JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 104, NO. D6, PAGES 6183–6191, MARCH 27, 1999

Albedo bias and the horizontal variability of clouds in subtropical marine boundary layers: Observations from ships and satellites

Robert Pincus[1, 2]
Laboratory for Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland

Sally A. McFarlane
Program in Atmospheric and Oceanic Science, University of Colorado, Boulder

Stephen A. Klein
Geophysical Fluid Dynamics Laboratory, NOAA, Princeton University, New Jersey

Abstract. Cloud optical properties vary dramatically at spatial scales smaller than typical grid cells in large-scale models, which can cause a significant overestimate of cloud albedo by the model. This plane parallel homogeneous (PPH) albedo bias exist may be reduced if the mean cloud optical thickness and the amount of variability are available, but little is known about how much variability exists in nature and to what factors it is sensitive. The authors combine 1331 observations made by volunteer surface observers with satellite imagery to assess the relationships between cloud fraction, cloud optical properties, and cloud type in marine boundary layer clouds off the coast of California during summer. Estimates of cloud fraction from the two datasets are in best agreement when a reflectance threshold between 0.09 and 0.10 is used. Satellite-derived cloud fraction increases slowly with sensor resolution at spatial scales from 1 to 32 km. Cloud fraction in scenes dominated by cumulus is much more sensitive to the reflectance threshold used for cloud detection than are scenes containing stratiform clouds. The mean magnitude of the PPH bias found here, 0.025, is considerably smaller than those found in other recent studies. When fit to the observed distributions of optical thickness both log-normal and gamma distributions substantially reduce the PPH bias. The mean and dispersion of log optical thickness are related to cloud type: optical thickness increases as cloud type changes from cumuliform to stratiform, while the relative amount of variability decreases. The authors suggest a basis for the parameterization of unresolved variability in large scale models.

1. Introduction

Clouds have long been recognized as being simultaneously one of the most important and sensitive aspects of the Earth-atmosphere system and one of the most difficult to simulate correctly. The representation of clouds in large-scale weather and climate models has improved greatly in the past decade as modelers have developed prognostic schemes for cloud water and ice [*Tiedtke*, 1993; *Del Genio et al.*, 1996; *Fowler et al.*, 1996]. These treatments couple large-scale dynamics, radiative processes, and the hydrologic cycle by predicting the average amount of liquid water and ice at each layer in each grid box.

Current prognostic cloud schemes add substantial realism to large-scale models, but do not explicitly represent the horizontal variability in cloud water that exists at spatial scales smaller then individual grid cells. One significant problem caused by this unresolved variability arises in the treatment of solar radiation. Cloud albedo is a convex function of optical thickness: given a constant increment in optical thickness, the change in albedo is larger for thin clouds than for thick clouds. For a fixed average value of optical thickness, therefore, a model grid cell exhibiting any internal horizontal variability will always be less reflective than a uniform grid cell. Depending on the amount of subgrid scale variability, this plane-parallel homogeneous albedo bias (PPH bias) can be quite large [*Cahalan et al.*, 1994].

The simplest treatment of cloud inhomogeneity is to allow clouds to occupy a variable fraction of each grid cell. Processes within the cell (including radiative transfer) can then be computed as a weighted average of clear- and cloudy-sky values. Model predictions of cloud fraction may also be tested against observations, since estimates of cloud cover have been made by surface observers and from satellite imagers for many years. Because the cloudy portion of each grid cell is assumed to be uniform, however, models which estimate cloud fraction are still subject to the PPH bias.

Techniques to treat the PPH bias in a computationally reasonable manner are a subject of active research. One possibility is a renormalization or rescaling of one or more of the quantities in the radiative transfer equation based on the amount of variability in optical thickness within each grid cell [*Cahalan et al.*, 1994; B. Cairns, A.A. Lacis, and B.E. Carlson, Absorption within inhomogenous clouds and its parameterization in general circulation models, submitted to *Journal of the Atmospheric Sciences*, 1998, hereinafter referred to as submitted manuscript, 1998] An alternative is to explicitly integrate a solution to the radiative transfer equation (i.e., the two-stream approximation) over a specified dis-

[1] Also at Joint Center for Earth Systems Technology, University of Maryland Baltimore Country, Baltimore
[2] Now at Department of Atmospheric and Oceanic Sciences, University of Wisconsin – Madison.

Paper number 1998JD200125.
0148-0227/99/1998JD200125$09.00

tribution [*Barker*, 1996; *Oreopoulos and Barker*, 1999]. Both methods require knowledge about the distribution of optical thickness that exists within each model grid cell.

Given evolving techniques for computing the domain-average albedo of inhomogeneous clouds, models must still determine the kind and amount of variability within a grid cell. Variability may be specified as a free parameter, but a more self-consistent approach would relate inhomogeneity to information available within the model. Variability is, in principle, dependent on a great number of interrelated factors including (but not limited to) mean cloud properties, cloud type, cloud microphysical state, geographic location, the level of atmospheric turbulence, and other meteorological conditions. Many of these relationships are quite difficult to observe: cloud properties and the state of the atmosphere are typically assessed using surface observations and profiling instruments including radiosondes, radar, and lidar, while albedo bias itself is most naturally observed with space-borne instruments.

This paper explores the relationships between cloud optical properties and cloud type in the subtropical marine boundary layer. We use surface observations of cloud amount and type to characterize the state of the atmosphere, and compare these to satellite observations of cloud amount and cloud optical thickness distributions at spatial scales between 1 and 128 km. After describing the data sets used in the study, we compare estimates of cloud fraction derived from satellite imagery with those made by surface observers, and explore the sensitivity of the satellite-based techniques to various parameters. We then examine the distributions of optical thickness derived from satellite images of marine boundary layer clouds and assess the accuracy with which these distributions may be represented by simple analytic functions. Finally, we relate the parameters of the optical thickness distributions to the underlying cloud type and propose ways in which these relationships might be used in large-scale models.

2. Simultaneous Observations of Clouds From Ships and Satellites

We combine two data sets collected during FIRE, the First ISCCP (International Satellite Cloud Climatology Project) Regional Experiment, between June 27 and July 25, 1987, in the eastern subtropical Pacific (110° W to 135° W; 20° N to 50° N) off the western coast of the United States and Baja California.

2.1 Observations From Ships of Opportunity

We employ individual synoptic surface cloud and meteorological observations from volunteer observing ships, which are taken from an extended version of the Edited Cloud Report Archive [*Hahn et al.*, 1996]. The archive is formed by applying quality control procedures to those observations of the Comprehensive Ocean Atmosphere Data Set [*Woodruff et al.*, 1987] that contain valid cloud information. These surface reports include estimates of total cloud cover and the cloud cover at the lowest level containing clouds. Cloud fraction is reported in octas (eighths) so that only nine discrete values are possible. Observers are instructed to report 1 octa if any cloud is visible, and 7 octas if any clear sky is visible. The observers also use the classification system provided by the World Meteorological Organization [1974] to record the dominant cloud type present at low, middle, and high levels. Mid- and high-level clouds are quite rare in our study area during summer.

2.2 Observations From Meteorological Satellites

We utilize daytime observations of narrow-band visible wavelength (0.65 μm) reflectance obtained by the VISSR instrument aboard the GOES6 platform. The instrument has 6-bit radiometric resolution and a nominal 1-km spatial resolution at nadir. Images are available every half-hour for locations east of 135° W longitude. Instrument counts are converted to reflectance using ISCCP calibration coefficients [*Rossow et al.*, 1992]. Pixels which are more reflective than a certain threshold are denoted cloudy, as discussed in section 3. The optical thickness for each cloudy pixel is then estimated using a look-up table retrieval [*Pincus et al.*, 1995] which assumes the presence of a midlatitude summer atmosphere. Solar zenith angle at scene center is distributed nearly normally across the set of images, with a mean value of 48.7°±7.9°; this narrow range of values makes stratification by solar zenith angle unenlightening.

2.3 Combining Observations

We merge these two data sets by matching pairs of ship and satellite observations of individual cloud fields in space and time. We discard ship observations further west than 134° W longitude and those closer to the California coast than 1°, since satellite images for these locations will be incomplete or will contain land. We associate each ship observation with the satellite image closest in time and extract from this image a scene 128 km on a side, centered on the ship location. This subimage covers an area somewhat larger than the area visible to the ship observer but between the sizes of grid cells used in climate and weather prediction models; we will show that the exact size of the subimage is unimportant. We discard ship observations for which a satellite image is not available within 45 min. We further restrict our attention to clear skies and fair weather boundary layer clouds: small and large cumulus (C_L 1 and 2), stratocumulus (C_L 5), fair weather stratus (C_L 6), mixtures of cumulus and stratocumulus (C_L 4 and 8), and sky-obscuring fog (as reported in the present weather code and denoted as C_L 11). The characteristics used to identify these clouds are noted in Table 1. We have excluded observations reporting stra-

Table 1. Low Cloud Classification.

C_L Code	Nontechnical Cloud-Type Description	Number
0	no stratocumulus, stratus, cumulus, or cumulonimbus	40
1	cumulus with little vertical extent and seemingly flattened, or ragged cumulus other than of bad weather, or both	124
2	cumulus of moderate or strong vertical extent, generally with protuberances in the form of domes or towers, either accompanied or not by other cumulus or by stratocumulus, all having their bases at the same level	138
4	stratocumulus from the spreading out of cumulus; cumulus may also be present	185
8	cumulus and stratocumulus other than that formed from the spreading out of cumulus; the base of the cumulus is at a different level than that of the stratocumulus	222
5	stratocumulus not resulting from the spreading out of cumulus	368
6	stratus in a more or less continuous sheet or layer, or in ragged shreds, or both, but no stratus fractus of bad weather	226

Cloud type 11 (sky obscuring fog) is not a WMO cloud type, but is used when the present weather code indicates sky-obscuring fog. After Hahn et al. [1996] and Norris [1998a].

tus fractus associated with synoptic storm systems (C_L 7) and cumulonimbus (C_L 3 and 9) to reduce the incidence of midlevel and high clouds which might not be accurately treated by our remote sensing algorithms. The resultant data set comprises 1331 pairs of ship and satellite observations.

3. Cloud Detection by Surface Observers and Satellite Radiometers

We begin our study of spatial variability in cloud properties by assessing sensitivities in the detection of clouds and the determination of cloud fraction (also called cloud amount or cloud cover). Cloud detection, by which cloud fraction is determined from satellite imagery, is the precursor to assessments of variability in cloud optical properties, since optical thickness distributions are properly computed only for cloudy pixels. Cloud fraction also provides the most direct link between the ship and satellite data sets because it is the single parameter the two data sets have in common.

3.1 Estimates of Cloud Fraction by Surface Observers

Table 1 documents the number of occurrences of each low cloud type in our data set of surface observations. Stratocumulus is the most commonly reported low cloud type, followed by stratus, then mixtures of stratocumulus and cumulus clouds. This is consistent with the long-term record from volunteer observing ships [*Norris*, 1998b]. These cloud types are associated with large values of cloud cover: completely cloudy skies (8 octas) are reported by the surface observers in more than half the observations in our data set, while cloud fractions of 6 octas or greater comprise 80% of the observations. Surface observer reports of cloud fraction are known to best agree with satellite estimates when skies are nearly overcast or nearly clear [*Henderson-Sellers et al.*, 1987].

3.2 Estimating Cloud Fraction From Satellite Imagery

Cloud fraction is typically estimated from satellite images by examining each pixel in turn and comparing the reflectance at visible wavelengths and/or the brightness temperature at infrared wavelengths to clear-sky values. Those pixels that are sufficiently colder and/or more reflective than the clear sky estimates are denoted cloudy. Cloud fraction is then computed by dividing the number of cloudy pixels in a scene by the total number of pixels. Cloud detection schemes differ primarily in which wavelengths are used and how the clear-sky values and temperature and reflectance increments (or thresholds) are determined [*Rossow et al.*, 1985; *Wielicki and Parker*, 1992].

We derive cloud fraction from satellite images using a simple visible reflectance threshold. Marine boundary layer clouds have small vertical extent but are quite reflective, so they exhibit much more contrast with the uniform ocean surface in visible wavelengths than in the infrared. More general cloud detection techniques such as the ISCCP algorithm [*Rossow and Garder*, 1993] also detect shallow, reflective clouds based on their visible wavelength reflectance.

3.3 Comparisons of Cloud Fraction: Sensitivity to Radiance Threshold, Sensor Resolution, and Cloud Type

Although estimates of cloud fraction are available from both surface reports and satellite imagery, the two values represent fundamentally different quantities. Satellite imagery tells us how much of the surface is obscured by clouds ("earth cover") while surface observers report how much of the celestial dome is obscured ("sky cover"). The extent to which the measures should be expected to agree in instantaneous observations is unclear. We therefore use surface estimates of cloud fraction primarily to elucidate the sensitivities inherent in the detection of clouds from space.

Cloud detection algorithms devote substantial effort to identifying the spectral signature of the clear sky, from which thresholds are set for cloud detection. In our visible-wavelength cloud detection scheme, this process reduces to choosing a reflectance threshold separating clear and cloudy pixels. The agreement between satellite and surface estimates of cloud fraction is thus a function of this threshold. Figure 1 shows the difference between satellite and surface estimates of cloud cover, averaged over all pairs of observations, as a function of the reflectance threshold. When the threshold is near zero every pixel (including clear-sky pixels) is considered cloudy, and the average difference corresponds to the average amount of clear sky in the surface observations. This clearly represents an overestimation of cloud cover. As the threshold increases fewer clear-sky pixels are misidentified. Agreement between ship observers and satellite estimates of cloud fraction is best when the reflectance threshold is between 0.09 and 0.10; this agrees with ISCCP algorithms given an ocean albedo of 0.06-0.07.

The amount of cloud determined from satellite imagery depends on the spatial resolution of the satellite sensor [*Wielicki and Parker*, 1992, hereinafter referred to as WP92], especially when threshold-based algorithms are used. We simulate the effects of varying sensor spatial resolution by averaging $N \times N$ groups of full resolution pixels and retrieving cloud fraction from these degraded images; in Figure 1 we show results for full resolution (1 km) pixels and for $N = \{2, 4, 8, 16, 32\}$. WP92 noted that optimal thresholds must be derived in different ways depending on sensor resolution. Full-resolution VISSR pixels are roughly the same size as individual clouds, so the "correct" threshold is that which strikes a balance between identifying partially cloudy pixels as fully cloudy and missing optically thin clouds. In the large pixel limit (WP92, p. 12813) "the 'correct' threshold is a complex func-

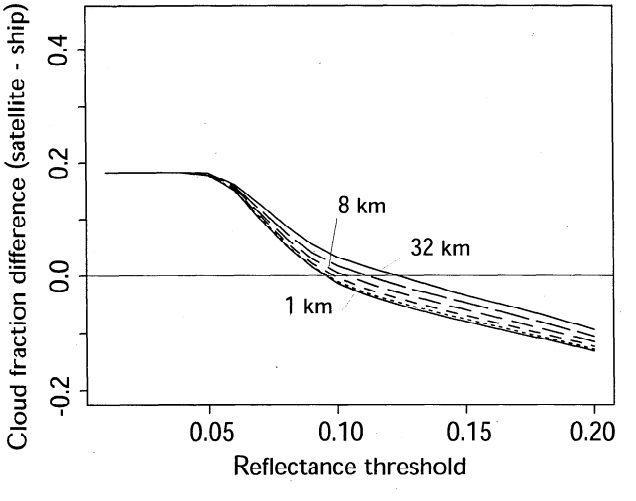

Figure 1. Average difference between satellite-derived and ship-based estimates of cloud fraction, as a function of reflectance threshold, at various sensor spatial resolutions. Cloud fraction is determined from each satellite image by counting the proportion of pixels which exceed the reflectance threshold. The average difference between these estimates and the cloud fraction reported by the corresponding surface observers is shown as a function of threshold. The VISSR instrument has a nominal spatial resolution of 1 km; results are also shown for degraded images in which groups of 2×2, 4×4, 8×8, 16×16, and 32×32 pixels are averaged before cloud fraction is determined. Cloud fraction as estimated from satellite imagery increases as sensor spatial resolution degrades.

tion of pixel size, cloud-cell size, cloud fraction, and the profile of cloud reflectance... from cloud edge to the center of the cloud cell" [WP92]. Our results support this view even at spatial scales exceeding 30 km, although the prevalence of optically thick clouds and high cloud fraction in our scenes makes the dependence of cloud fraction on spatial resolution weaker in our study than in WP92.

The characteristics that affect satellite determination of cloud fraction (i.e. visible features of cloud reflectance hundreds to thousands of meters in extent) are similar to those used by surface observers to identify cloud type. It is therefore unsurprising that the sensitivity of cloud fraction to threshold varies by cloud type. Figure 2 shows this sensitivity by presenting the average difference between satellite-derived and surface-reported values of cloud fraction for all clouds and for subsets of cumulus, stratus, stratocumulus, mixtures of cumulus and stratocumulus, and sky-obscuring fog, as a function of reflectance threshold. Cloud fraction for scenes dominated by cumulus show strong sensitivity to threshold because there are many pixels at cloud edges which are only marginally more reflective than the ocean surface. Stratiform clouds, on the other hand, are both more extensive and optically thicker, so changes in threshold have a smaller impact. WP92 showed similar relationships between cloud type (as inferred from satellite imagery) and the sensitivity of cloud fraction to sensor spatial resolution.

The relationships between cloud type and the amount of cloud detected at a given radiance threshold do not depend strongly on the size of the satellite image. We have repeated the analysis shown in Figure 2 for images 64 km on a side and find that the average amount of cloud detected by the satellite changes by less 0.5% regardless of threshold. Satellite-derived cloud amounts for cumulus increase by about 2.5% in aggregate; this improves the agreement with ship observations at realistic thresholds. Sky-obscuring fog shows an opposite trend, and the changes for other cloud types are less than 1%.

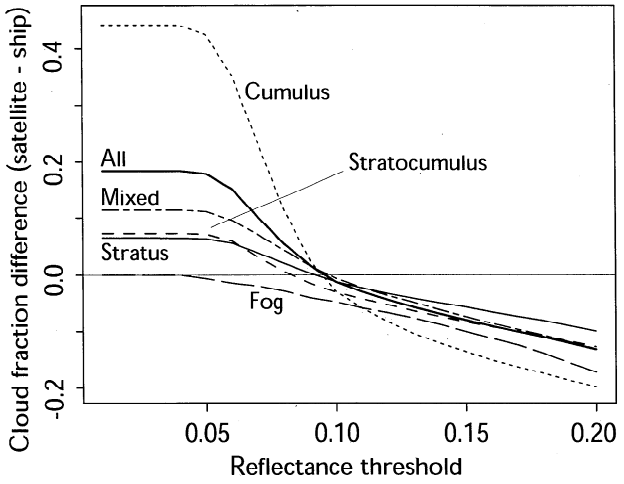

Figure 2. Average difference between satellite-derived and ship-based estimates of cloud fraction, as a function of reflectance threshold, for various cloud types as reported by surface observers. Cumulus are associated with somewhat lower cloud fraction, and the greater fraction of marginally cloudy pixels causes a higher sensitivity to the reflectance threshold. Uniform stratus and stratocumulus are much less sensitive. Fog is assumed by the surface observer to cover the entire scene but is sometimes more localized, yielding lower satellite-derived values of cloud fraction at reasonable thresholds.

4. Parameterizing Distributions of Cloud Optical Thickness

Once the cloudy pixels in each satellite image have been identified, we can determine the probability distribution of cloud optical thickness for each scene. If we assume that there is no net transfer of radiation between pixels (as is consistent with the assumptions underlying the retrieval algorithm), we can compute the albedo of each pixel independently. Averaging over pixels provides the average cloudy-sky albedo, as would be required by a large-scale model. The difference between the albedo computed using the domain averaged optical thickness and the albedo computed using this Independent Pixel Assumption (IPA) [*Cahalan et al.*, 1994] yields the cloudy sky plane parallel homogeneous (PPH) albedo bias:

$$\begin{aligned} B_{PPH} &= R_{PPH} - R_{IPA} \\ &= R(\bar{\tau}, \mu_0) - \overline{R(\tau_{obs}, \mu_0)} \end{aligned} \tag{1}$$

where the overbar denotes averaging over cloudy pixels, and the dependence of albedo on the cosine of the solar zenith angle μ_0 is explicit. Note that B_{PPH} is always greater than 0. The IPA is accurate and provides a benchmark calculation, but it is impractical to implement in a large-scale model because it requires a large number of radiative transfer calculations.

Many of the techniques that exist to circumvent the computational burden of the IPA [*Barker*, 1996; *Cahalan et al.*, 1994; *Oreopoulos and Barker*, 1999] rely on assumptions about the shape of the optical thickness distribution, which is often modeled as either a lognormal or a gamma distribution. With our data set we can assess the utility and accuracy of these two distributions in representing the optical thickness distributions in our data set.

The lognormal distribution is a function of two parameters, which may by determined by computing the first and second moments (mean and variance) of the logarithm of the observed values of optical thickness. Using distribution moments to estimate the parameters of the gamma distribution, however, can lead to serious errors [*Oreopoulos and Davies*, 1998b]. We therefore compute the gamma distribution parameters using a maximum likelihood estimator [see, e.g., *Wilks*, 1995].

4.1 Removing the Plane Parallel Homogeneous Albedo Bias

To test the skill of the lognormal distribution to reduce the PPH bias, we compute the parameters of the lognormal distribution for each scene, then integrate the analytic reflectance function of *Cahalan et al.* [1994] over the lognormal distribution, using the solar zenith angle at the scene center. The difference between this lognormal IPA and the IPA albedo provides the lognormal error:

$$E_{l-n} = \overline{R(\tau_{l-n}, \mu_0)} - \overline{R(\tau_{obs}, \mu_0)} \tag{2}$$

We compute the gamma error in an analogous fashion.

Figure 3 shows histograms of the PPH, lognormal, and gamma albedo errors for the 1331 scenes in our study. A few scenes show PPH biases as large as 0.175, but on average the bias is a moderate 0.025. The magnitude of this bias is smaller than some other studies have shown; we discuss this further in section 6. Both the lognormal and gamma IPA treatments greatly reduce both the range and the mean value of the error. The lognormal IPA is somewhat more effective, producing a range of errors about a factor of 2 and a mean value a factor of 3.5 smaller than the gamma IPA.

In principle, PPH albedo bias is largest for intermediate optical thickness [*Cahalan et al.*, 1994] and increases with the amount of

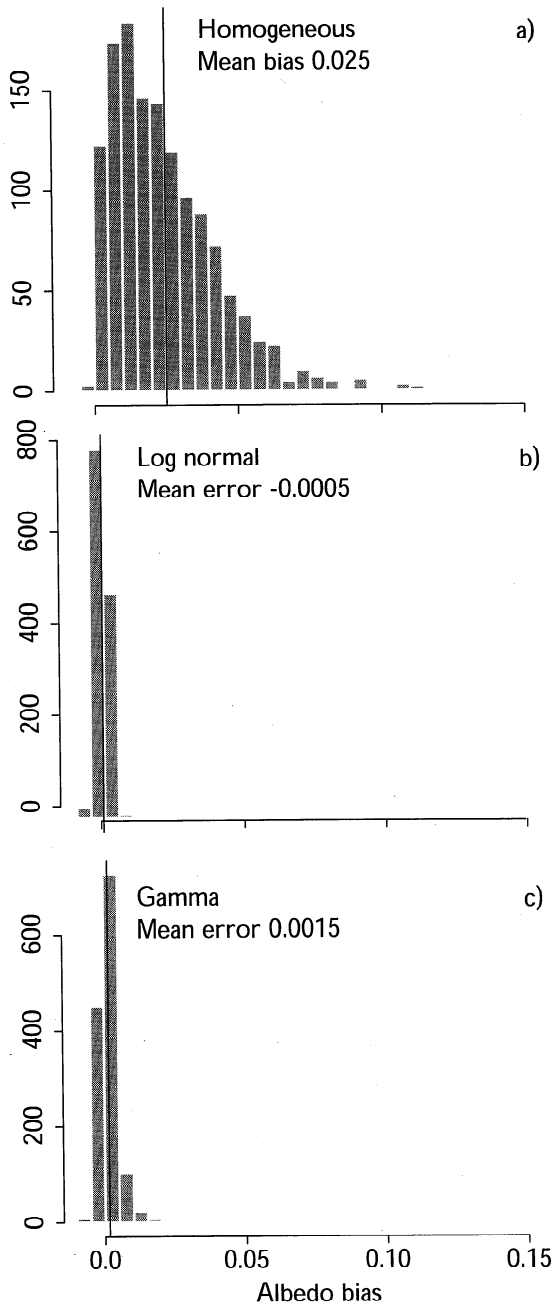

Figure 3. Histogram of (a) albedo bias for each scene computed with the plane parallel homogeneous model, and (b, c) the error made using log-normal and gamma distributions. In each case the albedo error is computed relative to the independent pixel approximation using the observed distribution of optical thickness for each scene. Both the lognormal and gamma distributions nearly eliminate the PPH albedo bias, though the lognormal distribution is somewhat more effective in reducing both the mean error and its range.

variability about the average. In nature, however, the average value and the amount of variability may be related both to each other and to other cloud properties. *Barker et al.* [1996], for example, divided 45 Landsat scenes into three classes denoted overcast, broken, and scattered, in order of decreasing mean optical thickness and increasing cloud fraction. They noted that PPH albedo biases increased as the clouds became thinner and more broken.

In Figure 4 we show joint probability density estimates for the co-occurrence of the mean and variability of the logarithm of optical thickness. Variability is characterized by the dispersion $D(\log\tau) \equiv \sigma(\log\tau)/\overline{\log\tau}$, where σ is the standard deviation. Solid contours indicate the value of the PPH bias calculated with fixed $\mu_0 = 0.8$ using a lognormal distribution at each $\overline{\log\tau}, D(\log\tau)$ pair. The abscissa is also labeled with effective optical thickness $\tau_{eff} \equiv \exp(\overline{\log\tau})$. This quantity provides a second-order accurate scaling of optical thickness for albedo calculations, though transmittance and absorption computations are not as well corrected if the clouds are absorbing (Cairns et al., submitted manuscript, 1998). The scenes in our data set exhibit a wide range of variability and a corresponding range in PPH albedo bias at each value of $\overline{\log\tau}$, but the range of the PPH bias does not increase dramatically with $\overline{\log\tau}$.

4.2 Testing for Goodness of Fit

Both the lognormal and gamma distributions can greatly reduce the PPH bias, but this success alone is not a guarantee that the analytic distributions accurately model the observed distribution of values. To answer this question, we turn to the Kolmogorov-Smirnov goodness of fit test [see, e.g., *Wilks*, 1995]. According to this test, we find that almost none of our fits to both lognormal and gamma distributions are statistically significant at a confidence levels larger than 10%. In other words, we cannot reject the null hypothesis that the observed distribution is drawn from an underlying distribution other than the specified distribution with the parameters we have computed. This somewhat surprising result may be due to a number of causes. First, the VISSR instrument has a relatively coarse 6-bit reflectance digitization, so optical thickness values in each image are clustered around discrete values. For large sample sizes the Kolmogorov-Smirnov test interprets the lack of continuous values of optical thickness as evidence that the fit is incorrect. We also find that in some scenes a few very large values of optical thickness have a disproportionate impact on the distribution parameters. The analytic distributions can still accurately reproduce the IPA albedo, however, because changes in optical thickness have a very small effect on albedo when optical thickness is very large.

5. Relating Variability in Optical Properties to Cloud Type

Although neither the lognormal or gamma functions provide good fits in the strictest statistical sense, their ability to reduce the PPH bias indicates that important aspects of the observed distributions of optical thickness can be summarized in terms of the distribution parameters. In order for this information to be useful in a large-scale model, however, cloud optical properties (including the amount of variability) must be related to quantities available in large-scale models. To this end, we examine the relationships between cloud properties and the information contained in the surface observations; the most robust connections exist between cloud optical properties and cloud type.

We show the distributions of the two parameters used in the lognormal distribution ($\overline{\log\tau}, D(\log\tau)$), and the corresponding distributions of PPH albedo bias, segregated according to the low cloud type reported by the surface observer, in Figure 5. The summary plots code the number of samples in each cloud type (see Table 1) in the width of each box. The median is indicated by a horizontal white line, with a notch showing the 95% confidence interval about the mean (i.e., if the notched areas of two samples do not overlap, we can assert that the medians are different with 95% or greater

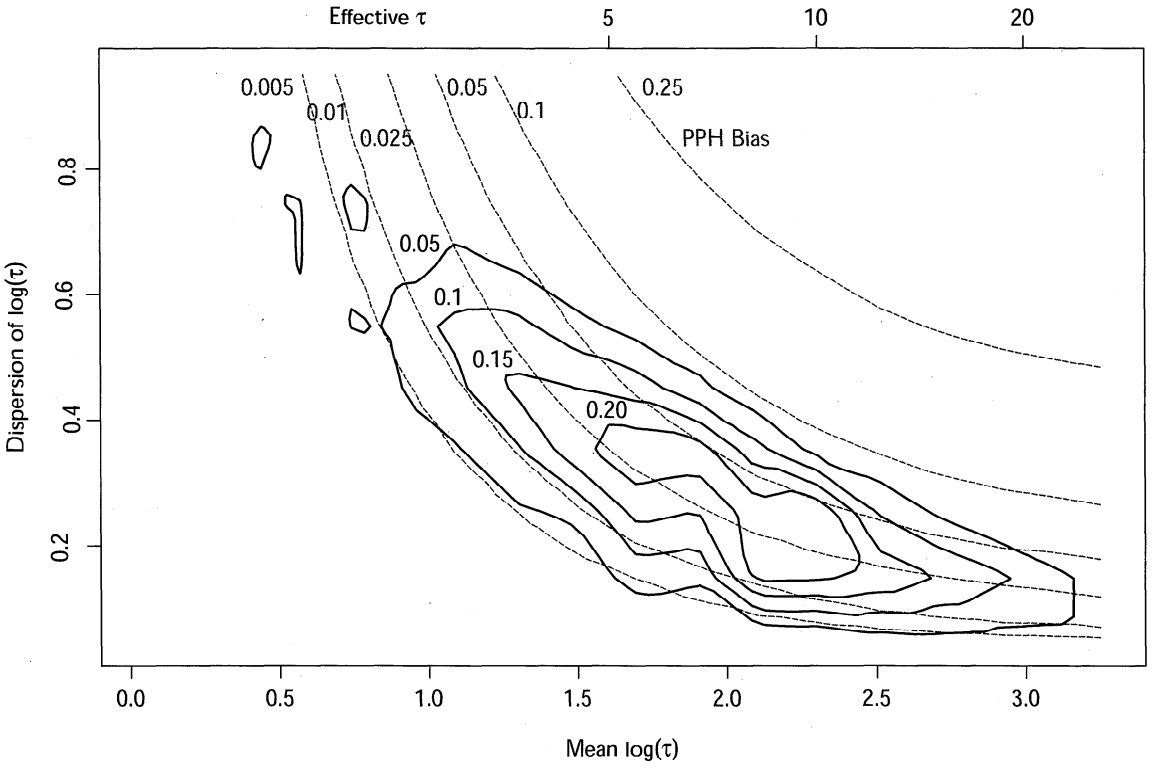

Figure 4. Probability density estimate for joint occurrence of the mean and dispersion (standard deviation divided by the mean) of the logarithm of optical thickness for 1331 images of subtropical marine boundary layer clouds. Also shown (as dashed contours) are the plane parallel homogeneous albedo bias for each value of mean and dispersion. The dispersion, and hence the PPH albedo bias, exhibits a wide range of values at each $\overline{\log\tau}$, but the relative amount of variability decreases as $\overline{\log\tau}$ increases.

confidence). The dark box contains 50% of the observations for each cloud type, while the whiskers extend to the limits of the data. Outliers (points further from the quartile than 1.5 times the interquartile range) are shown separately.

The correspondence between what might be expected for a particular cloud type and what is observed is remarkable, and underscores the value of the long term record from ships of opportunity. In general, cloud optical thickness increases as cloud type changes from cumuliform (C_L 1 and 2) to stratiform (C_L 4, 5, 6, and 8), while the relative amount of variability decreases. Small cumulus are often optically thinner than cumulus with greater vertical extent but are about as variable. Stratocumulus formed by the spreading out of cumulus (C_L 4) are likely to be optically thinner and somewhat more variable than other stratiform cloud types, but the optical properties of stratocumulus are unaffected by the presence of underlying cumulus (compare C_L 4 and C_L 5). Stratus and stratocumulus are indistinguishable. Sky obscuring fog tends to be optically thinner than any other stratiform cloud type. The distribution of PPH albedo bias, however, does not vary substantially by cloud type. With the exception of fog and shallow cumulus, increases in average optical thickness are offset by decreases in the amount of variability.

Comparisons of ISCCP-derived cloud properties with ship observations (C.J. Hahn, W.B. Rossow, and S.G. Warren, ISCCP cloud properties associated with standard cloud types identified in individual surface observations, submitted to *Journal of Climate*, 1998) indicate that cloud radiative properties vary nearly as much within a given cloud type as between cloud types; the distributions in Figure 5 support this view. It is not possible to identify the underlying cloud type based solely on the radiative properties of a scene, but the converse is not true: the parameters governing the distribution of optical thickness vary among cloud types by statistically significant amounts. Given a cloud type, therefore, one could choose a representative value of the mean and dispersion of log optical thickness from within a distinct distribution.

6. Discussion: On the Magnitude of the Plane Parallel Homogenous Albedo Bias

In section 4 we showed that the plane parallel albedo bias in our 1331 scenes has a mean value of $\overline{B_{PPH}} = 0.025$. The PPH bias computed in four related studies equals or exceeds this amount. The spread in PPH bias estimates is primarily a result of differences in observing practices and the methods of relating observations to large-scale models. Before assessing the sources and importance of these differences, however, we first explore the spatial and temporal scales at which horizontal variability is relevant to large-scale models and to remote sensing applications.

6.1 Spatial and Temporal Scales and the PPH Bias

Cloud optical properties exhibit variability over an enormous range of scales in both time and space. Instantaneous spatial variability might bias retrievals of optical thickness which, like large-scale models, are based on the assumption that cloud properties are horizontally uniform within each pixel. Fortunately, we know that this bias is small in unbroken clouds such as those in our sample. Cloud properties are indeed inhomogeneous within each 1-km pixel in our images, but the impact of this variability on our retrievals is negligible for two reasons. In the first place, the amount of variability at spatial scales smaller than 1 km is substantially less

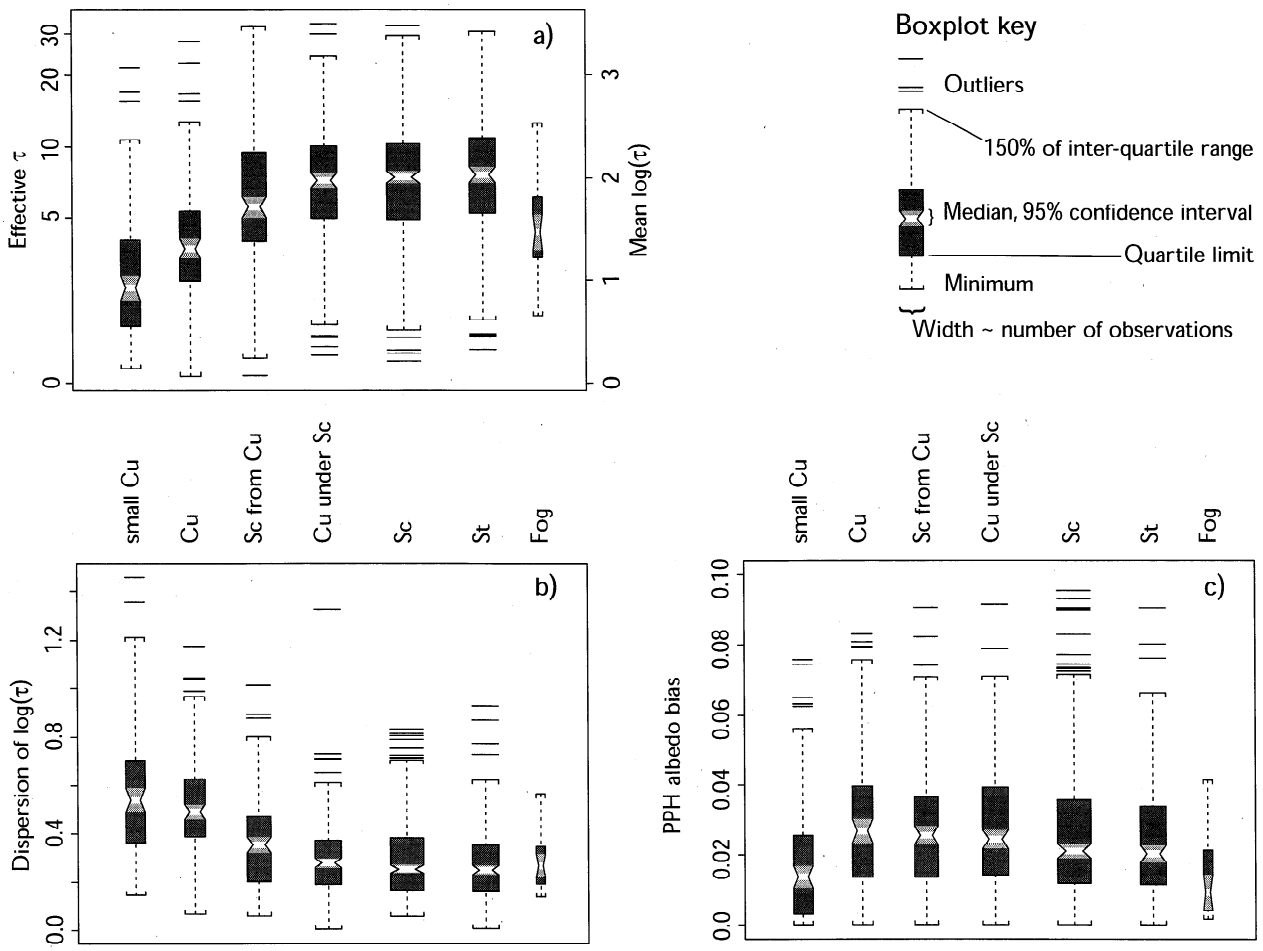

Figure 5. Distributions of the (a) mean and (b) dispersion of the log of optical thickness, and (c) the resultant PPH albedo bias segregated by cloud type as reported by the surface observer. The distribution of optical thickness for each of 1331 scenes is fit to a lognormal distribution, and contributes a single point to this figure. The distributions are displayed with boxplots: the white bar shows the median value, with shaded notches indicating 95% confidence levels about the median; the dark bar contains 50% of the observations; the whiskers extend to the closer of the limits of the data or to 1.5 times the interquartile range. Both effective optical thickness and the dispersion vary widely within each cloud type, but different cloud types have significantly different populations. The distribution of absolute PPH albedo bias, however, is not dependent on cloud type.

than the variability at larger scales. In addition, radiative transfer tends to smooth out spatial structure at scales below the photon mean free path, which is several hundred meters in marine boundary layer clouds [*Marshak et al.*, 1995]. The two effects combine to make the total bias in optical thickness retrievals small for 1-km pixels in unbroken clouds [*Davis et al.*, 1997].

Weather prediction and climate models, however, operate on much larger spatial and temporal scales than remote sensing retrievals. These models exhibit a bias associated with the variability that exists between the radiative smoothing scale and the size of one model grid cell. Our computations of PPH bias for 128-km scenes may be extended to the spatial scales appropriate to specific model grid sizes following *Oreopoulos and Davies* [1998a]. For example, we estimate that the bias within 300-km grid cells is roughly 1.5 times the amount at 128 km. The variability of cloud optical thickness at spatial scales larger than one model grid cell and at temporal scales larger than the interval between radiation calculations (about 3 to 12 hours), will presumably be simulated by prognostic cloud schemes. Indeed, the spatial and temporal variations of cloud water produced by a cloud parameterization might be compared with observations as a means of assessing the performance of the scheme.

6.2 Observations of the PPH Bias

Two studies have employed measurements of liquid water path made by ground based upward looking microwave radiometers. *Cahalan et al.* [1994] accumulated a month's worth measurements made off the coast of California and found that the PPH bias varied between about 0.03 and 0.10, depending on the time of day. A similar study [*Cahalan et al.*, 1995] in the Madeiras, where cloud type is more variable, found values as high as 0.125. These studies report higher values of PPH bias than we in part because they consider substantially larger regions than ours, and the bias increases with the size of the region being considered [*Oreopoulos and Davies*, 1998a]. Aggregation over time compounds this problem by introducing another source of variability. By way of comparison, we compute a PPH bias of 0.063 (using the average solar zenith angle of 50°) if all pixels in all scenes of our sample are combined into a single population. This considerably larger value would only

apply to a large-scale model which called its radiation scheme once a month using the space-time mean optical thickness.

Two additional studies have employed satellite observations of optical thickness to assess the instantaneous PPH bias. *Barker et al.* [1996] employed 45 high spatial resolution Landsat images of marine boundary layer clouds. These images were nearly the same size as our VISSR scenes, and were classified as either overcast, broken, or scattered according to cloud fraction. For overcast clouds the PPH bias was 0.03 (cf. our value of 0.025), but this value increased by about a factor of 3 when cloud fraction decreased. We suggest that this increase is associated with the ability of Landsat to distinguish between small-scale holes and optically thin clouds, as well as three-dimensional radiative transfer effects in the Landsat retrievals. In addition, the average biases reported by *Barker et al.* [1996] are increased by several scenes in which the solar illumination is very oblique, especially for cumulus clouds.

Finally, *Oreopoulos and Davies* [1998a] examined several thousand AVHRR scenes taken over the Atlantic during summer and fall; for $(110\ \mathrm{km})^2$ regions they report a mean PPH bias of 0.08 to 0.17, depending on the assumptions employed. At least two factors contribute to the large bias. With respect to observational strategies, we note that the solar zenith angle is more oblique in most of their scenes than in ours. Since optical thickness retrievals become more uncertain as the solar zenith angle increases [*Pincus et al.*, 1995], low sun angles act to increase the variability within each scene. Additionally, the geographic area considered by *Oreopoulos and Davies* [1998a] is quite large and contains a variety of cloud types (including ice clouds, which are not accounted for in the retrieval scheme), many of which may exist within individual scenes. Some of the PPH bias observed in complex images may be attributed to vertical variability of cloud optical properties (different cloud top and base heights, the presence of clouds at multiple vertical levels, etc.). This vertical variability is already treated in shortwave radiative transfer schemes for large-scale models, which employ cloud overlap assumptions and explicitly consider vertical profiles of cloud fraction.

7. Implications: On the Treatment of Horizontal Variability in Large-Scale Models

The magnitude of the plane parallel homogeneous albedo bias for the clouds we have described is moderate. Nonetheless, we suggest that the bias is still of concern. Systematic errors in the partitioning of energy between reflection, absorption in the atmosphere, and transmission to the Earth's surface may well affect other components of the climate system simulations. We note too that the relatively small PPH albedo bias we report applies only to fair weather marine boundary layer clouds. The large number of scenes we have considered make it likely that our results are robust for this geographic location and season, but they should not be extrapolated to other locations, seasons, or cloud types, where other observations would be welcome.

Our results imply that large-scale models cannot properly address the effects of subgrid scale variability in cloud optical properties without developing techniques for the diagnosis or prediction of that variability. We suggest that consideration of cloud type may be a useful approach. Cloud type might be diagnosed from profiles of temperature and humidity according to observed relationships [*Norris*, 1998a], or by considering the parameterized processes acting to form cloud within a grid cell (i.e., shallow convection versus large-scale saturation). Once cloud type is determined, the amount of variability could be drawn at random from an appropriate population, such as the distributions shown in Figure 5. Such an approach would use information about physical processes occurring in the model more directly than do parameterizations based on cloud fraction [*Barker et al.*, 1996] and/or mean optical thickness [*Oreopoulos and Davies*, 1998b]. Indeed, optical thickness adjustments based on cloud type are already included in some large-scale models [*Rotstayn*, 1997].

One of our goals has been to focus attention on the general problem of unresolved spatial variability in large-scale models of the atmosphere. Many physical processes besides radiative transfer depend nonlinearly on the concentration of water vapor or condensate. The representation of all these processes may benefit from a uniform consideration of spatial variability. Some models already treat the variability in one or more processes: the Hadley Centre Unified Model, for example, uses estimates of the subgrid scale variability of water vapor to predict the grid-averaged relative humidity at which clouds begin to form [S. Cusack, J. M. Edwards, and R. Kershaw, A parametrization of the subgrid-scale variability of saturation for use in GCM cloud schemes, submitted to *Quarterly Journal of the Royal Meteorological Society*, 1998]. Autoconversion may also be a good candidate for such an approach [*Chen and Cotton*, 1987]. We suggest that a unified treatment of subgrid scale variability may help improve the physical basis of a wide variety of model parameterizations.

Acknowledgments. We profited from a series of engaging discussions with Lazaros Oreopoulos, helpful suggestions by Howard Barker, and the thoughtful comments of two anonymous reviewers. This work was begun under three generous fellowship programs: RP held a National Research Council-NASA/GSFC Research Associateship; SAM participated in the NASA/Goddard Summer Institute in Atmospheric and Hydrospheric Sciences, and SAK was a visiting scientist in the Program in Atmospheric and Oceanic Sciences at GFDL. We also received support from the U.S. Department of Energy under grant DE-A105-90ER61069 to NASA's Goddard Space Flight Center as part of the Atmospheric Radiation Measurement program.

References

Barker, H.W., A parameterization for computing grid-averaged solar fluxes for inhomogeneous marine boundary layer clouds, I, Methodology and homogeneous biases, *J. Atmos. Sci.*, *53*, 2289-2303, 1996.

Barker, H.W., B.A. Wielicki, and L. Parker, A parameterization for computing grid-averaged solar fluxes for inhomogeneous marine boundary layer clouds, II, Validation using satellite data, *J. Atmos. Sci.*, *53*, 2304-2316, 1996.

Cahalan, R.F., W. Ridgway, W.J. Wiscombe, T.L. Bell, and J.B. Snider, The albedo of fractal stratocumulus clouds, *J. Atmos. Sci.*, *51*, 2434-2455, 1994.

Cahalan, R.F., D. Silberstein, and J.B. Snider, Liquid water path and plane-parallel albedo bias during ASTEX, *J. Atmos. Sci.*, *52*, 3002-3012, 1995.

Chen, C., and W.R. Cotton, The physics of the marine stratocumulus-capped mixed layer, *J. Atmos. Sci.*, *44*, 2951-2977, 1987.

Davis, A., A. Marshak, R. Cahalan, and W. Wiscombe, The Landsat scale break in stratocumulus as a three-dimensional radiative transfer effect: Implications for cloud remote sensing, *J. Atmos. Sci.*, *54*, 241-260, 1997.

Del Genio, A.D., M.-S. Yao, W. Kovari, and K.K.W. Lo, A prognostic cloud water parameterization for global climate models, *J. Climate*, *9*, 270-304, 1996.

Fowler, L.D., D.A. Randall, and S.A. Rutledge, Liquid and ice cloud microphysics in the CSU general circulation model, I, Model description and simulated microphysical processes, *J. Climate*, *9*, 489-529, 1996.

Hahn, C.J., S.G. Warren, and J. London, Edited synoptic cloud reports from ships and land stations over the globe, 1982-1991, Dep. of Energy, Washington, DC, 1996.

Henderson-Sellers, A., G. Seze, F. Drake, and M. Desbois, Surface-observed and satellite-retrieved cloudiness compared for the 1983 ISCCP special study area over Europe, *J. Geophys. Res.*, *92*, 4019-4033, 1987.

Marshak, A., A. Davis, W. Wiscombe, and R. Cahalan, Radiative smoothing in fractal clouds, *J. Geophys. Res.*, *100*, 26247-26261, 1995.

Norris, J.R., Low cloud type over the ocean from surface observations, I, Relationship to surface meteorology and the vertical distribution of temperature and moisture, *J. Atmos. Sci.*, *11*, 369-382, 1998a.

Norris, J.R., Low cloud type over the ocean from surface observations, II, Geographic and seasonal variations, *J. Atmos. Sci.*, *11*, 383-403, 1998b.

Oreopoulos, L., and R. Davies, Plane parallel albedo biases from satellite observations, I, Dependence on resolution and other factors, *J. Climate*, *11*, 919-932, 1998a.

Oreopoulos, L., and R. Davies, Plane parallel albedo biases from satellite observations, II, Parameterizations for bias removal, *J. Climate*, *11*, 933-944, 1998b.

Oreopoulos, L., and H.W. Barker, Accounting for subgrid-scale cloud variability in a multi-layer, 1D solar radiative transfer algorithm, I, Model development, *Q. J. R. Meteorol. Soc.*, *126*, 301-330, 1999.

Pincus, R., M. Szczodrak, J. Gu, and P. Austin, Uncertainty in cloud optical depth estimates made from satellite radiance measurements, *J. Climate*, *8*, 1453-1462, 1995.

Rossow, W.B., and L.C. Garder, Cloud detection using satellite measurements of infrared and visible radiances for ISCCP, *J. Climate*, *6*, 2341-2369, 1993.

Rossow, W.B., F. Mosher, E. Kinsella, A. Arking, M. Desbois, E. Harrision, P. Minnis, E. Ruprecht, G. Seze, C. Simmer, and E. Smith, ISCCP cloud algorithm intercomparison, *J. Climate Appl. Meteor.*, *24*, 877-903, 1985.

Rossow, W.B., Y. Desormeaux, C.L. Brest, and A.W. Walker, International Satellite Cloud Climatology Project (ISCCP) Radiance Calibration Report, World Climate Res. Program, Geneva, 1992.

Rotstayn, L.D., A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models, *Q. J. R. Meteorol. Soc.*, *123*, 1227-1282, 1997.

Tiedtke, M., Representation of clouds in large-scale models, *Mon. Weather Rev.*, *121*, 3041-3061, 1993.

Wielicki, B.A., and L. Parker, On the determination of cloud cover from satellite sensors: The effect of sensor spatial resolution, *J. Geophys. Res.*, *97*, 12799-12823, 1992.

Wilks, D.S., *Statistical Methods in the Atmospheric Sciences: An Introduction*, 467 pp., Academic, San Diego, Calif., 1995.

Woodruff, S.D., R.J. Slutz, R.L. Jenne, and P.M. Steurer, A comprehensive ocean-atmosphere data set, *Bull. Am. Meteorol. Soc.*, *68*, 1239-1250, 1987.

World Meteorological Organization, Manual on codes, *WMO Publ. 306*, World Meteorol. Org., Geneva, 1974.

S. A. Klein, Geophysical Fluid Dynamics Laboratory, Princeton University, Box 308, Princeton, NJ, 08542. (email: sak@gfdl.gov)

S. A. McFarlane, Program in Atmospheric and Oceanic Science, Box 311, University of Colorado, Boulder, CO, 80309. (email: mcfarlan@nit.colorado.edu)

R. Pincus, Department of Atmospheric and Oceanic Sciences, 1225 W Dayton St., University of Wisconsin, Madison, WI, 53706. (email: robert@meteor.wisc.edu)

(Received July 12, 1998; revised November 25, 1998; accepted December 3, 1998.)