Overlap of fractional cloud for radiation calculations in GCMs: A global analysis using CloudSat and CALIPSO data

Howard W. Barker[1]

[1] Assumptions made by global climate models (GCMs) regarding vertical overlap of fractional amounts of clouds have significant impacts on simulated radiation budgets. A global survey of fractional cloud overlap properties was performed using 2 months of cloud mask data derived from CloudSat-CALIPSO satellite measurements. Cloud overlap was diagnosed as a combination of maximum and random overlap and characterized by vertically constant decorrelation length $\mathcal{L}^*_{cf}$. Typically, clouds overlap between maximum and random with smallest $\mathcal{L}^*_{cf}$ (medians $\rightarrow$ 0 km) associated with small total cloud amounts $\widehat{C}$, while the largest $\mathcal{L}^*_{cf}$ (medians $\sim$3 km) tend to occur at $\widehat{C}$ near 0.7. Global median $\mathcal{L}^*_{cf}$ is $\sim$2 km with a slight tendency for largest values in the tropics and polar regions during winter. By crudely excising near-surface precipitation from cloud mask data, $\mathcal{L}^*_{cf}$ were reduced by typically <1 km. Median values of $\mathcal{L}^*_{cf}$ when Sun is down exceed those when Sun is up by almost 1 km when cloud masks are based on radar and lidar data; use of radar only shows minimal diurnal variation but significantly larger $\mathcal{L}^*_{cf}$. This suggests that sunup inferences of $\mathcal{L}^*_{cf}$ might be biased low by solar noise in lidar data. Cloud mask cross-section lengths L of 50, 100, 200, 500, and 1000 km were considered. Distributions of $\mathcal{L}^*_{cf}$ are mildly sensitive to L thus suggesting the convenient possibility that a GCM parametrization of $\mathcal{L}^*_{cf}$ might be resolution-independent over a wide range of resolutions. Simple parametrization of $\mathcal{L}^*_{cf}$ might be possible if excessive random noise in $\widehat{C}$, and hence radiative fluxes, can be tolerated. Using just cloud mask data and assuming a global mean shortwave cloud radiative effect of -45 W m^{-2}, top of atmosphere shortwave radiative sensitivity to $\mathcal{L}^*_{cf}$ was estimated at 2 to 3 W m^{-2} km^{-1}.

1. Introduction

[2] Owing to speed and size limitations of most computers, Global Climate Models (GCMs) discretize the Earth-atmosphere system into columns with horizontal cross-sectional areas generally exceeding 10^4 km^2. Hence, many processes and fluctuations are unresolved by GCMs and must be parametrized. Although GCM columns are partitioned into relatively thin slabs, the horizontal expanse of these slabs makes it difficult to describe the statistical nature of vertical correlations of processes and fluctuations. A longstanding line of work on this problem involves representation of horizontal and vertical fluctuations of clouds for the purpose of computing column-wide (i.e., domain-average) radiative flux profiles.

[3] Up until the late 1970s, radiative flux profiles were computed assuming that the horizontal position of a uniform slab cloud in a cloudy layer was completely uncorrelated with like clouds in all other layers [*Manabe and Strickler*, 1964]. This is the random overlap model. In 1979, *Geleyn and Hollingsworth* [1979] introduced the maximum-random overlap (MRO) model and an accompanying radiative flux algorithm. The MRO method assumes that clouds in adjacent layers overlap maximally while those separated by a cloudless stretch overlap randomly. Through the 1980s the MRO model received some attention [e.g., *Morcrette and Fouquart*, 1986; *Tian and Curry*, 1989] and increased much in popularity during the 1990s. Currently, approximately half of the operational weather and climate models employ the MRO model [*Barker et al.*, 2003; Q. Fu, personal communication, 2004].

[4] The MRO method is, however, untenable simply because it does not describe cloud overlap well [e.g., *Barker et al.*, 1999; *Hogan and Illingworth*, 2000; *Mace and Benson-Troth*, 2002]; it is an inflexible scheme that does not accommodate inclusion of well-known cloud structural characteristics, and, like random overlap, its performance depends much on GCM vertical resolution [*Barker et al.*, 1999; *Oreopoulos and Khairoutdinov*, 2003; *Stephens et al.*, 2004].

[1]Cloud Physics and Severe Weather Research Section, Environment Canada, Toronto, Ontario, Canada.

[5] *Hogan and Illingworth* [2000] combined the components of MRO via linear combination in which the weighting factor depends on distance between layers and a decorrelation length that effectively eliminates the dependence of cloud overlap on vertical resolution. Shortly thereafter, the Monte Carlo Independent Column Approximation (McICA) for computing broadband radiative fluxes emerged [*Barker et al.*, 2002; *Pincus et al.*, 2003] and has set a new standard for GCM radiation routines [*Pincus et al.*, 2005; *Räisänen et al.*, 2005; *Morcrette et al.*, 2008]. McICA utilizes stochastic subcolumn cloud generators [e.g., *Räisänen et al.*, 2004] that accommodate nicely Hogan and Illingworth's model. In the context of McICA, *Räisänen et al.* [2004] and *Barker and Räisänen* [2005] examined output from an array of cloud system-resolving models (CSRMs) [*Randall et al.*, 2003] and concluded, in accord with previous studies, that overlap and horizontal variance of water content must be treated together and that cloud overlap generally lies between maximum and random. The latter conclusion was also arrived at using data from both surface-based cloud remote sensing instruments [*Hogan and Illingworth*, 2000; *Mace and Benson-Troth*, 2002] and CSRM data [*Barker et al.*, 1999; *Oreopoulos and Khairoutdinov*, 2003; *Pincus et al.*, 2005].

[6] The purpose of this study was to assess cloud fraction overlap on a global scale using data collected from the CloudSat and Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellites [*Stephens et al.*, 2002; *Winker et al.*, 2003]. Both satellites have been orbiting in the A-train since May 2006, and suitable data for this study became available in July 2007. The following section explains the motives, and outlines the objectives, of this study. It is followed by a description of the data and the method of analysis, with results and conclusions forming the final sections.

2. Motivation and Objective

[7] According to the independent column approximation (ICA), radiative flux averaged across a GCM grid cell can be expressed as

$$F_{\text{ICA}} = \int_0^\infty p(\tau)F(\tau)\,d\tau \qquad (1)$$

where τ is cloud optical depth, $F(\tau)$ is a multilayer solution of the 1-D radiative transfer equation, which of course depends on many variables other than τ, and $p(\tau)$ is the normalized density function of τ over the domain [*Stephens et al.*, 1991; *Barker*, 1996]. If

$$p(\tau) = \left(1 - \widehat{C}\right)\delta(\tau) + \widehat{C}\,\widehat{p}(\tau)$$

where $\widehat{C}$ is total cloud fraction and $\widehat{p}(\tau)$ is a normalized density function of τ for the cloudy portion of the domain, (1) becomes

$$F_{\text{ICA}} = \left(1 - \widehat{C}\right)F(0) + \widehat{C}\int_0^\infty \widehat{p}(\tau)F(\tau)\,d\tau. \qquad (2)$$

If the atmosphere is divided into M layers, (1) can be rewritten as

$$F_{\text{ICA}} = \left(1 - \widehat{C}\right)F(0) + \sum_{m=1}^M \varepsilon_m \int_0^\infty \widehat{p}_m(\tau)F(\tau)\,d\tau \qquad (3)$$

where ε_m is fraction of the domain having cloudtops in layer m exposed to space, $\widehat{C} = \sum_{m=1}^M \varepsilon_m$, and $\widehat{p}_m(\tau)$ is frequency distribution of τ for clouds of any geometric thickness but whose tops are exposed to space in layer m. The attraction of this representation of the ICA is that $\widehat{C}, \varepsilon_m$, and $\widehat{p}_m(\tau)$ are seen clearly as the basic cloud structural quantities needed to estimate radiative fluxes, they all have the potential of being inferred, on a global scale, from satellite data, and can be diagnosed easily in a GCM. Moreover, if a GCM simulates these quantities, and hence TOA fluxes, well, it follows that corresponding flux profiles are likely to be accurate too.

[8] While GCMs, do not estimate ε_m and $\widehat{p}_m(\tau)$ explicitly, they are present implicitly at every timestep and depend on fractional amounts and mass of layer clouds, as well as assumptions regarding vertical overlap of layered clouds and horizontal fluctuations of water content. In the McICA method [*Barker et al.*, 2002], with its reliance on stochastic generation of unresolved clouds [*Räisänen et al.*, 2004], exact computation of ε_m and $\widehat{p}_m(\tau)$ is simple. Currently within GCMs, cloud overlap and horizontal variability are either assigned or highly parametrized. According to analyses by *Barker et al.* [1999] and *Barker and Räisänen* [2005], it is crucial that both descriptors be addressed simultaneously, particularly when computing shortwave radiative fluxes. The lingering question is how accurate and sophisticated must parametrizations of these descriptors be? To some extent, the answer is governed by the implicit sunset clause built into this line of research: ever-shrinking grid spacings in GCMs will give way to global CSRMs whereupon properties such as cloud overlap will be reduced to diagnostic measures and effectively emancipated from parametrization. For the foreseeable future, however, parametrizations of these descriptors are required, but as yet there are no global analyses of them so the required level of parametrization is still not clear. It appears, however, that the basic cloud masks derived from CloudSat-CALIPSO data should be able to furnish an initial perspective on the global distribution of characteristics pertaining to, at least, overlap of fractional cloud.

3. Data

[9] To illustrate some global results of cloud overlap properties, data collected by CloudSat's 94-GHz cloud-profiling radar (CPR) and CALIPSO's dual wavelength lidar during January 2007 and August 2007 were examined. Both satellites were launched together on 29 April 2006, spent a few weeks manoeuvring into position with the A-train, and began transmitting data about a month after launch. The months selected here have among the largest amounts of data to work with and almost maximize seasonal disparity. Analyses were performed using CloudSat's

Table 1. Values Assigned to Each Radar Volume in CloudSat's *CPR_Cloud_mask* Field[a]

Value	Explanation
0	no cloud detected
1	likely bad data
5	likely ground clutter
5–10	weak detection found using along track integration
20–40	cloud detected; the larger the value, the smaller the chance of a false detection

[a]For this study, values greater than 20 were taken as cloud.

CPR_Cloud_mask and *Radar_Reflectivity* fields from the 2B-GEOPROF database and the *CloudFraction* field from 2B-GEOPROF-LIDAR [*Mace*, 2006, 2007]. (These data are available at http://cloudsat.cira.colostate.edu/dataSpecs.php.) In essence, *CPR_Cloud_mask* and *CloudFraction* indicate the likelihood of cloud in a radar volume. Table 1 lists the values that *CPR_Cloud_mask* can take. *CloudFraction* is the fraction of lidar volumes in a radar volume identified as containing hydrometeors.

[10] There are typically 37,081 CPR profiles reported per orbit implying that each column is effectively 1.1 km long (along-track). In actuality, the radar footprint is ∼1.4 km long and integrated for ∼1 km. Clearly there is some oversampling. They are approximately 1.4 km wide (across-track) and each volume is 0.24 km deep. Because of ground-clutter contamination, the lowest 2 or 3 volumes are unworkable. According to *Mace* [2007, Table 1], resolutions of the processed lidar data are: in the along-track direction 1 km and 0.3 km for altitudes below and above 8.2 km, respectively; in the vertical direction 0.03 km and 0.075 km for altitudes below and above 8.2 km, respectively;

1. locate the surface bin J_{srf} ; provided in 2B–GEOPROF files

2. is $CM(J_{\mathrm{srf}} - 3)$ or $CM(J_{\mathrm{srf}} - 4) \geq 20$?   ; precipitation near surface?

3a. if 2 = .false., move to next column and go to 1   ; no near–surface precipitation

3b. else, find bin $J_{\max}$ with $\max_j (Z(j))$; find bin with maximum radar refl.

4. set $CM(J_{\mathrm{srf}}, \ldots, J_{\max}) = 0$ for $CF(j) < 0.99$; make cells from J_{srf} to $J_{\max}$ cloudless

5. move to next column and go to 1

$$(5)$$

and 0.3 km across-track. For this study, a CPR volume was classed as *cloud* if

$$
\begin{cases}
CPR\_Cloud\_mask \geq 20 \\
Z \geq -30 \ \mathrm{dBZ} \\
CloudFraction \geq 99\%
\end{cases}
\tag{4}
$$

were all satisfied, where Z is radar reflectivity; CloudSat's minimum detectable signal is purportedly −30 dBZ [*Stephens et al.*, 2002; D. Vane, personal communication, 2007].

[11] Given that CloudSat's radar reflectivity responds to the sixth power of particle size [*Doviak and Zrnić*, 1993], *CPR_Cloud_mask* captures most of the precipitation but certainly not all "cloud." The concern is easy to appreciate: almost by definition, precipitation displays stronger vertical correlation than cloud, and so CloudSat-derived estimates of vertical correlation for *cloud* might be biased high. This is potentially important because GCM cloud-radiation algorithms do not account for precipitation which has shorter atmospheric residence times than clouds and relatively week extinction at wavelengths important for Earth's radiation budget. Hence, as long as GCM cloud-radiation routines disregard precipitation, ideal parametrization of cloud overlap based on observations should have precipitation factored out.

[12] This points to the distinction between *intrinsic* and *extrinsic* cloud properties that appears whenever one attempts to reconcile models with observations [e.g., *Stephens*, 1988]. At the time of writing, retrievals pertaining to cloud water content and particle size were available from CloudSat's database, but the intermittency of the data made it difficult to use as a screen for precipitation (or at least volumes with large particles and small visible extinction). Hence, knowing that CloudSat's *CPR_Cloud_mask* certainly contains precipitation in addition to cloud, a very simple and crude precipitation screen was applied to *CPR_Cloud_mask* in an attempt to remove at least shafts of precipitation that approached the surface. This procedure is not being advocated for general use; it was used here simply to illustrate the impact on estimates of cloud overlap due to removal of what was likely precipitation approaching/reaching the surface.

[13] The precipitation screen consists of using (4) and then for each column, where $CM(j)$ and $CF(j)$ represent values of *CPR_Cloud_mask* and *CloudFraction* in the jth bin, applying this sequence of tests:

Values of CM are checked in the third and fourth bins above the surface cell so as to avoid ground-clutter contamination. The declouding process up to bin $J_{\max}$ is admittedly an arbitrary choice that likely catches several volumes that contained a mixture of precipitation and cloud. However, by leaving volumes unaltered if their lidar *CloudFraction* is >0.99 should help avoid eliminating cells with significant visible extinction.

[14] Cross sections of length 50, 100, 200, 500, and 1000 km were used to span GCM-size domains. The issue of optimal length of cross section to represent the statistics of the (square) domain from which it was drawn has been

addressed in the past with respect to total cloud fraction [e.g., *Astin and Di Girolamo*, 1999]. It will be addressed for cloud overlap in a separate study.

4. Cloud Fraction Overlap

4.1. Methodology

[15] Throughout this analysis, overlap of clouds is treated as a linear combination of maximum and random overlap [*Hogan and Illingworth*, 2000; *Bergman and Rasch*, 2002]. To illustrate, if c_k and c_l are fractional amounts of cloud in layers k and l at altitudes z_k and z_l, then total, vertically projected cloud fraction $c_{k,l}$ for the two layers is defined as

$$c_{k,l} = \alpha_{k,l} \underbrace{\max(c_k, c_l)}_{\text{maximum overlap}} + (1 - \alpha_{k,l}) \underbrace{(c_k + c_l - c_k c_l)}_{\text{random overlap}};$$

$$\alpha_{k,l} \in \left[\frac{\max(c_k, c_l)}{\max(c_k, c_l) - 1}, 1 \right] \qquad (6)$$

where $\alpha_{k,l}$ is the weighting parameter. The maximum component is unambiguous: positions of clouds in one layer are determined by, and overlay, those in the other. The random component is the expectation value for the distribution of total cloud fraction that occurs when the positions of clouds in both layers are completely uncorrelated. When clouds overlap less than expected when cloud positions in one layer are completely independent of those in the other,

$$c_{k,l} > c_k + c_l - c_k c_l$$

which leads to $\alpha_{k,l} < 0$ in (6). Hence, the smallest values that $\alpha_{k,l}$ can attain are

$$\alpha_{k,l} = \frac{\max(c_k, c_l)}{\max(c_k, c_l) - 1}$$

which occur when clouds do not overlap at all and $c_{k,l} = c_k + c_l \leq 1$. Clearly,

$$\lim_{\max(c_k, c_l) \to 1} \alpha_{k,l} = -\infty.$$

[16] It has become almost customary, and with fairly good reason [see *Astin and Di Girolamo*, 2006], to define

$$\alpha_{k,l} \equiv \exp\left[-\int_{z_k}^{z_l} \frac{dz}{\mathcal{L}_{cf}(z)} \right], \qquad (7)$$

where $\mathcal{L}_{cf}$ is decorrelation length for overlapping fractional cloud which, in general, varies with z. Note that this definition does not admit $\alpha_{k,l} < 0$ [cf. *Mace and Benson-Troth*, 2002; *Oreopoulos and Khairoutdinov*, 2003] and so $c_{k,l}$ is always within [max (c_k, c_l), $c_k + c_l - c_k c_l$]. Nor does it allow $\widehat{C} = 1$ when max $(c(z)) < 1$, where $c(z)$ is layer cloud fraction profile; though these cases can, and do, occur.

[17] Within most conventional GCMs, all the relevant information one has pertaining to cloud structure inside a column is $c(z)$. There are, therefore, infinitely many ways to overlap the fractions $c(z)$ and thus produce corresponding total cloud fractions C that can, in principle, take values within [max$(c(z))$, 1]. With (6) and (7), a convenient, and potentially useful, way to assess overlap for nonovercast scenes, is to enforce a vertically invariant $\mathcal{L}_{cf}$ on $c(z)$ [*Barker and Räisänen*, 2005]. This leads to a unique function $C(\mathcal{L}_{cf}) \in$ [max$(c(z))$, 1]. This function is, by definition, nonanalytic, but can be evaluated numerically using a stochastic cloud generator. The generator used here was developed originally by *Räisänen et al.* [2004] for use with McICA [*Barker et al.*, 2002; *Pincus et al.*, 2003]. Unless stated otherwise, all applications of the generator used 25,000 subcolumns.

[18] When dealing with a 3-D field of cloud or a 2-D cross section, obtained from either a CSRM or observations, one can compute the domain's actual total cloud fraction $\widehat{C}$ (along with $c(z)$, and even $\mathcal{L}_{cf}(z)$), and thus define an effective decorrelation length $\mathcal{L}_{cf}^*$ as the solution to

$$C(\mathcal{L}_{cf}^*) = \widehat{C}. \qquad (8)$$

Again, however, this cannot be done with GCM data for all one has to work with is $c(z)$.

[19] One way to solve for $\mathcal{L}_{cf}^*$ is Brent's method [*Brent*, 1973] which is ideal when computed values of a function are all that are known, and roots are one-dimensional. Brent's method combines the safety of bisection methods with the convergence speed of higher-order methods, provided the root lies between extremes set by the user. The extremes used here were 0 km and 20 km. If $\mathcal{L}_{cf}^* > 20$ km the algorithm defaults to 20 km, which is often very close to maximum overlap. The convergence criterion was 0.05 km. It was set small to diminish the impact of stochastic fluctuations in returned values of C. The number of iterations required for convergence was typical 7 or 8 but ranged from 5 to 12.

[20] Figure 1 shows a comparison of roots between those based on the settings just mentioned and a benchmark that used an upper limit of 100 km, a convergence criterion of 0.01 km, and 250,000 subcolumns. The latter required generally 10 to 15 iterations and took about 20 times more CPU time than the operational settings. For $\mathcal{L}_{cf}^* < 20$ km, the operational settings lead to errors in $\mathcal{L}_{cf}^*$ that are typically less than 0.1 km, well below the resolution of the data and thus adequate for the purpose at hand. While errors for $\mathcal{L}_{cf}^* > 20$ km are exclusively bias, and of magnitude $\mathcal{L}_{cf}^* - 20$ km, only 50 of the 11,498 cross sections for $L = 1000$ km had $\mathcal{L}_{cf}^* > 20$ km. Moreover, based on experience [*Barker and Räisänen*, 2005], $dF_{\text{ICA}}/d\mathcal{L}_{cf}^*$ are usually very small for $\mathcal{L}_{cf}^* > 20$ km and so errors are negligible. While 1000 km transects were used to produce Figure 1, samples for shorter cross sections yielded virtually identical results.

4.2. Example

[21] Before moving to the global analysis, consider an example of the methodology proposed in the previous subsection and the impact of the precipitation screen.

[22] Figure 2 contrasts (4) with its precipitation-screened counterpart for a typical 1000 km sample drawn from an orbit on 1 January 2007. Clearly, this screening goes too far in some cases and not far enough in others, but generally speaking, it appears to move matters in the desired direc-
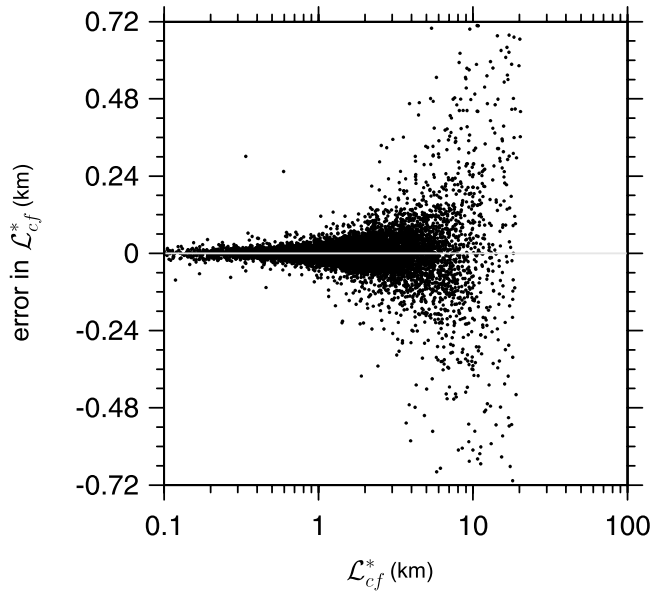
Figure 1. Error in the estimation of $\mathcal{L}_{cf}^*$ via Brent's root-finding method when the convergence criterion was 0.1 km and 25,000 subcolumns were used in the stochastic cloud generator. The benchmark solution was taken to be a convergence criterion of 0.01 km along with 250,000 subcolumns in the stochastic cloud generator.

tion. The plot of layer cloud fraction profile in the lower left corner of Figure 2 shows that screening eliminated volumes up to about 3 km above the surface. Since application of (5) never increases $\widehat{C}$, and rarely reduces it (it went from 0.876 to 0.875 in this case), sizable reductions in layer cloud fractions mean that there is less cloud in a profile needed to make-up $\widehat{C}$. In general this means that screening will reduce $\mathcal{L}_{cf}^*$; in this instance, solving (8) for the unscreened and screened profiles yielded $\mathcal{L}_{cf}^*$ of 2.0 km and 1.5 km, respectively. When these values of $\mathcal{L}_{cf}^*$ are used in the stochastic generator they produce fields that by design, had the correct profile of layer cloud fraction and associated $\widehat{C}$, but because $\mathcal{L}_{cf}^*$ was set constant with height, profiles of cumulative cloud fraction were slightly incorrect, as seen in the central plot in Figure 2. The rightmost plot in Figure 2 shows, however, that the generated fields have fractional amounts of cloud exposed to space, ε_m, that not only approximate the observed amounts well but that are impacted only slightly by alteration of $\mathcal{L}_{cf}^*$ induced by precipitation screening.

[23] To further this example, Figure 3 shows the function $C(\mathcal{L}_{cf})$ for both the unscreened and screened cases shown in Figure 2. It also shows the progression of steps in Brent's method using 25,000 subcolumns in the generator. The left plot in Figure 4 shows cumulative distribution for $\mathcal{L}_{cf}^*$ for the unscreened case. It deviates slightly from a step function because of the generator's intrinsic stochastic noise. Median $\mathcal{L}_{cf}^*$ is 2.02 km and interquartile range is just 0.03 km. Moreover, 90% of the solutions took 8 or fewer iterations to converge; 0.3% took 11. Use of this distribution of roots in the stochastic generator, this time using 100,000 subcolumns, yields the cumulative distribution for C shown in the right plot of Figure 4. The actual value is 0.876 while the

median of the generated fields is 0.8756. The interquartile range is just 0.004. Interquartile ranges for the screened counterpart are almost identical. By almost all standards, uncertainties of these magnitudes are inconsequential.

5. Results

5.1. Global Overlap Characteristics From CloudSat and CALIPSO Data

[24] Figure 5 shows frequency distributions of $\mathcal{L}_{cf}^*$ conditional upon $\widehat{C}$ obtained from CloudSat-CALIPSO 2-D cross-sections of lengths $L = 200, 500,$ and 1000 km for all transects with total cloud fractions between 0.05 and 0.99 observed during January 2007. These cross sections did not have the precipitation screen applied. The limits on $\widehat{C}$ were imposed, and used hereinafter, because of poor statistics expected at very small $\widehat{C}$ and $\mathcal{L}_{cf}^*$ being undefined at $\widehat{C} = 1$; overlap of nonovercast layers when $\widehat{C} = 1$ is addressed later. Values were composited into six latitudinal bands as listed on the plots. For $L = 200$ km there were 33,250 suitable cross sections, while for $L = 1000$ km there were 11,139. As expected, distributions for both variables broaden as L decreases; specially those for $\widehat{C}$ which become increasingly J- and U-shaped [*Rossow*, 1989]. Without exception, all distributions exhibit peak median values of $\mathcal{L}_{cf}^*$ at intermediate values of $\widehat{C}$. The largest values are in the northern polar region where the 75th percentile (or third quartile) of $\mathcal{L}_{cf}^*$ reach upward of 6 km for $\widehat{C}$ near 0.6, which signifies strong vertical correlation, while at the other end the 25th percentile hovers near 1 km for most locations regardless of L.

[25] Figure 6 shows zonal medians and interquartile ranges of $\mathcal{L}_{cf}^*$ for several cross-section lengths L. There is very little suggestion that the median depends much on L. The interquartile range, however, clearly increases with decreasing L. This is due exclusively to the emergence of tails in the distribution toward large $\mathcal{L}_{cf}^*$ as few collimated clouds within short distances give rise to some enormous $\mathcal{L}_{cf}^*$ (i.e., near-maximum overlap). Sample sizes used to create Figure 6 are shown in Figure 7. Note that they do not scale linearly with L. This is because as L decreases an increasing number of disregarded clear and overcast skies cross sections are encountered.

[26] Given that at least medians of $\mathcal{L}_{cf}^*$ depend weakly on L [cf. *Mace and Benson-Troth*, 2002], the majority of results shown hereinafter are for $L = 500$ km. This length probably provides adequate sampling for typical GCM grid cells [cf. *Astin and Di Girolamo*, 1999] and provides fairly large numbers of samples.

[27] Figure 8 indicates that median values of $\mathcal{L}_{cf}^*$ as functions of $\widehat{C}$ can differ by 0.5 km to 1 km between land and ocean for January 2007 with perhaps a slight tendency for larger values over ocean. Why this should be the case is not clear at the moment. Figure 9 suggests that for extratropical areas, values of $\mathcal{L}_{cf}^*$ during winter tend to be about 0.5 km to 1 km larger than during summer with the seasonal disparity increasing poleward. Near 30°N the values presented here are somewhat smaller, particularly during summer, then values shown by *Mace and Benson-Troth* [2002] for CPR data gathered at ARM's Southern Great Plains site and do not display strong seasonality as seen in the ARM data. Note that Mace and Benson-Troth solved for $\mathcal{L}_{cf}^*$
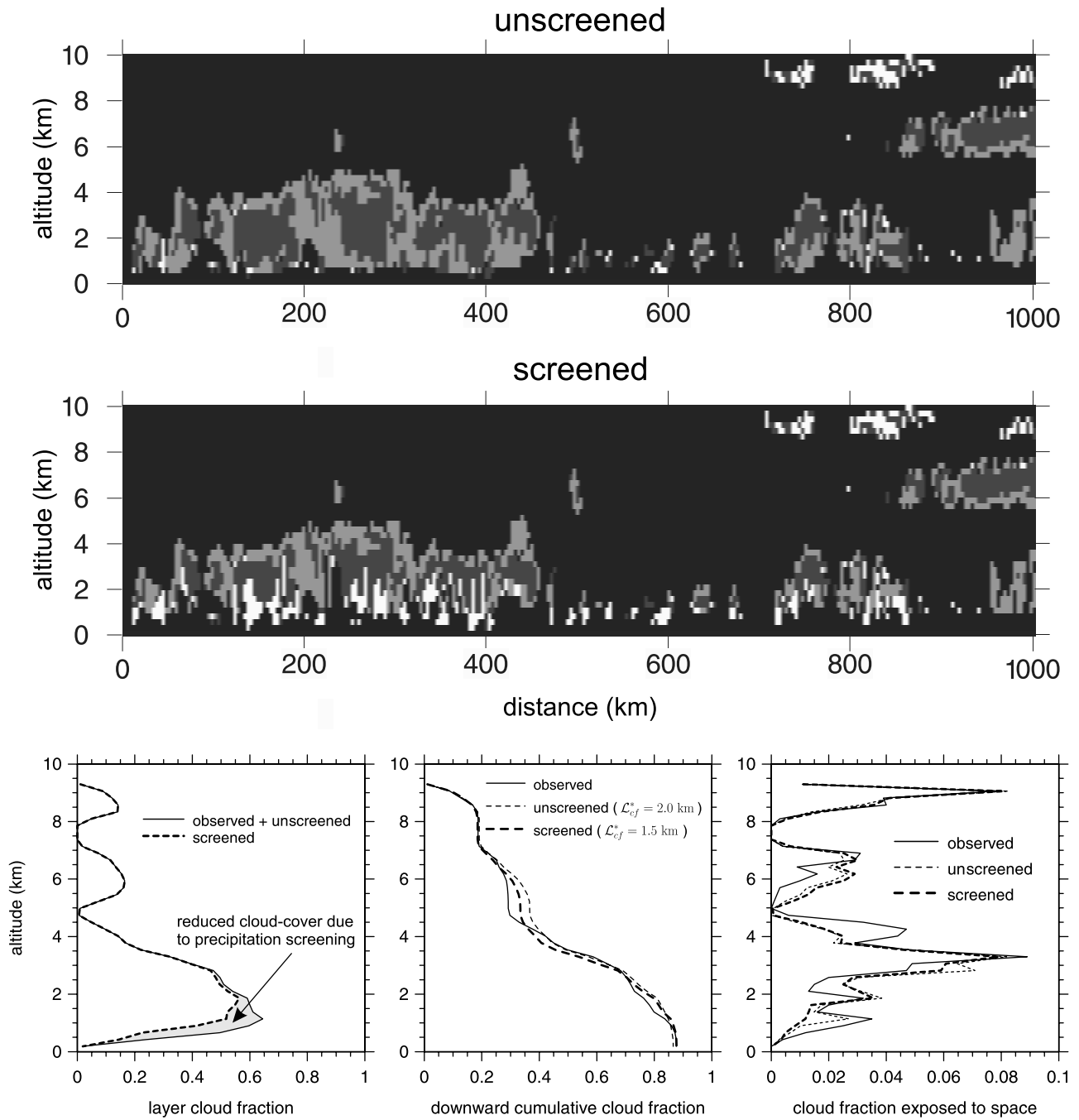
Figure 2. (top) A 1000 km long cloud mask cross section produced by a merger of CloudSat and CALIPSO data. The four shades are defined as follows: (1) white corresponds to radar volumes with 99% or more of their lidar volumes identified as containing cloud though the radar did not register cloud; (2) light gray represents radar volumes in which only the radar registered cloud; (3) dark gray is when both conditions 1 and 2 are satisfied; and (4) black corresponds to no cloud detected. This information is straight from the CloudSat database. (middle) The same as the upper except the precipitation mask was applied (see (5)). Plots along the bottom show profiles pertaining to the two cross sections. Leftmost plot shows layer cloud fractions; shaded area indicates the reduction in cloud cover due to application of (5). Center plot shows downward cumulative cloud fractions for the cross sections, as well as for fields produced by the stochastic cloud generator using $\mathcal{L}_{cf}^* = 2.0$ km and $\mathcal{L}_{cf}^* = 1.5$ km, which are applicable to the unscreened and screened cases, respectively.
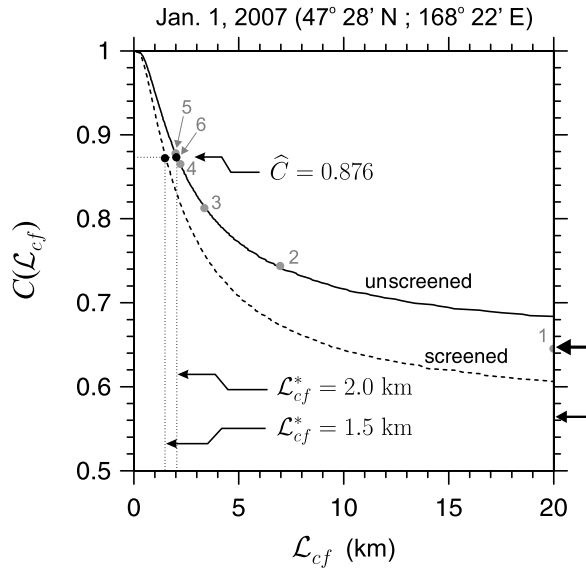
Jan. 1, 2007 (47° 28' N ; 168° 22' E)

Figure 3. Examples of total cloud fraction C as a function of $\mathcal{L}_{cf}$ for the cross sections shown in Figure 2. Bounds for the functions are random overlap, where $\mathcal{L}_{cf} = 0$ and $C \simeq 1$, and maximum overlap at $\mathcal{L}_{cf} \rightarrow \infty$; the latter values are indicated by arrows on the right. Values for $\mathcal{L}_{cf}^*$ are solutions to (8) using $\widehat{C} = 0.876$. The shaded dots and numbers correspond to values at successive steps in Brent's method of finding the root to $C(\mathcal{L}_{cf}^*) = 0.876$.

(which they called Δz_0) via a method that differed from that used here and they used rather short cross-sections. Their method requires much more computation than (8), and unlike $\mathcal{L}_{cf}^*$, there is no guarantee that use of their Δz_0 in a stochastic generator will return an unbiased estimate of $\widehat{C}$. Regardless, Mace and Benson-Troth's *mean* values of Δz_0 for the tropical ARM sites were ~5 km for relatively short cross sections and vertical layering close to that of the satellite data used here. Tropical *median* values shown in

Figure 9 for $L = 50$ or 100 km are only ~2.5 km. Note, however that the tail is thickening at these small L (see Figure 6) and corresponding tropical *means* are ~6 km. Hence, the ARM and satellite results seem to be in fair agreement.

[28] Figure 10 shows the same data as Figure 9 except they have been partitioned by $\widehat{C}$ and grouped into broader zonal bands. It shows that the large seasonal differences seen in Figure 9 for the regions 90°S–60°S and 60°N–90°N come from cases with $\widehat{C} \approx 0.5$. By the magnitudes of $\mathcal{L}_{cf}^*$ across to the rest of Earth, this suggests that there is greater, and quite substantial, vertical coherence during the polar winter. Presumably, this is due to ice crystals forming aloft and sedimenting through air with little shear [cf. *Hogan and Illingworth*, 2003]. Likewise, *Mace and Benson-Troth*'s [2002] estimate of ~1 km for Δz_0 from data gathered at the ARM North Slope of Alaska site agree nicely with satellite mean values for 70°N–80°N; particularly after precipitation screening which reduced median values shown in Figure 9 by ~0.5 km.

[29] As mentioned, $\widehat{C} = 1$ can occur in conjunction with $\max(c_k) < 1$. When this happens, the algorithm defaults to $\mathcal{L}_{cf}^* = 0$. However, when $\max(c_k) = 1$, there can be numerous other layers with $c_k < 1$ that overlap according to $\mathcal{L}_{cf}^* > 0$. These can be investigated easily by simply removing clouds from overcast layers and proceeding to compute $\mathcal{L}_{cf}^*$. For $L = 50$ km these situations occur about 10% of the time; for the vast majority of cross sections of length $L \geq 100$ km, removal of overcast layers results in total cloud fractions greater than 0.98 where differences in $\mathcal{L}_{cf}^*$ have minor impacts. Upon removing overcast layers from cross sections of length $L = 50$ km, values of $\mathcal{L}_{cf}^*$ relative to the entire population were typically 1 to 2 km larger; an increase of 50% to 100% (see Figure 6). This is partly due to the fact that even after removal of overcast layers, total cloud fractions generally exceeded 0.8, and as Figures 5, 8, and 10 indicate, if cross sections with small $\widehat{C}$ are neglected, median $\mathcal{L}_{cf}^*$ will increase. While such an increase in $\mathcal{L}_{cf}^*$ is large, the fact remains that the cross

Jan. 1, 2007 (47° 28' N ; 168° 22' E)
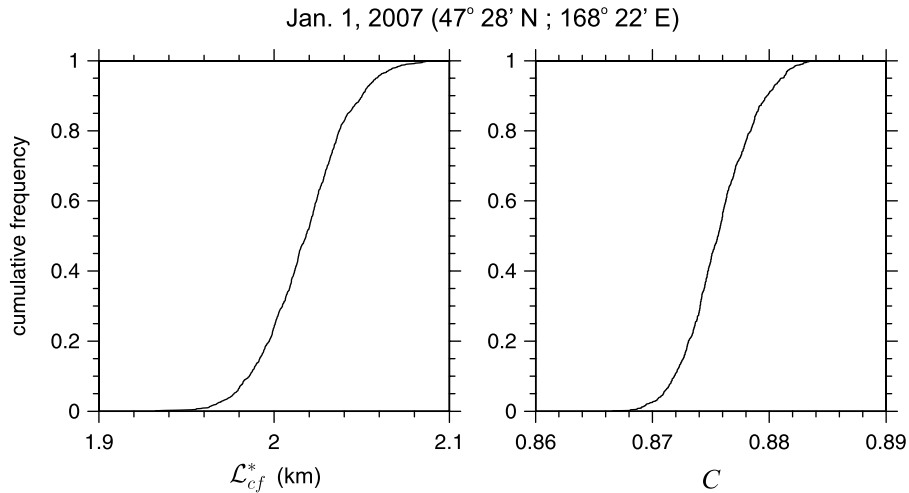

Figure 4. (left) Cumulative distribution of $\mathcal{L}_{cf}^*$ obtained via Brent's method for the unscreened curve shown in Figure 3. These are for convergence criterion of 0.1 km and 25,000 stochastic subcolumns in the cloud generator. Using this distribution of $\mathcal{L}_{cf}^*$ in the stochastic generator, with 100,000 subcolumns, yields (right) the cumulative distribution of total cloud fractions C.

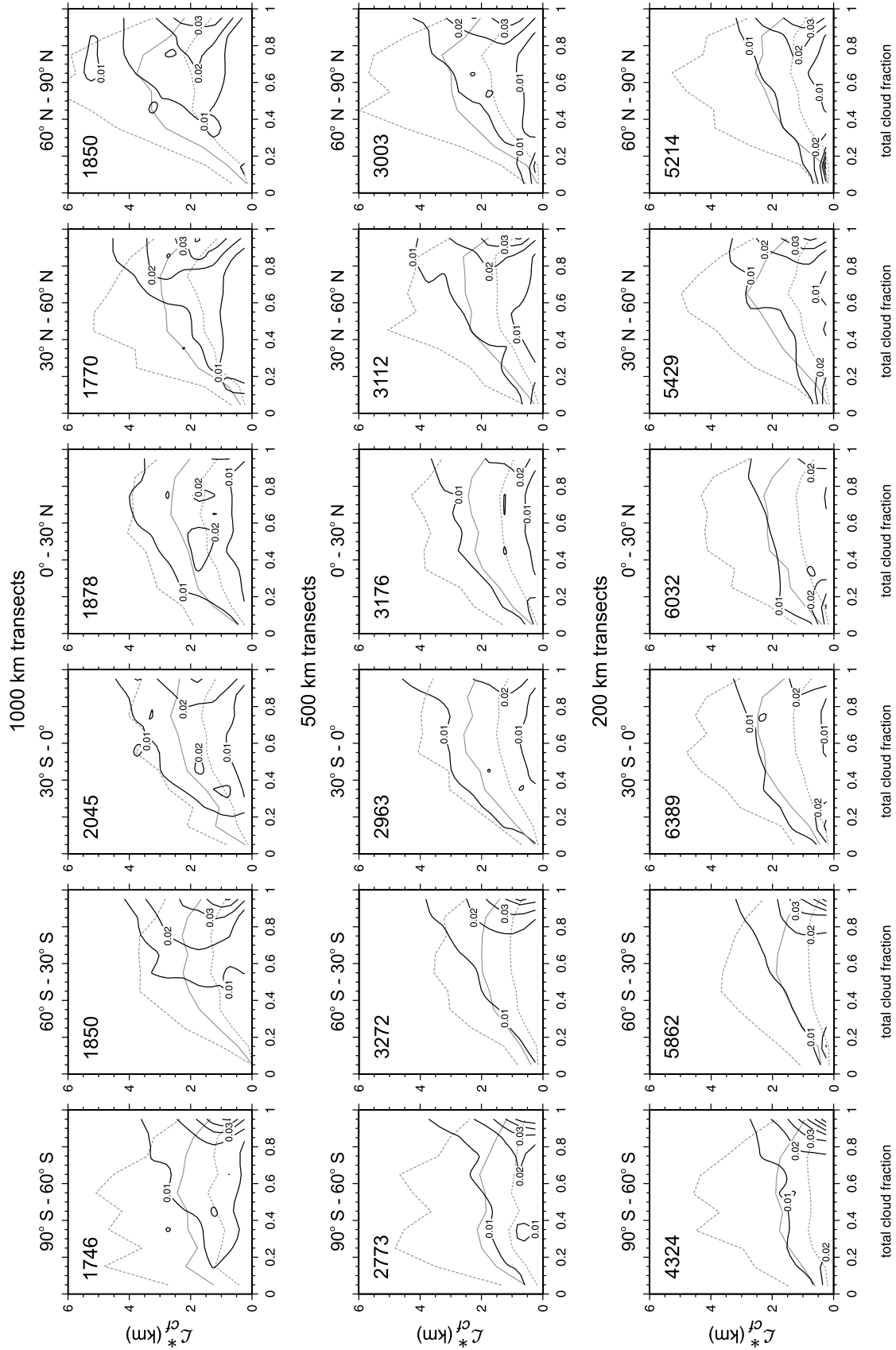**Figure 5.** Contours (black lines) indicate frequency of occurrence of $\mathcal{L}_{cf}^{*}$ for a given total cloud fraction $\widehat{C}$ for all unscreened satellite cross sections with $\widehat{C}$ between 0.05 and 0.99 for January 2007. Frequency distributions are integrals over zonal bands, as indicated in the title of each plot. Sample size is indicated in the upper left corner of each plot. Upper row applies to 1000 km long cross sections, while the middle and lower rows correspond to cross sections of lengths 500 km and 200 km, respectively. Shown on each plot is median $\mathcal{L}_{cf}^{*}$ (solid gray line) as well as corresponding 25th and 75th percentiles (dashed gray lines) as functions of $\widehat{C}$.

January 2007

median of $\mathcal{L}^*_{cf}$

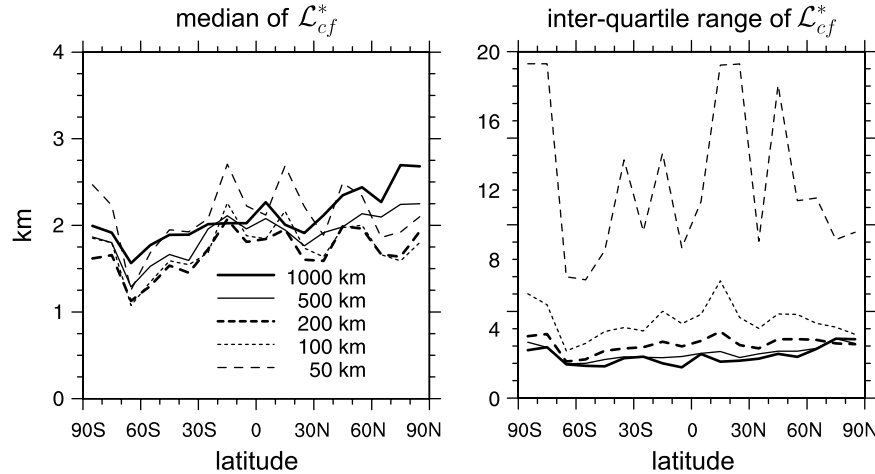**inter-quartile range of $\mathcal{L}^*_{cf}$**

Figure 6. (left) Median $\mathcal{L}^*_{cf}$ for January 2007 as a function of latitude for unscreened cross sections with lengths listed on the plot. (right) Corresponding interquartile ranges. Note that the largest allowable $\mathcal{L}^*_{cf}$ was 20 km.

section is overcast and realigning nonovercast layers via changes to $\mathcal{L}^*_{cf}$ does not alter $\widehat{C}$; it alters only the variance of cloud optical depth [see *Barker and Räisänen*, 2005].

[30] To round out these comparisons, cross sections were partitioned according to whether the Sun was up or down (according to *Solar_zenith* in CloudSat's MODIS-AUX files). This was motivated partly by the expectation of finding some cloud structural differences depending on whether the Sun is up or not [e.g., *Wang et al.*, 1999] but also CALIPSO's lidar signal is noisier during sunup (K. Strawbridge, personal communication, 2007) and this presumably impacts cloud detection and hence estimates of cloud fraction and $\mathcal{L}^*_{cf}$. The left plot in Figure 11 shows zonal median $\mathcal{L}^*_{cf}$ for 500 km cross sections for January 2007. During sundown, estimates of $\mathcal{L}^*_{cf}$ are typically ~0.75 km larger than during sunup. These are sufficiently large differences that might warrant being captured in parametrizations of $\mathcal{L}^*_{cf}$. Smaller values during sunup are, however, consistent with the hypothesis that additional noise in lidar data fosters minor random fluctuations in cloud identification thereby affecting a shift in $\mathcal{L}^*_{cf}$ to smaller values.

[31] The right plot in Figure 11 shows the same analysis as in Figure 10 except just CloudSat radar data were used (i.e., *CPR_Cloud_mask*). It is immediately clear that when lidar data are neglected the day/night distinction in $\mathcal{L}^*_{cf}$ disappears. This too supports the claim that $\mathcal{L}^*_{cf}$ are biased low by random fluctuations in identification of cloud by the lidar during sunup. Moreover, note that radar-only estimates of $\mathcal{L}^*_{cf}$ are typically 0.5 to 1 km larger than sundown estimates using radar and lidar. This is likely because the lidar's primary contribution to cloud detection is thin high cloud that is often missed by the radar which, as mentioned early, tends to detect lower precipitating cloud best. Likewise, the radar can sometimes miss low nonprecipitating cloud that could be detected by the lidar given a fairly clear line of sight. Since thin high clouds and low clouds are often essentially decoupled from each other, inclusion of the lidar will lead, rightly so, to smaller $\mathcal{L}^*_{cf}$.

5.2. Discussion: On the Prospect of Using Parametrized $\mathcal{L}^*_{cf}$ in GCMs

[32] Given the importance of $\widehat{C}$ for, at least, radiation calculations in GCMs, it is recognized generally that a means for describing how clouds overlap and thus give rise to $\widehat{C}$, is required. As mentioned earlier, while $\mathcal{L}_{cf}$ will almost always vary with height through a column [cf. *Hogan and Illingworth*, 2003; *Räisänen et al.*, 2004], there is a certain attraction to using something as simple as the effective overlap decorrelation length $\mathcal{L}^*_{cf}$. The questions that have to be answered are can it be parametrized well and will its simplicity result in undesirable effects that might be avoided with a, justifiably, more elaborate parametrization?

[33] Consider first some of the limitations associated with the use of $\mathcal{L}^*_{cf}$. *Räisänen et al.*'s [2004] subcolumn cloud generator will always produce, on average, the correct
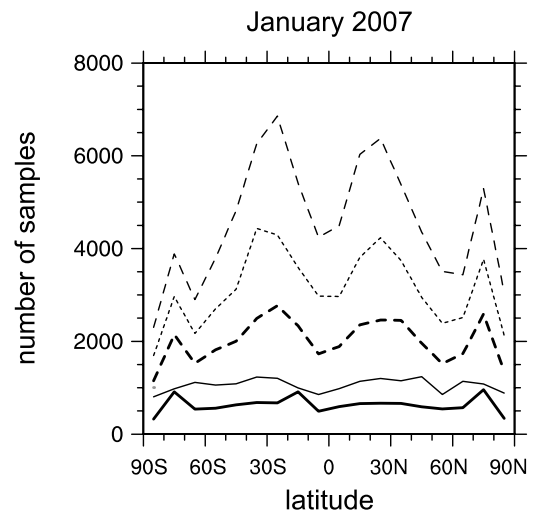
January 2007

Figure 7. Number of samples per 10° wide latitude band used to produce results shown in Figure 6. Cross-section lengths are indicated on Figure 6.
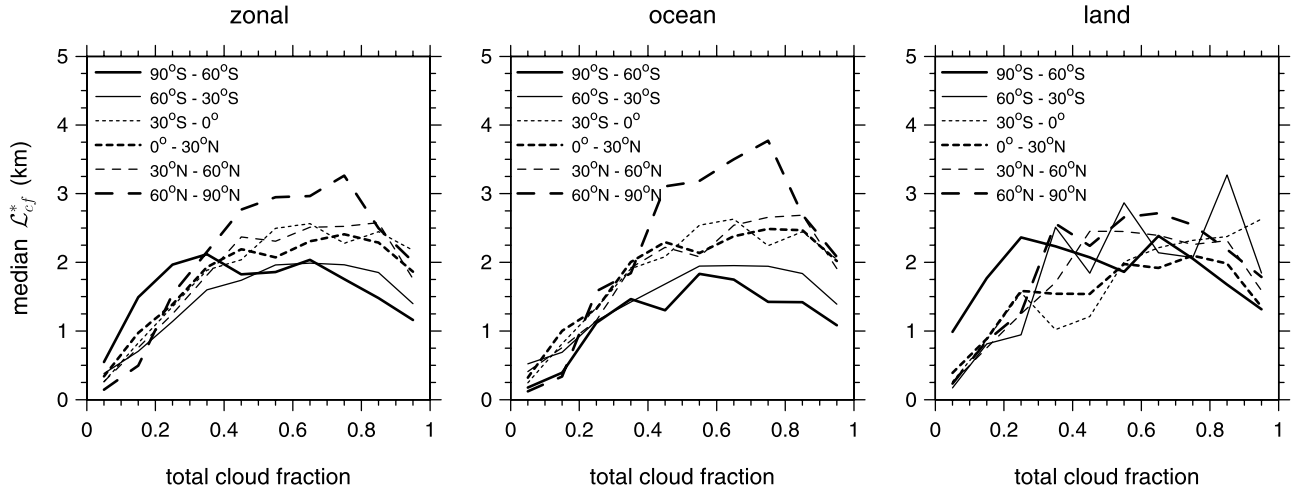
January 2007

Figure 8. Median $\mathcal{L}_{cf}^*$ for 500 km long cross sections for January 2007 as a function of total cloud fraction for unscreened cross sections for six zonal bands as indicated on the plots. Leftmost plot is for all cross sections in a latitude band while the other plots are for ocean and land areas.

profile of layer cloud fractions. Likewise, if supplied with the *correct* $\mathcal{L}_{cf}^*$, it will also produce a distribution, due to random sampling, of total cloud fractions $p(C)$. It will usually return, however, an erroneous distribution of cloud-top areas exposed to space. This could be important for computation of longwave radiative heating rates and outgoing LW to space.

[34] Figure 12 shows zonal average layer cloud fraction profiles for six latitude bands for 500 km cross sections observed during January 2007. By definition, the expectation profiles, in the statistical sense, produced by the stochastic cloud generator are identical to those shown in Figure 12. Figure 13 shows corresponding zonal mean profiles of fractional amount of cloud exposed to space for the observed data and the generated fields using proper $\mathcal{L}_{cf}^*$. The right and left *error bars* on the plots are defined as

$$\frac{\sum\limits_{n=1}^{N}\varepsilon_n^{\mathrm{obs}}}{N} - \frac{\sum\limits_{\varepsilon_n^{\mathrm{mod}}<\varepsilon_n^{\mathrm{obs}}}^{N_<}\left(\varepsilon_n^{\mathrm{mod}}-\varepsilon_n^{\mathrm{obs}}\right)}{N_<} \; ; \; \frac{\sum\limits_{n=1}^{N}\varepsilon_n^{\mathrm{obs}}}{N} + \frac{\sum\limits_{\varepsilon_n^{\mathrm{mod}}\geq\varepsilon_n^{\mathrm{obs}}}^{N_\geq}\left(\varepsilon_n^{\mathrm{mod}}-\varepsilon_n^{\mathrm{obs}}\right)}{N_\geq}$$

(9)

where $\varepsilon_n^{\mathrm{obs}}$ and $\varepsilon_n^{\mathrm{mod}}$ are observed and modelled fractions exposed to space, N is the total number of samples, and $N_<$ and $N_\geq$ are the number of samples in which $\varepsilon_n^{\mathrm{mod}} < \varepsilon_n^{\mathrm{obs}}$ and $\varepsilon_n^{\mathrm{mod}} \geq \varepsilon_n^{\mathrm{obs}}$, respectively. Hence, the bars represent mean negative and positive deviations from the observed means. For the most part, mean bias errors are very small but there are some obvious serial correlations: at altitudes above $\sim$3 km the generated fields have too much cloud exposed to space, while at lower altitudes they have too little. Just as with the example in Figure 2, this indicates that values of $\mathcal{L}_{cf}$ are smaller than $\mathcal{L}_{cf}^*$ near the surface and larger aloft. This was true also for MMF data analyzed by *Räisänen et al.* [2004, Figure 3]. The small bias errors shown here suggest that if provided with good estimates of $\mathcal{L}_{cf}^*$, generated profiles should be satisfactory.

[35] With such small errors associated with mean cloud fraction exposed to space, one might expect that errors in $\mathcal{L}_{cf}^*$ will matter little. However, consider the following test in which constant values of $\mathcal{L}_{cf}^*$ are applied to the entire set of CloudSat profiles. On the basis of Figures 6, 9, and 11, if one had to select a global-constant $\mathcal{L}_{cf}^*$ for use in a GCM, it would likely be between 1 and 3 km. Figure 14 shows differences between stochastically generated C based on $\mathcal{L}_{cf}^* = 1$, 2, and 3 km and their respective $\widehat{C}$ as a function of $\widehat{C}$ for 3000 randomly sampled 500 km long cross sections. Clearly, $\mathcal{L}_{cf}^* = 1$ km is too small as many estimates of $\widehat{C}$ are too large. For $\mathcal{L}_{cf}^* = 3$ km the reverse is true. While $\mathcal{L}_{cf}^* = 2$ km yields a small bias error, a large fraction of errors exceed 10% of $\widehat{C}$. Stochastic noise was not an issue here as 100,000 subcolumns were used to generate C. It is now
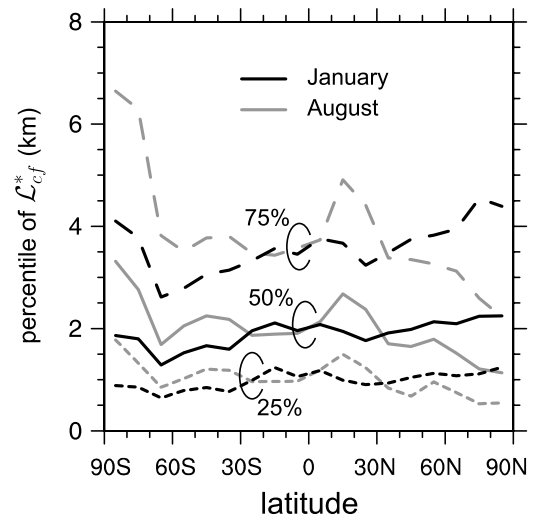

Figure 9. The 25th, 50th, and 75th percentiles of $\mathcal{L}_{cf}^*$ for 500 km long unscreened cross sections as functions of latitude for January and August 2007.
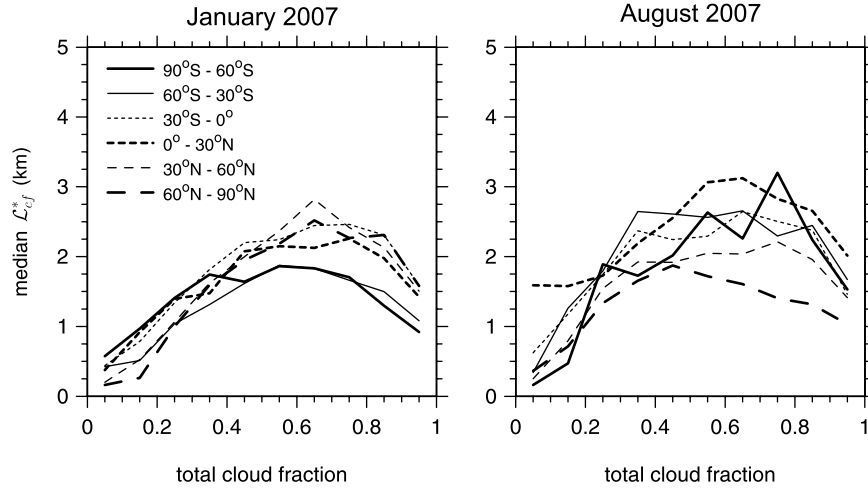
Figure 10. Median $\mathcal{L}_{cf}^*$ for 500 km long unscreened cross sections for January and August 2007 as functions of total cloud fraction for six zonal bands as indicated on the plots.

known, however, through experiments with McICA (H. W. Barker et al., The Monte Carlo Independent Column Approximation: An assessment using several global atmospheric models, submitted to *Quarterly Journal of the Royal Meteorological Society*, 2008), that climate simulations can digest very large amounts of unbiased noise, and so errors like those shown in Figure 14 may be more acceptable than they appear. Another study is needed to elucidate the radiative implication of these errors, and their impact on GCM simulations, before simple *parametrizations* of $\mathcal{L}_{cf}^*$ can be dismissed.

5.3. Discussion: Importance of Cloud Overlap for Radiation Calculations in GCMs

[36] *Barker and Räisänen* [2005] addressed the question of radiative sensitivities due to cloud overlap decorrelation length by assessing $\partial F_{\text{ICA}}/\partial \mathcal{L}_{cf}$, where F_{ICA} is net solar flux at the TOA, using the concept of $\mathcal{L}_{cf}^*$ and performing extensive radiative transfer calculations on data from a

global array of CSRMs [*Khairoutdinov and Randall*, 2001]. It is possible, however, to get a global estimate of $\partial F_{\text{ICA}}/\partial \mathcal{L}_{cf}$ using only the data presented thus far; that is, without having to perform radiative transfer simulations and deal with gaps in CloudSat water content and particle size retrievals.

[37] Differentiating (2) with respect to $\mathcal{L}_{cf}$, and assuming that $\partial F(0)/\partial \mathcal{L}_{cf} = 0$, gives

$$\frac{\partial F_{\text{ICA}}}{\partial \mathcal{L}_{cf}} = -F(0)\frac{\partial \widehat{C}}{\partial \mathcal{L}_{cf}} + F_{\text{cld}}\frac{\partial \widehat{C}}{\partial \mathcal{L}_{cf}} + \widehat{C}\frac{\partial F_{\text{cld}}}{\partial \mathcal{L}_{cf}} \qquad (10)$$

where

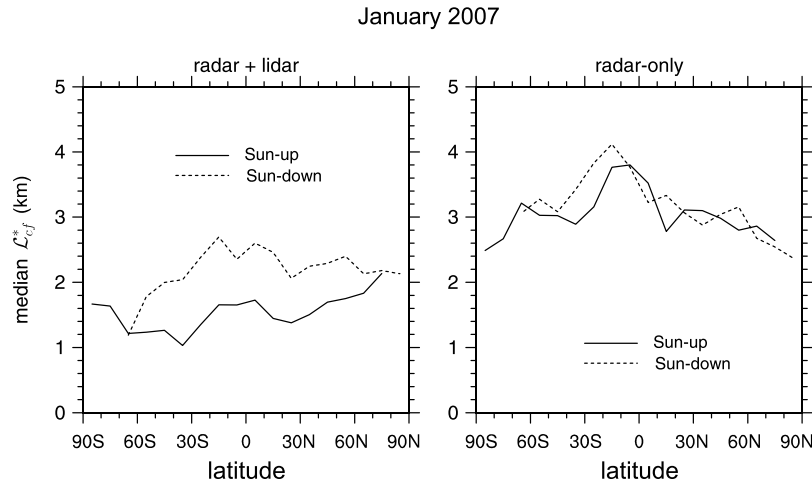$$F_{\text{cld}} = \int_0^\infty \widehat{p}(\tau)F(\tau)\,d\tau.$$

Figure 11. Median $\mathcal{L}_{cf}^*$ for January 2007 as a function of latitude conditional upon whether the Sun is irradiating or not cross sections of length 500 km. The left plot is for cloud masks inferred from radar and lidar data while the right is for radar only.
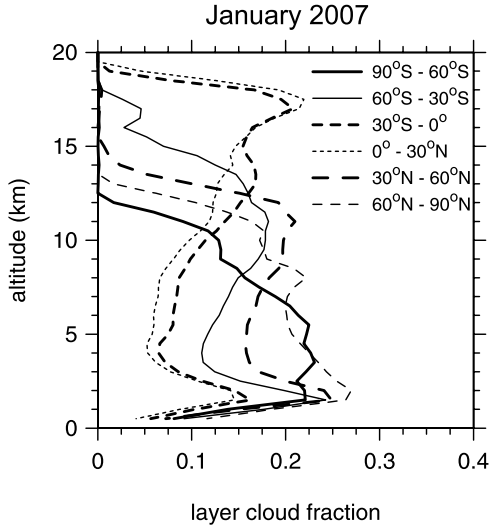
Figure 12. Zonal average layer cloud fraction profiles as functions of altitude for unscreened CloudSat-CALIPSO cloud mask data for January 2007.

Since changing cloud overlap changes distributions of cloud optical depth $\widehat{p}(\tau)$ [*Barker et al.*, 1999; *Barker and Räisänen*, 2005], the last term in (10) can be expanded as

$$\frac{\partial F_{\text{cld}}}{\partial \mathcal{L}_{cf}} \simeq \frac{\partial F_{\text{cld}}}{\partial \overline{\tau}} \frac{\partial \overline{\tau}}{\partial \mathcal{L}_{cf}} + \frac{\partial F_{\text{cld}}}{\partial \sigma_{\tau}} \frac{\partial \sigma_{\tau}}{\partial \mathcal{L}_{cf}} \tag{11}$$

where $\overline{\tau}$ is mean cloud optical depth, σ_{τ} is standard deviation of $\widehat{p}(\tau)$, and changes to higher-order moments of $\widehat{p}(\tau)$ have been neglected [cf. *Barker and Davis*, 2005]. Using the common definition for TOA shortwave (SW) cloud radiative effect of

$$CRE \equiv F_{\text{ICA}} - F(0) = \widehat{C}[F_{\text{cld}} - F(0)]$$

and substituting (11) into (10) gives

$$\frac{\partial F_{\text{ICA}}}{\partial \mathcal{L}_{cf}} \simeq CRE \frac{1}{\widehat{C}} \frac{\partial \widehat{C}}{\partial \mathcal{L}_{cf}} + \widehat{C} \left[\underbrace{\frac{\partial F_{\text{cld}}}{\partial \overline{\tau}} \frac{\partial \overline{\tau}}{\partial \mathcal{L}_{cf}}}_{<0} + \underbrace{\frac{\partial F_{\text{cld}}}{\partial \sigma_{\tau}} \frac{\partial \sigma_{\tau}}{\partial \mathcal{L}_{cf}}}_{\sim O(0)} \right] \tag{12}$$
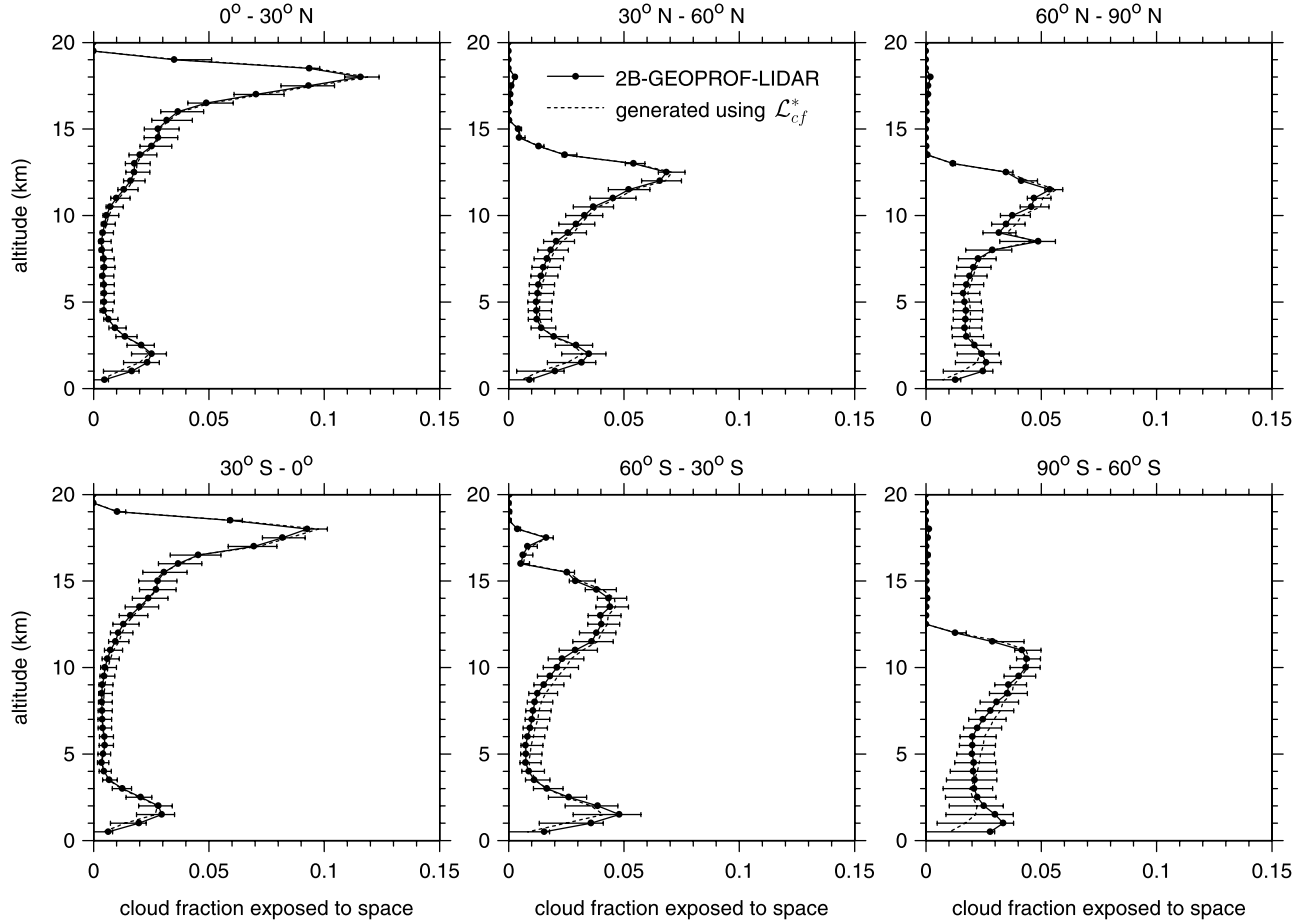

Figure 13. Zonal mean profiles of cloud fractions exposed to space (see (3)) as functions of altitude for six latitude bands. Values for unscreened CloudSat-CALIPSO cloud-mask data for January 2007 are indicated by the dots. Dashed lines are corresponding means that result from using $\mathcal{L}_{cf}^{*}$ in the stochastic cloud generator. "Error" bars indicate mean negative and positive deviations of generated values from observed values as defined in (9).
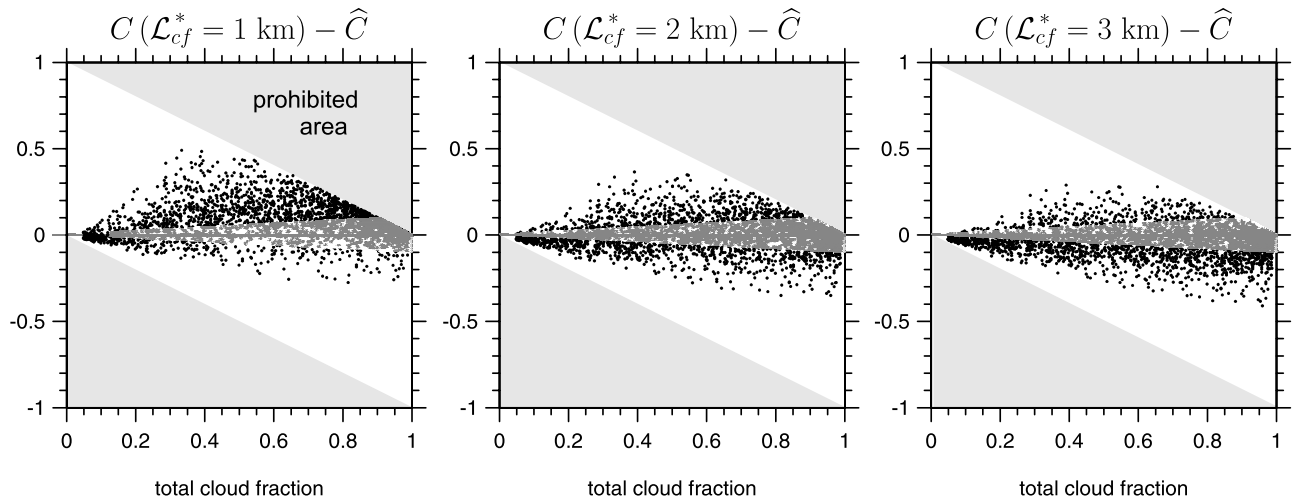
Figure 14. Deviations of stochastically generated total cloud fractions C based on 100,000 subcolumns, observed cloud fraction profiles, and use of either $\mathcal{L}_{cf}^* = 1, 2$, or 3 km (as indicated in the titles) plotted against corresponding true cloud fractions $\widehat{C}$ for 3000 unscreened cross-sections of length 500 km for January 2007. Shaded regions are prohibited areas for one can only be in error so much. Shaded transparent wedges indicate errors smaller than 10%.

As noted in (12), the first term in brackets is negative since net TOA flux decreases as mean τ increases, and mean τ increases with $\mathcal{L}_{cf}$ for clouds become increasingly overlapped yet mass is conserved. On the basis of *Barker and Räisänen*'s [2005] analyses, this term is likely close to, but less than, ~ 1 W m^{-2} km^{-1}. The second term in brackets is likely very small as $\partial\sigma_\tau/\partial\mathcal{L}_{cf}$ can be positive or negative [*Barker and Räisänen*, 2005]. Integrating (12) over annual cycle and Earth, yields

$$\left\langle \frac{\partial F_{\mathrm{ICA}}}{\partial\mathcal{L}_{cf}} \right\rangle \simeq \langle CRE \rangle * \left\langle \frac{\partial \ln \widehat{C}}{\partial\mathcal{L}_{cf}} \right\rangle - O(\lesssim 1) \qquad (13)$$

where angular brackets indicate time and space averages. Using -45 W m^{-2} (based on CERES satellite data for 2001–2004) (J. Cole, personal communication, 2007) for global mean, TOA SW CRE and -0.08 km^{-1} for the median of $\partial\ln\widehat{C}/\partial\mathcal{L}_{cf}$ for the unscreened, diurnal integrated, 1000 km cross sections for January 2007 gives a maximum sensitivity of ~ 3.5 W m^{-2} km^{-1} which is close to *Barker and Räisänen*'s [2005] lengthy model-based estimate of ~ 2 W m^{-2} km^{-1}. Corresponding sensitivity for longwave radiation is expected to be smaller than 1 W m^{-2} km^{-1} [see *Barker and Räisänen*, 2005].

6. Conclusions and Recommendations

[38] At this stage in the development of GCMs, the need to consider cloud overlap structure depends on the genre of the GCM. For conventional GCMs, a description of cloud overlap for unresolved cloud fields must be provided, via parametrization, to carry out, among other things, radiative transfer calculations. In the more avant garde multiscale GCMs, with imbedded CSRMs [*Randall et al.*, 2003], cloud overlap ceases to be a parametrization issue and becomes a diagnostic variable that can help to assess CSRM clouds.

[39] This study examined the vertical overlapping properties of fractional amounts of cloud on a global scale using two months worth of "cloud mask" products derived from CloudSat and CALIPSO data [*Mace*, 2007]. Cloud overlap was diagnosed assuming that total cloud cover C can be described as a linear combination of maximum and random overlap of layer clouds [*Hogan and Illingworth*, 2000]. The weighting factor was assumed to depend on layer separation and a decorrelation length $\mathcal{L}_{cf}$. If $\mathcal{L}_{cf}$ is forced to be constant with height, then for a given profile of layer cloud amounts there exists a unique function $C(\mathcal{L}_{cf})$. By computing the actual total cloud fraction $\widehat{C}$ for a cross section of satellite data, one can solve $C(\mathcal{L}_{cf}^*) = \widehat{C}$ where $\mathcal{L}_{cf}^*$ is effective decorrelation length. $\mathcal{L}_{cf}^*$ was used here as a diagnostic variable for an otherwise cumbersome characteristic to assess.

[40] In very general terms, median values of $\mathcal{L}_{cf}^*$ tend to 0 km for very small $\widehat{C}$, increase approximately linearly with $\widehat{C}$ to maxima between 2 and 3 km near $\widehat{C} \simeq 0.7$, and decrease to ~ 1.5 km as $\widehat{C} \to 1$. Not only does this appear to be a fair description regardless of location and time, it also applies for cross-section lengths L ranging from 100 km to 1000 km. There are, however, some interesting, and potentially important, spatial and temporal variations in the statistics of $\mathcal{L}_{cf}^*$. For instance, the largest $\mathcal{L}_{cf}^*$ appear to occur in polar regions during their respective winters and in the northern tropics during boreal summer. Although $\mathcal{L}_{cf}^*$ were computed differently than by *Mace and Benson-Troth* [2002] and came from a very different data set, their general agreement is encouraging.

[41] A concern with CloudSat is that precipitation, whose visible extinction is generally much smaller than that of cloud and not acknowledged in GCMs, figures too strongly

in its cloud masks. The issue is that precipitation is often correlated well in the vertical and so might bias $\mathcal{L}_{cf}^*$ high, specially if one is assessing $\mathcal{L}_{cf}^*$ for the express purpose of defining cloud structure for radiation calculations in GCMs. Hence, a very rough precipitation-screening algorithm was devised and applied to CloudSat-CALIPSO cloud masks. The impact was to reduce values of $\mathcal{L}_{cf}^*$ from typically $\sim$2 km to $\sim$1.5 km.

[42] Synergy between radar and lidar was essential for the analyses presented here. As shown, when cloud masks were based on just radar data, typical values of $\mathcal{L}_{cf}^*$ increased by $\sim$1 km. This is clearly due to the radar having missed high tenuous clouds which are decoupled from lower clouds and presumably too much weight given to large (precipitating) particles that are readily detected by CloudSat's radar and are correlated in the vertical better than cloud. At the same time, however, when radar and lidar data were used, it appeared that during sunup periods random noise in lidar data, and hence cloud detection, yielded systematic underestimates of $\mathcal{L}_{cf}^*$. When just radar data were used, day/night differences in $\mathcal{L}_{cf}^*$ vanished.

[43] It was shown that a global estimate of sensitivity of TOA solar fluxes to changes in $\mathcal{L}_{cf}^*$ can be estimated without performing radiative transfer calculations. A systematic error of $\sim$1 km in $\mathcal{L}_{cf}^*$ can be expected to alter global-average net solar flux at the TOA by $\sim$3 W m^{-2}. This estimate, which is comparable to that expected for a 10% to 20% systematic change in cloud particle size, agrees with *Barker and Räisänen*'s [2005] estimate derived from extensive model-based calculations.

[44] Although $\mathcal{L}_{cf}^*$ was used here essentially as a diagnostic tool, it may yet find meritable use in GCMs. It remains to be shown whether it can be parametrized based on resolved variables and whether such a parametrization, preferably as simple as possible, can capture a satisfactory fraction of radiative flux variance. The extremely simple *parametrization* of setting $\mathcal{L}_{cf}^*$ to a global constant is attractive but has to be tested. It might be that having two distinct values, one for stable (stratiform) and another for unstable (cumuliform) conditions, might effectively velcro overlap to the cloud feedback conundrum, not as a feedback in itself but rather as a conditional setting that modulates the GCM's radiation budget in accordance with changing frequencies of occurrence of stratiform and cumuliform cloud.

[45] On this note, it was shown that use of a global-constant $\mathcal{L}_{cf}^*$ leads to substantial random noise in C which would, of course, be passed on to radiation fields. A decade ago this might have seemed unacceptable, but experiments with the McICA method (H. W. Barker et al., submitted manuscript, 2008) have shown that GCMs can withstand much random noise; at the expense of quashing bias errors. It is worth mentioning that a similar brand of random noise is implicitly present in more idealized schemes such as maximum random overlap.

[46] This study represents just an initial step in the use of A-train data to study cloud structural properties. There are several obvious subsequent steps including linking $\mathcal{L}_{cf}^*$ to resolved variables in GCMs, assessment of cloud condensate overlap, further investigation of diurnal variations in $\mathcal{L}_{cf}^*$, assessment of overlap in cloud system-resolving models (including superparametrized GCMs), and ultimately impacts on radiation budgets and GCM simulations. Finally,

conventional GCM sensitivity studies should be conducted to help elucidate how much work is warranted in the construction of a cloud overlap parametrization. In so doing, however, attention has to be paid to the verisimilitude of the GCM's cloudiness, for if too incorrect, it can easily skew the role of overlap.

References

Astin, I., and L. Di Girolamo (1999), A general formalism for the distribution of the total length of a geophysical parameter along a finite transect, *IEEE Trans. Geosci. Remote Sens.*, 37, 508–512.

Astin, I., and L. Di Girolamo (2006), The relationship between a and the cross-correlation of cloud fraction, *Q. J. R. Meteor. Soc.*, 132, 2475–2478.

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

Barker, H. W., and A. B. Davis (2005), Approximation methods in atmospheric 3D radiative transfer, part 2: Unresolved variability and climate applications, in *3D Radiative Transfer in Cloudy Atmospheres*, edited by A. Marshak and A. B. Davis, 686 pp., Springer, Heidelberg, Germany.

Barker, H. W., and P. Räisänen (2005), Radiative sensitivities for cloud structural properties that are unresolved by conventional GCMs, *Q. J. R. Meteorol. Soc.*, 131, 3103–3122.

Barker, H. W., G. L. Stephens, and Q. Fu (1999), The sensitivity of domain-averaged solar fluxes to assumptions about cloud geometry, *Q. J. R. Meteorol. Soc.*, 125, 2127–2152.

Barker, H. W., R. Pincus, and J.-J. Morcrette (2002), The Monte Carlo Independent Column Approximation: Application within large-scale models, paper presented at GCSS-ARM Workshop on the Representation of Cloud Systems in Large-Scale Models, GEWEX Cloud Syst. Sci., Kananaskis, Alberta, Canada. (Available at http://www.met.utah.edu/skrueger/gcss-2002/Extended-Abstracts.pdf.)

Barker, H. W., et al. (2003), Assessing 1D atmospheric solar radiative transfer models: Interpretation and handling of unresolved clouds, *J. Clim.*, 16, 2676–2699.

Bergman, J. W., and P. J. Rasch (2002), Parameterizing vertically coherent cloud distributions, *J. Atmos. Sci.*, 59, 2165–2182.

Brent, R. P. (1973), *Algorithms for Minimization Without Derivatives*, chap. 4. Prentice-Hall, Englewood Cliffs, N. J.

Doviak, R. J., and D. S. Zrnić (1993), *Doppler Radar and Weather Observations*, 562 pp., Academic, San Diego, Calif.

Geleyn, J.-F., and Hollingsworth (1979), An economical analytical method for the computation of the interaction between scattering and line absorption of radiation, *Cont. Atmos. Phys.*, 52, 1–16.

Hogan, R. J., and A. J. Illingworth (2000), Deriving cloud overlap statistics from radar, *Q. J. R. Meteor. Soc.*, 128, 2903–2909.

Hogan, R. J., and A. J. Illingworth (2003), Parameterizing ice cloud inhomogeneity and the overlap of inhomogeneities using cloud radar data, *J. Atmos. Sci.*, 60, 756–767.

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.

Mace, G. G. (2006), CloudSat project: Level 2 GEOPROF product process description and interface control document algorithm version 5.3, Coop. Inst. for Res. in the Atmos., Colo. State Univ., Fort Collins. (Available at http://cloudsat.cira.colostate.edu/dataICDlist.php?go=list&path=/2B-GEOPROF)

Mace, G. G. (2007), CloudSat project: Level 2 radar-lidar GEOPROF product VERSION 1.0 process description and interface control document, Coop. Inst. for Res. in the Atmos., Colo. State Univ., Fort Collins. (Available at http://cloudsat.cira.colostate.edu/dataICDlist.php?go=list&path=/2B-GEOPROF-LIDAR)

Mace, G. G., and S. Benson-Troth (2002), Cloud-layer overlap characteristics derived from long-term cloud radar data, *J. Clim.*, 15, 2505–2515.

Manabe, S., and R. F. Strickler (1964), Thermal equilibrium of the atmosphere with a convective adjustment, *J. Atmos. Sci.*, 21, 361–385.

Morcrette, J.-J., and Y. Fouquart (1986), The overlapping of cloud layers in shortwave radiation parameterizations, *J. Atmos. Sci.*, 43, 321–328.

Morcrette, J.-J., H. W. Barker, J. N. S. Cole, M. J. Iacono, and R. Pincus (2008), Impact of the new radiation package, McRAD, in the ECMWF integrated forecast system, *Mon. Weather Rev.*, doi:10.1175/2008MWR2363.1, in press.

Oreopoulos, L., and M. F. Khairoutdinov (2003), Overlap properties of clouds generated by a cloud-resolving model, *J. Geophys. Res.*, *108*(D15), 4479, doi:10.1029/2002JD003329.

Pincus, R., H. W. Barker, and J.-J. Morcrette (2003), A fast, flexible, approximate technique of computing radiative transfer for inhomogeneous clouds, *J. Geophys. Res.*, *108*(D13), 4376, doi:10.1029/2002JD003322.

Pincus, R., C. Hannay, S. A. Klein, K.-M. Xu, and R. Hemler (2005), Overlap assumptions for assumed probability distribution function cloud schemes in large-scale models, *J. Geophys. Res.*, *110*, D15S09, doi:10.1029/2004JD005100.

Räisänen, P., H. W. Barker, M. Khairoutdinov, J. Li, and D. A. Randall (2004), Stochastic generation of subgrid-scale cloudy columns for large-scale models, *Q. J. R. Meteorol. Soc.*, *130*, 2047–2067.

Räisänen, P., H. W. Barker, and J. Cole (2005), The Monte Carlo independent column approximation's conditional random noise: Impact on simulated climate, *J. Clim.*, *17*, 4715–4730.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski (2003), Breaking the cloud parameterization deadlock, *Bull. Am. Meteorol. Soc.*, *84*, 1547–1564.

Rossow, W. B. (1989), Measuring cloud properties from space: A review, *J. Clim.*, *2*, 201–213.

Stephens, G. L. (1988), Radiative transfer through arbitrary shaped optical media: Part II: Group theory and simple closures, *J. Atmos. Sci.*, *45*, 1837–1848.

Stephens, G. L., P. M. Gabriel, and S.-C. Tsay (1991), Statistical radiative transfer in one-dimensional media and its application to the terrestrial atmosphere, *Trans. Theory Stat. Phys.*, *20*, 139–175.

Stephens, G. L., et al. (2002), The CloudSat Mission and the A-Train, *Bull. Am. Meteorol. Soc.*, *83*, 1771–1790.

Stephens, G. L., N. B. Wood, and P. M. Gabriel (2004), An assessment of the parameterization of subgrid-scale cloud effects on radiative transfer. Part I: Vertical overlap, *J. Atmos. Sci.*, *61*, 715–732.

Tian, L., and J. A. Curry (1989), Cloud overlap statistics, *J. Geophys. Res.*, *94*, 9925–9935.

Wang, J., W. B. Rossow, T. Uttal, and M. Rozendaal (1999), Variability of cloud vertical structure during ASTEX observed from a combination of rawinsonde, radar, ceilometer, and satellite, *Mon. Weather Rev.*, *127*, 2484–2502.

Winker, D. M., J. Pelon, and M. P. McCormick (2003), The CALIPSO mission: Spaceborne lidar for observation of aerosols and clouds, *Proc. SPIE*, *4893*, 1–11.

———————————

H. W. Barker, Cloud Physics and Severe Weather Research Section, Environment Canada, 4905 Dufferin Street, Toronto, ON M3H 5T4, Canada. (howard.barker@ec.gc.ca)