud objects, and LWP of stratus cloud objects (compared to IWP of deep convection). The differences between the March 1998 and March 2000 periods are not significant in most cloud microphysical and optical properties and the top-of-the-atmosphere albedo of the boundary layer cloud object types, but are significant in some cloud macrophysical properties and OLR. Similar results are also obtained in the northeast Pacific region (not shown). Further investigation of the boundary layer cloud objects will be focused on many smaller geographic regions using long-term data products.

## 5. Summary and discussion

This study has presented an objective classification methodology that uses Earth Observing System (EOS) satellite data to identify distinct cloud objects defined by cloud-system types, sizes, geographic locations, and matched large-scale environments. This analysis method identifies a cloud object as a contiguous region of the earth with a single dominant cloud-system type, independently of where it is located relative to some coordinate grid of the earth. It determines the shape and size of the cloud object from the satellite data and the cloud-system selection criteria. This method enhances our ability to study the cloud feedback processes directly for a single cloud-system type, as opposed to the multiple cloud-system types that are often simultaneously present in a fixed region of the earth using monthly, seasonally, or yearly averaged satellite/surface data. Because there are thousands of observed cloud objects, statistical characteristics of these cloud objects can be computed for different climate regimes or large-scale circulations, and thus the cloud-radiative feedbacks can be isolated for specific cloud-system types and climate regimes. That is, the partial derivatives of

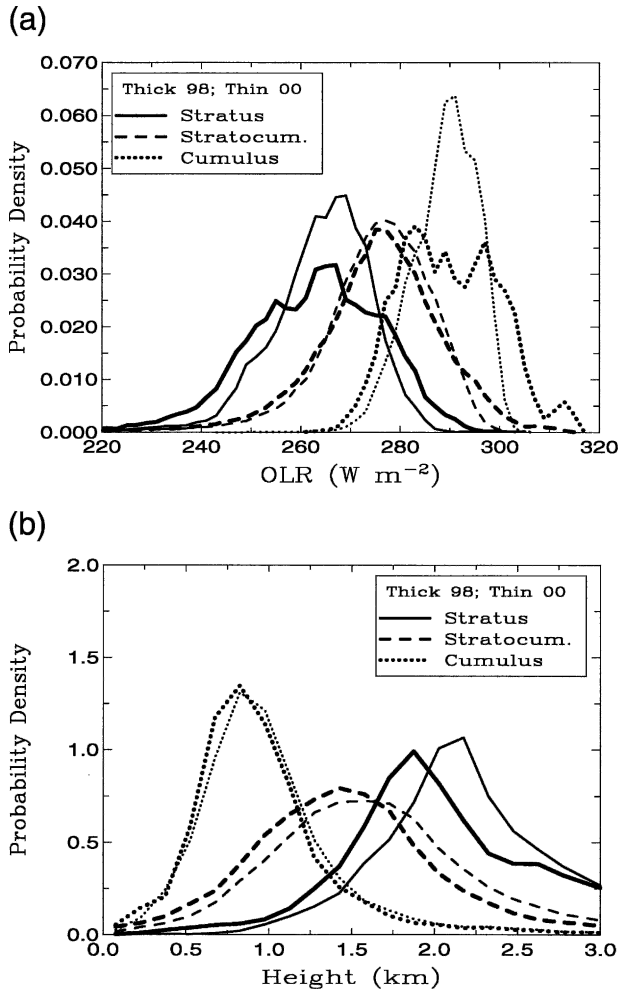


FIG. 11. Same as in Fig. 9, but for (a) OLR flux and (b) cloud-top height.

cloud radiative forcings with respect to atmospheric states can be computed for each cloud-system type or climate regime.

This study has focused on the analysis of the statistical properties of cloud objects in terms of their PDFs, which are produced from the CERES SSF data on the EOS satellites. The PDFs of cloud microphysical, macrophysical, and optical properties and radiative fluxes can be usefully compared with cloud model simulations of a large ensemble of cloud objects driven by matched ECMWF meteorological data. Here, cloud models include large-eddy simulations, cloud-resolving models, and cloud parameterizations/single-column models. These models have been extensively tested with field experiment data (e.g., Xu and Randall 1996, 2000; Ghan et al. 2000; Xu et al. 2002; Siebesma et al. 2003). With the cloud object data product, the new approach can take cloud model evaluation beyond the more traditional tests that use a few field experiment case studies to tests that use large, statistically robust ensembles of matched meteorological states and satellite cloud object data. In the analysis of simulation outputs, the same cloud-system selection criteria are applied to select model columns for the PDF comparison because more than one cloud-system type may be produced for a given meteorological situation. Eitzen and Xu (2005) provided a preliminary study of model evaluation using tropical deep convective cloud object data.

Improvement of cloud models will be a major goal for using this advanced data product with hundreds and thousands of cloud objects. Models can be tuned for a few cases, as many modelers do, with the expectation of some improved performance. But it will be more difficult to do so for hundreds or thousands of cloud object cases. This is particularly true for the higher-order statistical moments of the errors and for the dependence of the cloud variables on different atmospheric-state parameters. The ability to quantify the errors as a function of the atmospheric state is expected to be a powerful tool in pinpointing model problems because the large cloud-object sample size will allow one to rigorously determine the accuracy of model simulations with respect to the atmospheric states. An improved cloud model can be evaluated against the cloud object data. Through an iterative process, this approach can improve a GCM's treatment of clouds, either by evaluating a conventional cloud parameterization or by evaluating a cloud-resolving model and using it in a multiscale modeling framework (MMF). The improved climate simulation from the MMF approach can be evaluated against the monthly mean satellite/surface data, but also against the detailed statistical properties from the satellite cloud-object data product and the frequency of occurrence for each specific cloud-system type in various geographic regions. Thus, the MMF approach can be more vigorously validated against the satellite cloud-object data.

The cloud object data product has been generated from the CERES data collected from the TRMM satellite for this study. Four distinct types of single-layer, oceanic cloud objects, that is, tropical deep convection, boundary layer cumulus, transition stratocumulus, and solid stratus, have been initially identified from the CERES data. Preliminary results have been presented from the analysis of these four distinct types of cloud objects associated with the strong 1997/98 El Niño and very weak 2000 La Niña events. Major results of this analysis are summarized below.

A majority of the CERES footprint statistical characteristics of observed tropical deep convection, such as the TOA albedo, cloud optical depth, emissivity, and

ice water path, are similar between the March 1998 and March 2000 periods in spite of their climatological contrast. There are, however, statistically significant differences in some cloud macrophysical properties such as the cloud-top height and cloud-top pressure and moderate differences in OLR, cloud-top temperature, and ice diameter. A majority of the analyzed parameters have nearly Gaussian distributions except for cloud optical depth and ice water path, which are distributed exponentially and lognormally, respectively. The total number of tropical deep convective cloud objects is approximately the same in these two periods, but the number of the largest cloud objects is larger during the El Niño period as a result of the higher SSTs from the same period. The geographic distributions are also quite different; that is, more cloud objects are identified in the central and eastern Pacific during the El Niño period.

The CERES footprint statistical characteristics of observed boundary layer clouds are distinctly different in all cloud microphysical, macrophysical, cloud optical properties and radiative fluxes between the three cloud-system (cumulus, stratocumulus, and stratus) types. The differences between the two periods are not significant in most cloud microphysical and optical properties and the TOA albedo, but are significant for some cloud macrophysical properties and the OLR. Furthermore, the partitioning of the boundary layer cloud object population is different between the two periods. An increase of the stratus cloud population and a decrease of the cumulus cloud population is evident for a region that consists of all the Tropics and subtropics. This change in the populations is also seen in the boundary layer cloud-dominated regions such as the southeast and northeast Pacific during the El Niño period, compared to the climatologically near-normal period.

Part II of this series of papers will provide a more detailed analysis of the cloud-object data product using the eight-month (January–August 1998) data collected from the TRMM satellite. The cloud object data from the *Terra* and *Aqua* satellites will become available soon. Because of the much larger sample sizes of cloud objects, the analysis will be performed based upon the matched atmospheric states and over much smaller geographic regions than those presented in this study.

The methodology proposed in this study is a unique tool for analyzing satellite/surface data with matched atmospheric states. The preliminary four cloud-system types can be expanded to include many more cloud-system types such as the midlatitude frontal cloud system and polar cloud system. This methodology will also be very effective with data collected from the NASA A-train satellites, which will provide vertical profiles of cloud liquid/ice contents and aerosol information, when these data are available.

*Acknowledgments.* The CERES data were obtained from the Atmospheric Sciences Data Center at the NASA Langley Research Center. This research has been supported by NASA EOS interdisciplinary study program and by NSF Grant ATM-0336762. The authors would also like to acknowledge Dr. Bruce Barkstrom of Langley Research Center and Professor David Randall of Colorado State University for their earlier insightful vision to this project and Professor Lisa Bloomer of Middle Tennessee State University for designing the statistical significance test used in this study.

## REFERENCES

Ackerman, T. P., and G. M. Stokes, 2003: The Atmospheric Radiation Measurement Program. *Phys. Today,* **56,** 38–44.

Barker, H. W., and B. A. Wielicki, 1997: Parameterizing grid-averaged longwave fluxes for inhomogeneous marine boundary layer clouds. *J. Atmos. Sci.,* **54,** 2785–2798.

——, ——, and L. Parker, 1996: A parameterization for computing grid-averaged solar fluxes for inhomogeneous marine boundary layer clouds. Part II: Validation using satellite data. *J. Atmos. Sci.,* **53,** 2304–2316.

Cess, R. D., M. Zhang, P.-H. Wang, and B. A. Wielicki, 2001: Cloud structure anomalies over the tropical Pacific during the 1997/98 El Niño. *Geophys. Res. Lett.,* **28,** 4547–4550.

Efron, B., and R. Tibshirani, 1993: *An Introduction to the Bootstrap.* Chapman & Hall, 436 pp.

Eitzen, Z. A., and K.-M. Xu, 2005: A statistical comparison of deep convective cloud objects observed by an Earth Observing System satellite and simulated by a cloud-resolving model. *J. Geophys. Res.,* **110,** D15S14, doi:10.1029/2004JD005086.

Ghan, S. J., and Coauthors, 2000: An intercomparison of single column model simulations of summertime midlatitude continental convection. *J. Geophys. Res.,* **105,** 2091–2124.

Grabowski, W. W., 2001: Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.,* **58,** 978–997.

——, 2003: MJO-like coherent structures: Sensitivity simulations using the cloud-resolving convection parameterization (CRCP). *J. Atmos. Sci.,* **60,** 847–864.

Hartmann, D. L., and K. Larson, 2002: An important constraint on tropical cloud–climate feedback. *Geophys. Res. Lett.,* **29,** 1951, doi:10.1029/2002GL015835.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, and D. Xiaosu, 2001: *Climate Change 2001: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC).* Cambridge University Press, 944 pp.

Jakob, C., and G. Tselioudis, 2003: Objective identification of cloud regimes in the tropical western Pacific. *Geophys. Res. Lett.,* **30,** 2082, doi:10.1029/2003GL018367.

Khairoutdinov, M. F., and D. A. Randall, 2001: A cloud-resolving model as a cloud parameterization in the NCAR Community

Climate System Model: Preliminary results. *Geophys. Res. Lett.,* **28,** 3617–3620.

Klein, S. A., and C. Jakob, 1999: Validation and sensitivity of frontal clouds simulated by the ECMWF model. *Mon. Wea. Rev.,* **127,** 2514–2531.

Kuettner, J. P., and D. E. Parker, 1976: GATE: Report on the field phase. *Bull. Amer. Meteor. Soc.,* **57,** 11–27.

Loeb, N. G., N. Manalo-Smith, S. Kato, W. F. Miller, S. K. Gupta, P. Minnis, and B. A. Wielicki, 2003: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Tropical Rainfall Measuring Mission satellite. Part I: Methodology. *J. Appl. Meteor.,* **42,** 240–265.

Matusita, K., 1955: Decision rules based on distance for problems of fit, two samples and estimation. *Ann. Math. Stat.,* **26,** 631–641.

Minnis, P., and Coauthors, 1997: Cloud optical property retrieval (subsystem 4.3). Clouds and the Earth's Radiant Energy System (CERES) Algorithm Theoretical Basis Document, Vol. III, Cloud Analyses and Radiance Inversions (Subsystem 4), 60 pp. [Available online at http://asd-www.larc.nasa.gov/ATBD/ATBD.html.]

Randall, D. A., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003: Breaking the cloud parameterization deadlock. *Bull. Amer. Meteor. Soc.,* **84,** 1547–1564.

Rossow, W. B., and R. A. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.,* **72,** 2–20.

Siebesma, A. P., and Coauthors, 2003: A large eddy simulation intercomparison study of shallow cumulus convection. *J. Atmos. Sci.,* **60,** 1201–1221.

Stokes, G. M., and S. E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed. *Bull. Amer. Meteor. Soc.,* **75,** 1202–1221.

Suttles, J. T., and Coauthors, 1988: Shortwave radiation. Vol. I, Angular Radiation Models for Earth–Atmosphere Systems, NASA Rep. RP-1184, 144 pp.

——, R. N. Green, G. L. Smith, B. A. Wielicki, I. J. Walker, V. R. Taylor, and L. L. Stowe, 1989: Longwave radiation. Vol. II, Angular Radiation Models for Earth–Atmosphere Systems, NASA Rep. RP-1184, 84 pp.

Tselioudis, G., and C. Jakob, 2002: Evaluation of midlatitude cloud properties in a weather and a climate model: Dependence on dynamic regime and spatial resolution. *J. Geophys. Res.,* **107,** 4781, doi:10.1029/2002JD002259.

——, Y. Zhang, and W. B. Rossow, 2000: Cloud and radiation variations associated with northern midlatitude low and high sea level pressure regimes. *J. Climate,* **13,** 312–327.

Wang, P.-H., P. Minnis, B. A. Wielicki, T. Wong, R. D. Cess, M. Zhang, L. B. Vann, and G. S. Kent, 2003: Characteristics of the 1997/1998 El Niño cloud distributions from SAGE II observations. *J. Geophys. Res.,* **108,** 4009, doi:10.1029/2002JD002501.

Webb, M., C. A. Senior, S. Bony, and J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dyn.,* **17,** 905–922.

Webster, P. J., and R. Lukas, 1992: TOGA COARE: The coupled ocean–atmosphere response experiment. *Bull. Amer. Meteor. Soc.,* **73,** 1377–1417.

Wielicki, B. A., and R. M. Welch, 1986: Cumulus cloud properties derived using Landsat satellite data. *J. Climate Appl. Meteor.,* **25,** 261–276.

——, and L. Parker, 1992: On the determination of cloud cover from satellite sensors: The effect of sensor spatial resolution. *J. Geophys. Res.,* **97,** 12 799–12 823.

——, R. D. Cess, M. D. King, D. A. Randall, and E. F. Harrison, 1995: Mission to Planet Earth: Role of clouds and radiation in climate. *Bull. Amer. Meteor. Soc.,* **76,** 2125–2153.

——, B. R. Barkstrom, E. F. Harrison, R. B. Lee III, G. L. Smith, and J. E. Cooper, 1996: Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment. *Bull. Amer. Meteor. Soc.,* **77,** 853–868.

——, and Coauthors, 2002: Evidence for large decadal variability in the tropical mean radiative energy budget. *Science,* **295,** 841–844.

Xu, K.-M., and D. A. Randall, 1996: Explicit simulation of cumulus ensembles with the GATE Phase III data: Comparison with observations. *J. Atmos. Sci.,* **53,** 3710–3736.

——, and ——, 2000: Explicit simulation of midlatitude cumulus ensembles: Comparison with ARM data. *J. Atmos. Sci.,* **57,** 2839–2858.

——, and Coauthors, 2002: An intercomparison of cloud-resolving models with the Atmospheric Radiation Measurement Summer 1997 IOP data. *Quart. J. Roy. Meteor. Soc.,* **128,** 593–624.