Parameterizing the Difference in Cloud Fraction Defined by Area and by Volume as Observed with Radar and Lidar

MALCOLM E. BROOKS,* ROBIN J. HOGAN, AND ANTHONY J. ILLINGWORTH

Department of Meteorology, University of Reading, Reading, United Kingdom

(Manuscript received 15 October 2003, in final form 30 October 2004)

ABSTRACT

Most current general circulation models (GCMs) calculate radiative fluxes through partially cloudy grid boxes by weighting clear and cloudy fluxes by the fractional area of cloud cover (C_a), but most GCM cloud schemes calculate cloud fraction as the volume of the grid box that is filled with cloud (C_v). In this paper, 1 yr of cloud radar and lidar observations from Chilbolton in southern England, are used to examine this discrepancy. With a vertical resolution of 300 m it is found that, on average, C_a is 20% greater than C_v, and with a vertical resolution of 1 km, C_a is greater than C_v by a factor of 2. The difference is around a factor of 2 larger for liquid water clouds than for ice clouds, and also increases with wind shear. Using C_a rather than C_v, calculated on an operational model grid, increases the mean total cloud cover from 53% to 63%, and so is of similar importance to the cloud overlap assumption.

A simple parameterization, $C_a = [1 + e^{(-f)}(C_v^{-1} - 1)]^{-1}$, is proposed to correct for this underestimate based on the observation that the observed relationship between the mean C_a and C_v is symmetric about the line $C_a = 1 - C_v$. The parameter f is a simple function of the horizontal (H) and vertical (V) grid-box dimensions, where for ice clouds $f = 0.0880\ V^{0.7696}\ H^{-0.2254}$ and for liquid clouds $f = 0.1635\ V^{0.6694}\ H^{-0.1882}$.

Implementing this simple parameterization, which excludes the effect of wind shear, on an independent 6-month dataset of cloud radar and lidar observations, accounts for the mean underestimate of C_a for all horizontal and vertical resolutions considered to within 3% of the observed C_a, and reduces the rms error for each individual box from typically 100% to approximately 30%. Small biases remain for both weakly and strongly sheared cases, but this is significantly reduced by incorporating a simple shear dependence in the calculation of the parameter f, which also slightly improves the overall performance of the parameterization for all of the resolutions considered.

1. Introduction

Uncertainty in the representation of clouds in general circulation models (GCMs) is one of the major causes of the broad spread of predicted future climate change (Mitchell 2000; Stocker 2001), mostly via the impact on radiative heat fluxes and the resultant cloud-climate feedbacks. Many climate and weather forecasting GCMs use two variables to represent the cloud properties in each grid box. Typically one represents the mean mass concentration of condensate across the grid box and the other represents the fraction of the grid box that contains cloud. This cloud fraction is defined either by area (C_a) or by volume (C_v) with the distinction shown schematically in Fig. 1. By definition, C_a will always be greater than or equal to C_v, and we can expect this difference to increase with the vertical dimension of the grid box.

Most schemes for determining the cloud fraction yield C_v, which is more easily related to the mass of condensate. For example, in the European Centre for Medium-Range Weather Forecasts (ECMWF) cloud scheme of Tiedtke (1993) and the Met Office Unified Model (Wilson and Ballard 1999; Smith 1990), large-scale cloud fraction is defined to be the proportion of the humidity distribution across the grid box that exceeds saturation, thus giving C_v directly. However, C_a is more appropriate for calculating the radiative effect of cloud (Stephens 1984; Edwards and Slingo 1996) and the representation of precipitation (Jakob and Klein 1999), but most GCMs assume that the cloudy area of a grid box fills the entire grid box in the vertical, thus setting C_a equal to C_v. This discrepancy may go some way in explaining discrepancies between the observed and modeled radiative fluxes discussed by Webb et al. (2001).

The only previous reference to the distinction be-

* Current affiliation: Met Office, Exeter, United Kingdom.

Corresponding author address: Malcolm E. Brooks, Met Office, FitzRoy Road, Exeter, EX1 3PB, United Kingdom.
E-mail: Malcolm.E.Brooks@metoffice.com

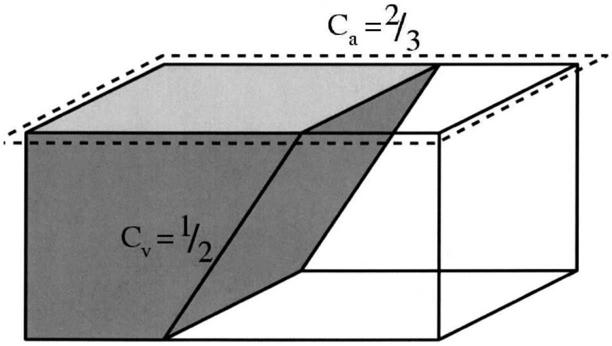

FIG. 1. Schematic of a distribution of clouds within a 3D grid box, where the cloud fraction by volume (C_v) is ½, but the cloud fraction by area (C_a) is ⅔.

tween C_a and C_v appears to be that in Del Genio et al. (1996), which gives $C_a = C_v^{2/3}$ by assuming that cloud within a grid box takes the form of a regular cube, which does not fill the grid box in the vertical. However when considering real clouds it is apparent that they are not regular cubes, and the real reason that C_a and C_v differ is that clouds have irregular geometries both in the horizontal and vertical, on scales much smaller than the grid box.

It is also worthy of note that the difference between C_a and C_v is not dealt with by the cloud overlap assumptions used in the models, as these describe the arrangement of cloud between different grid boxes in the vertical (see Morcrette and Fouquart 1986; Tian and Curry 1989). In contrast, the difference between C_a and C_v is more a representation of subgrid cloud geometry, and the combination of the parameterization detailed in this paper with a realistic overlap assumption such as Hogan and Illingworth (2000), would enable an unbiased total cloud cover to be determined from a profile of individual C_v values.

In this paper we use radar and lidar to directly measure the values of C_a and C_v that would be simulated in a model, and develop a parameterization to enable models to determine C_a from C_v as a function of grid-box size, cloud phase, and wind shear. The ability of radar and lidar to measure the profile of cloud occurrence with high resolution was demonstrated by Mace et al. (1998), Clothiaux et al. (2000), and Hogan et al. (2001). Indeed Hogan et al. (2001) compared three months of observed C_v with the values held in the ECMWF model and found that although the model simulated the frequency of cloud occurrence well, the amount of cloud when present was overestimated above 6 km and underestimated below 6 km. The underestimate below 6 km would become even more severe if observed C_a were compared with the values of C_v currently used within such GCMs for the partitioning of radiative fluxes.

In section 2 of this paper, we outline the observational methods used to derive values of C_a and C_v, for comparison in section 3. In section 4, a parameterization is developed for use within a GCM to obtain the C_a for cloud from the currently held C_v, which is evaluated in section 5.

2. Observational methods and data

The primary observations are made from the 94-GHz Galileo cloud radar and Vaisala CT75K lidar ceilometer, at Chilbolton in southern England. At this site, stratiform and frontal cloud are predominantly observed. The radar and lidar are vertically pointing and operated near-continuously from 1 May 1999 until 26 May 2000. Examples of the radar and lidar observations for 9 May 1999 are shown in Figs. 2a and 2b.

a. *The 94-GHz cloud radar*

The 94-GHz cloud radar records vertical profiles of reflectivity factor (Z), with a vertical resolution of 60 m, averaged over a 30-s time period, and is particularly sensitive to larger cloud particles. The cloud radar was calibrated by comparison with the Chilbolton 3-GHz weather radar in drizzle, where the droplets are small enough to Rayleigh scatter at both 3 and 94 GHz whereas larger hydrometeors will Mie scatter at 94 GHz. The 3-GHz weather radar at Chilbolton was absolutely calibrated using the redundancy of the polarization parameters in heavy rain (Goddard et al. 1994; Hogan et al. 2003a). Heavy rainfall results in significant attenuation of the 94-GHz signal, and so all occasions when the rainfall rate exceeds 0.5 mm h^{-1}, as measured by a rapid response drop counting rain gauge, were excluded from the analysis.

b. *Lidar ceilometer*

To complement the radar observations, a 905-nm Vaisala CT75K lidar ceilometer was used. This records profiles of lidar backscatter (β), which is approximately proportional to particle diameter to the second power. In comparison to the Z returned from a radar, the lidar return is much more sensitive to high concentrations of small droplets, so it is therefore able to distinguish the liquid cloud base from any precipitation falling from the cloud, as well as detect the cloud base of even thin liquid water clouds that may be undetectable by the radar alone. The disadvantage of the lidar instrument is that the lidar beam is rapidly attenuated by the presence of any liquid water clouds, so producing lidar information only at or slightly above cloud base.

c. *Model fields*

The model wind and temperature data used in this study were taken from daily operational ECMWF fore-

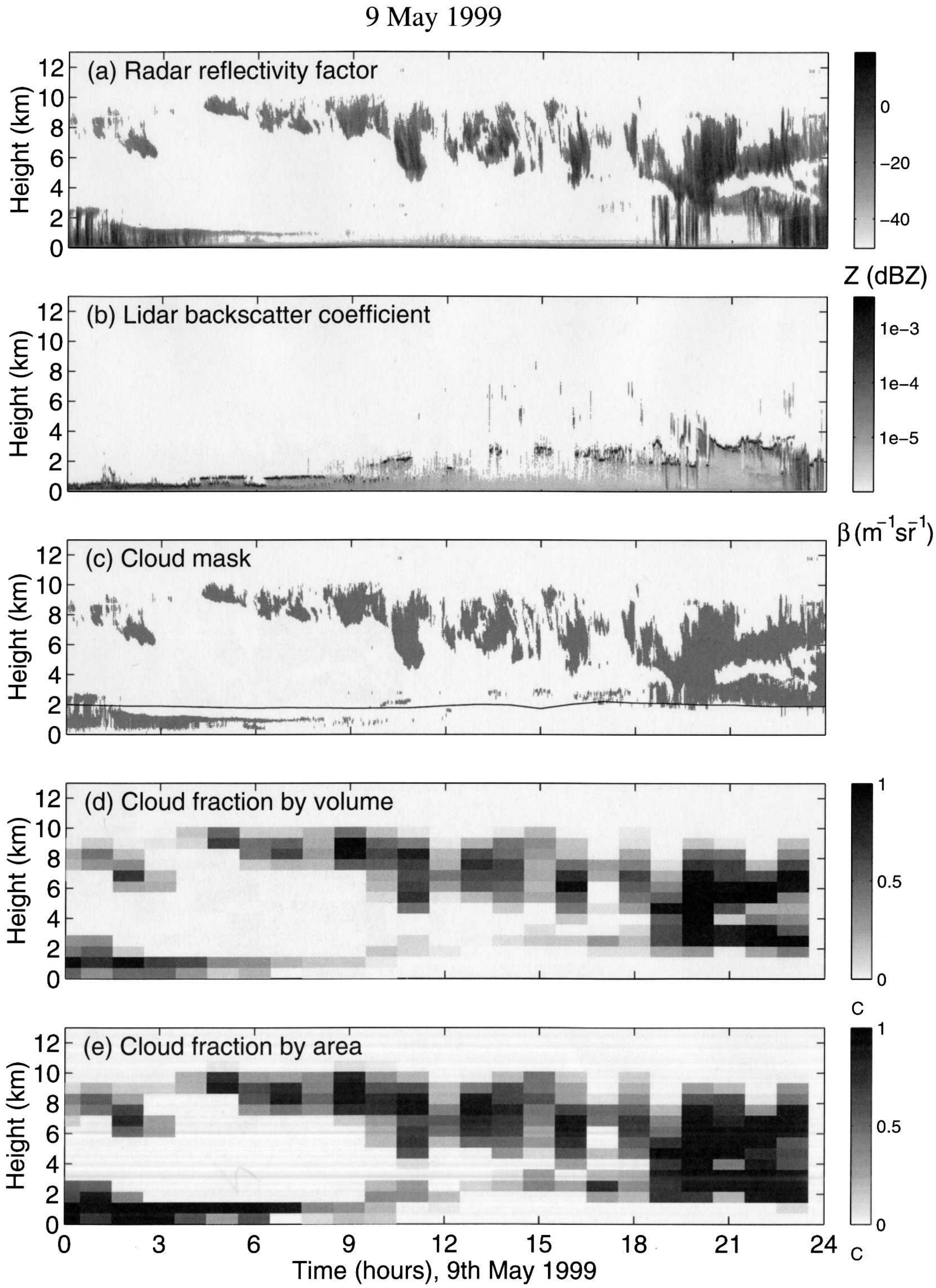

FIG. 2. Time–height sections of (a) 94-GHz radar reflectivity, (b) lidar backscatter coefficient, and (c) the corresponding cloud mask with the model wet-bulb 0°C isotherm superposed. In this example, cloud fractions by (d) volume and (e) area are calculated for a grid box with 720-m vertical resolution, and a temporal resolution of 1 h (roughly equivalent to 65-km horizontal resolution, based on the annual mean tropospheric wind speed of 18 m s^{-1}).

casts and Met Office operational mesoscale forecasts. The data were extracted for the model grid box closest to Chilbolton and a long time series was created by concatenating consecutive model forecasts, taken daily at 12 to 36 h after the analysis time (termed T+12 to T+36) for the ECMWF and every 6 h at T+4 to T+10 for the Met Office.

During the period of this study the ECMWF spectral model used T_L319 truncation, corresponding to a horizontal resolution of around 60 km, and with 25 model levels below 15 km. The Met Office model operated with a horizontal resolution of 0.11° × 0.11°, which is equivalent to approximately 11 km, and with 32 model levels below 15 km.

d. Defining cloud fraction

To derive the cloud fraction from the radar and lidar observations, the radar and lidar data are analyzed in grid boxes, which contain a large number of the radar pixels, of height 60 m and a time of 30 s, corresponding to the sampling of the instrument. The radar and lidar returns from each pixel are then analyzed and each pixel is determined as either cloudy or cloud free, producing a cloud mask. Then C_v is simply the fraction of the pixels within the grid box that are deemed to be cloudy, while C_a is the fraction of the area of the grid box that contains cloud when viewed from above or below.

To determine whether an individual pixel is cloudy, we adopt the methodology of Hogan et al. (2001). A radar return is determined to be from cloud when it is located either above the freezing level (defined as the height at which the model wet-bulb temperature is 0°C), or above the lidar cloud base. This excludes any rain or drizzle falling from the base of the cloud, or insects below cloud base, but retains the ice. Mittermaier and Illingworth (2003) compared the freezing levels in the operational Met Office and ECMWF models and found them to be in close agreement with the radar observations.

Thin, nonprecipitating liquid water clouds are also observed by the lidar that the radar may not be able to detect. The bases of these clouds are clearly indicated by high values of β ($>6 \times 10^{-4}$ sr^{-1} m^{-1}) observed by the lidar, and are included in the cloud mask with an assumed thickness of 180 m (3 radar pixels). Clouds thicker than this typically develop sufficient drizzle droplets to become observable by the radar.

The resultant array of cloudy or noncloudy pixels, hereafter referred to as the cloud mask and shown in Fig. 2c, forms a time–height section of cloud at the radar's high temporal (30 s) and spatial (60 m) resolution, from which C_a and C_v are calculated as described above. The example cloud fractions shown in Fig. 2 are estimated over an hour with a vertical resolution of 720 m, and so are based on 1440 individually cloudy or noncloudy radar pixels. Although the cloud mask is a two-dimensional slice, given sufficient samples it is believed to be representative of the three-dimensional cloud field.

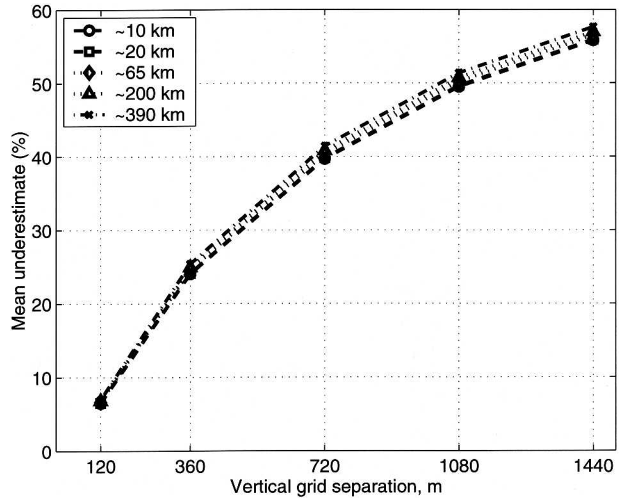

FIG. 3. Mean underestimate of C_a by the use of C_v, as a percentage of the observed mean C_a, against vertical grid dimension, for radar and lidar observations analyzed for the presence of cloud and gridded at various different horizontal resolutions. The analysis includes all events, both cloudy and noncloudy, observed at Chilbolton in odd numbered months between 1 May 1999 and 26 May 2000. Note the insensitivity of the underestimate to the horizontal resolution.

To investigate the degree of difference between cloud fraction by volume and cloud fraction by area, C_a and C_v were calculated from the cloud masks on a variety of regular grids, with time intervals of 10, 20, 60, 180, and 360 min and 120, 360, 720, 1080, and 1440 m in the vertical. The time-averaging periods are converted to horizontal resolutions using the ECMWF wind speeds, and on average, the selected time intervals are equivalent to horizontal resolutions of approximately 10, 20, 65, 200, and 390 km. Initially only odd numbered months will be analyzed, so that the even numbered months can be used as an independent dataset upon which to test the emerging parameterization.

3. Evaluating cloud fraction by area and volume

To gauge the significance of the difference between C_a and C_v, Fig. 3 shows the mean underestimate of C_a by C_v as function of vertical dimension of the grid box, for a variety of horizontal resolutions, for all events in the odd numbered months in the yearlong dataset. It is significant that the mean underestimate is all but independent of the horizontal grid-box dimension (H), and strongly dependent on the model vertical grid-box dimension (V). This property is based on an average of 6 months of radar and lidar observations. Examining a single month's data shows similar results, but with rather more noise.

For a typical numerical weather prediction (NWP) model, vertical resolution is currently of the order of 500 to 1000 m in the midtroposphere, which corresponds to an underestimate of C_a by C_v of 30% to 50%. This underestimate is reduced with the vertical grid-box dimension, but even with a vertical resolution of 120 m, which is a typical vertical resolution of an NWP model within the boundary layer, C_v is still a 7% underestimate of the observed C_a. This would indicate that to eliminate the problem solely through increasing vertical resolution would require vertical resolutions throughout the troposphere that may take many decades to be achievable in operational NWP or climate prediction models.

Now that the significance of the underestimate of C_a by C_v has been evaluated, we will examine this bias in more detail. In doing so it is necessary to examine only partially cloudy grid boxes, where $0 < C_v < 1$, as by definition, C_a and C_v are equal for either fully clear or cloudy grid boxes. Figure 4 shows the mean underestimate of C_a by C_v for partially cloudy grid boxes with V and H.

When looking only at partially cloudy grid boxes, the effect of increasing V is still apparent, but the mean underestimate decreases as H increases. This is due to the subdivision of a given cloud field producing more grid boxes that are either completely clear or completely cloudy, in which C_a and C_v are equal. In the remaining partially cloudy grid boxes, the underestimate of C_a by C_v is correspondingly greater. When averaging over all grid boxes, as in Fig. 3, these two effects of varying H cancel out, but in developing a relationship between C_v and C_a we must examine only partially cloudy grid boxes, where the effect of varying H is significant.

With a view to examining the physical processes that lead to the generation of this bias, in Fig. 5 the ECMWF model outputs were used to classify each grid box according to temperature (T) and vertical wind shear (s), which was derived by examining the vector change in the horizontal wind velocity with height. This approach has been successfully used by Hogan and Illingworth (2003) when examining ice-fall streaks and cloud inhomogeneities and is an approach often used in GCMs themselves. For example, it is an intrinsic part of the calculation of turbulent fluxes, and the local Richardson number used in the cloud and boundary layer schemes in the Met Office Unified Model (Smith 1990; Martin et al. 2000). In this analysis it is assumed that within a grid box with $T < -15°C$, only a very small proportion of any cloud is still in liquid phase, so it

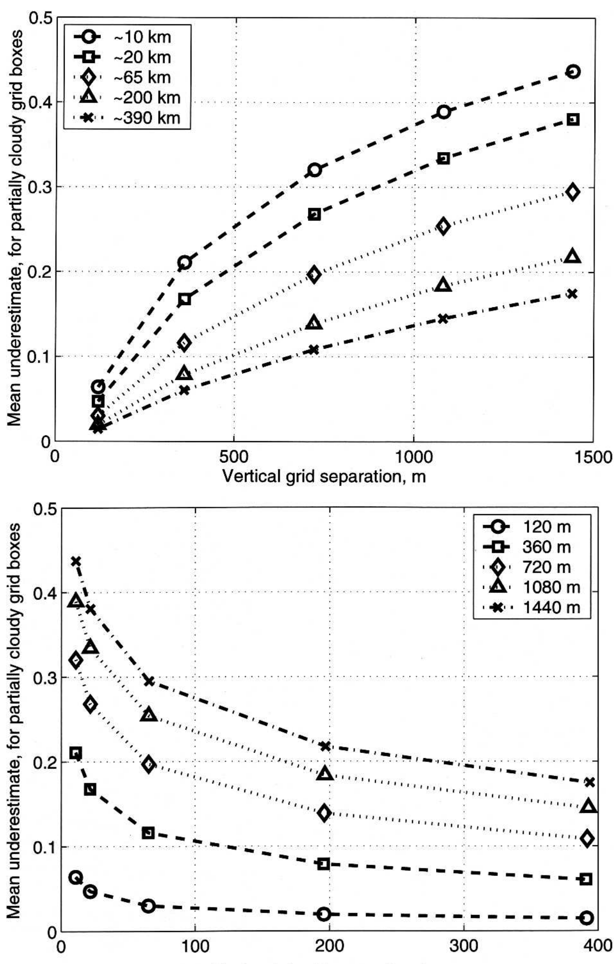

FIG. 4. Mean underestimate of C_a by the use of C_v, with (top) vertical and (bottom) horizontal grid dimension. The analysis includes only partially cloudy events, in odd numbered months between 1 May 1999 and 26 May 2000.

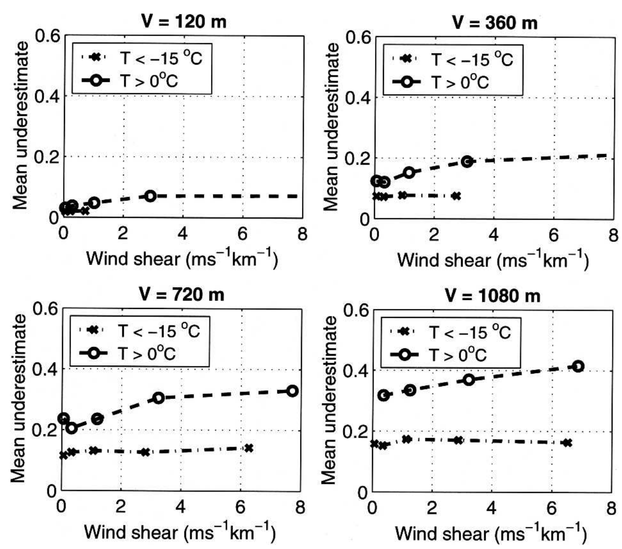

FIG. 5. Mean underestimate of C_a by the use of C_v, varying with wind shear and temperature for grid boxes approximately 65 km in the horizontal, with differing vertical dimension. The analysis includes only partially cloudy events, in odd numbered months between 1 May 1999 and 26 May 2000.

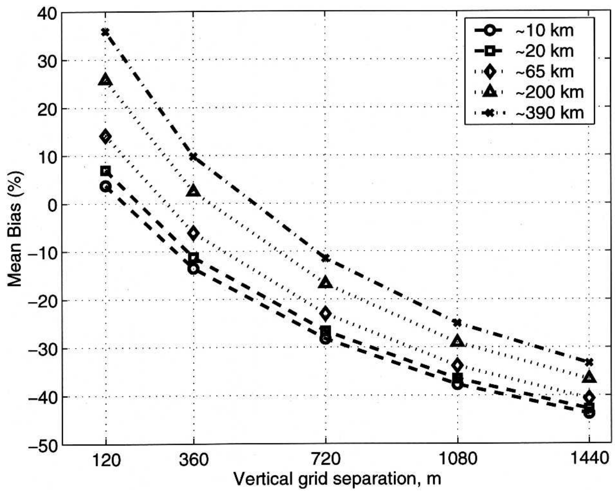

FIG. 6. Mean underestimate of C_a by the observed C_v, corrected as described in Del Genio et al. (1996), as a percentage of the observed mean C_a, against vertical grid dimension, for Chilbolton cloud masks gridded at various different horizontal resolutions. The analysis includes all events, both cloudy and noncloudy, observed in odd numbered months between 1 May 1999 and 26 May 2000.

is classified as ice cloud. These temperature boundaries are selected to remove the majority of the mixed phase clouds (Hogan et al. 2003b; Hogan et al. 2004), clouds with a temperature between 0° and −15°C are ignored.

Figure 5 demonstrates that the mean underestimate of C_a by C_v is most significant for liquid water cloud and increases with wind shear. This arises because liquid water clouds typically form in thinner layers than ice clouds; Mace et al. (1997) found that mean cirrus thickness is 1.6 km, whereas for stratocumulus and supercooled altocumulus clouds, 100 to 200 m is typical, so liquid water clouds are less likely to fill a grid box in the vertical.

As an investigation of the sensitivity of these results to the assumed thickness of those liquid water clouds detected only by lidar, this thickness was raised from 180 to 360 m. The effect of this was to reduce the observed underestimate of C_a by C_v by only a fraction of a percent at the resolutions considered above.

The only reported parameterization for obtaining C_a is that of Del Genio et al. (1996), which uses C_v to mimic C_a by raising it to the power ⅔. Given the resolution dependence of the observed underestimate of C_a by C_v, it is to be expected that the Del Genio et al. (1996) correction will also be highly resolution dependent. Figure 6 shows the mean bias in C_v against the observed C_a after correction by the Del Genio et al. (1996) method, as a percentage of the mean C_a. For the larger vertical resolutions there is still a significant underestimate in C_a of up to 40%, although this is a reduction from almost 60% for unmodified C_v in Fig. 3. As the vertical grid separation decreases, so the Del Genio et al. (1996) becomes an overestimate of C_a by up to 40%. There is hence a need to seek a more appropriate parametric from relating C_v and C_a.

4. Parameterizing cloud fraction by area

In section 3 we examined the underestimate of C_a by the use of C_v and found that, for partially cloudy grid boxes, the underestimate increased with vertical grid separation (V), and decreased with horizontal grid separation (H). Also this underestimate was more significant for liquid water clouds than for ice, and increased with wind shear. In this section we aim to produce a simple parameterization that captures these observed trends. To determine an appropriate functional form for a parameterization of C_a from C_v, we will consider the observed relationship between the two cloud properties. Figure 7 shows scatterplots of C_a versus C_v for two example resolutions. Note that only a specific range of horizontal grid-box dimensions are shown. The discretization of the C_a values in Fig. 7a is due to the smaller sample size when averaging over a smaller period of time. Overlaid are the mean values C_a and C_v in bins of C_v, which a successful parameterization should be able to reproduce.

One possible parameterization would be based on the Del Genio et al. (1996) parameterization, but with the exponent of C_v allowed to vary, which we will refer to as the power law parameterization, as shown in Eq. (1):

$$C_a = C_v^D. \tag{1}$$

A second possible parameterization is formulated based on the observation that the line produced by the mean C_a versus C_v in Fig. 7 is roughly symmetrical on reflection in the line $C_a = 1 - C_v$. One suitable equation for such a relationship is given in Eq. (2), and will be referred to as the symmetric parameterization:

$$C_a = [1 + e^{(-f)}(C_v^{-1} - 1)]^{-1}, \tag{2}$$

where the f parameter controls, for $f > 0$, the extent to which C_a is greater than C_v, with $f = 0$ giving $C_a = C_v$, as shown in Fig. 8.

The solid lines overplotted in Fig. 8 show the relationship between C_v and C_a using Eq. (1) and the value of D that best fits observed mean values of C_a in the bins of C_v, in terms of least squares. Similarly, the dashed lines show the relationship between C_a and C_v, using Eq. (2) with a value of f that best fits the mean values of C_a in the C_v bins.

For cases where C_v is small (less than 0.2 and 0.3 in the examples presented), the C_a produced by the

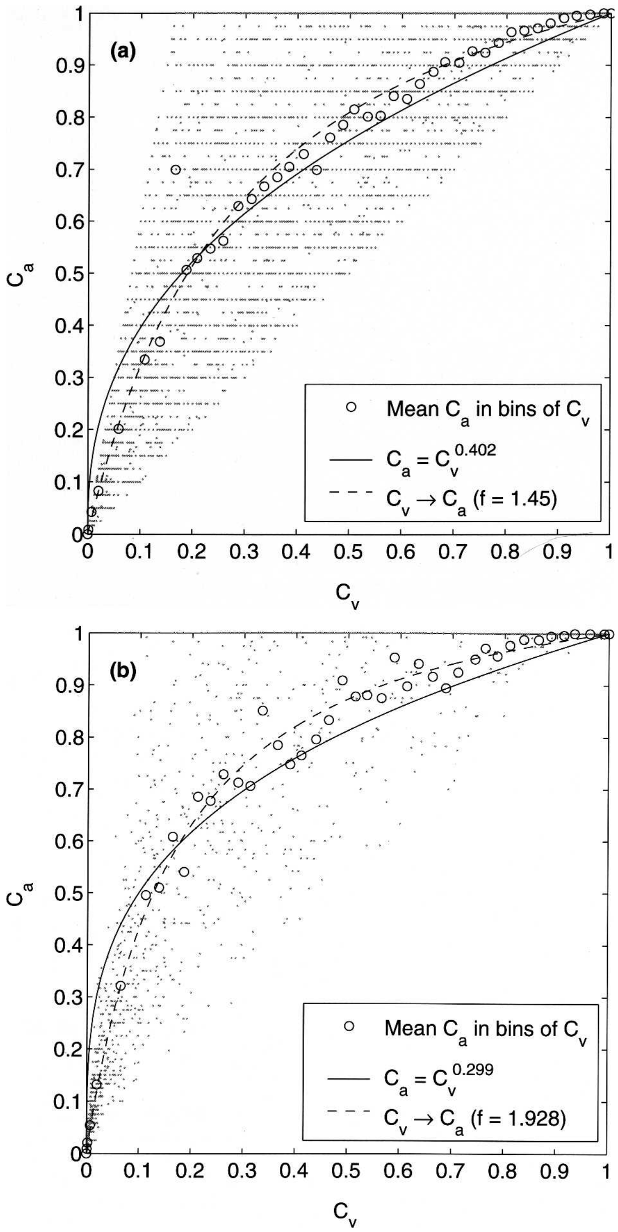

FIG. 7. Scatterplots of cloud fractional area (C_a) vs volume (C_v), as observed by radar and lidar observations. Overlaid are the mean values of C_a and C_v in bins of C_v (the circles), and the results of two possible parameterizations that are discussed in the text.

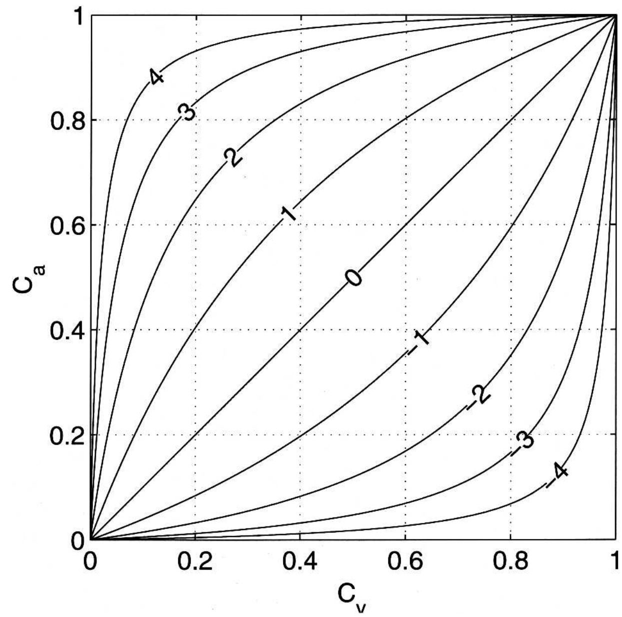

FIG. 8. A demonstration of the relationship between C_a and C_v proposed in Eq. (2), for varying values of f, as labeled.

power law parameterization is significantly overpredicted to the extent that, for a C_v of 0.05, the C_a predicted by (1) is approximately double the mean observed value of C_a. Conversely, for higher values of C_v, C_a is underestimated by the power law parameterization. Neither of these problems is evident when comparing the symmetric parameterization. Although only two example resolutions have been shown here, the problems with the power law (solid lines) are apparent in all of the resolutions examined. If the value of D in Eq. (1) is selected by minimizing the rms error against the raw data points, rather than the mean values of C_a and C_v in C_v bins, then the value of D is altered to reflect the distribution of C_v, to the extent that the overestimate of low cloud fractions is slightly reduced, at the expense of drastically increasing the underestimate of C_a for larger C_v. We therefore conclude that the symmetric parameterization, given in Eq. (2) best describes the mean dependence of C_a on C_v, and we will continue in this section to develop this parameterization further.

To determine the appropriate values of f, the cloud fraction data were reanalyzed, splitting the observed values of C_a and C_v according to their vertical (V) and horizontal (H) grid sizes, where the ECMWF wind speed, at the specific time and height, was used determine the actual horizontal distance which a given averaging period represents. The variability of wind speeds with height is such that the number distribution of cases with a given H will vary with height, but the variability of the wind speed with time ensures that all H bins examined are well populated at all heights. Such an approach significantly improves the fits of f with H compared to using a single wind speed with height to convert to a horizontal distance. The resultant variation of f with H and V is shown in Fig. 9.

It is apparent from Fig. 9 that f increases with V, and decreases with H, which is consistent with the relationship between the mean underestimate of C_a by C_v, and H and V shown in Fig. 4. To determine

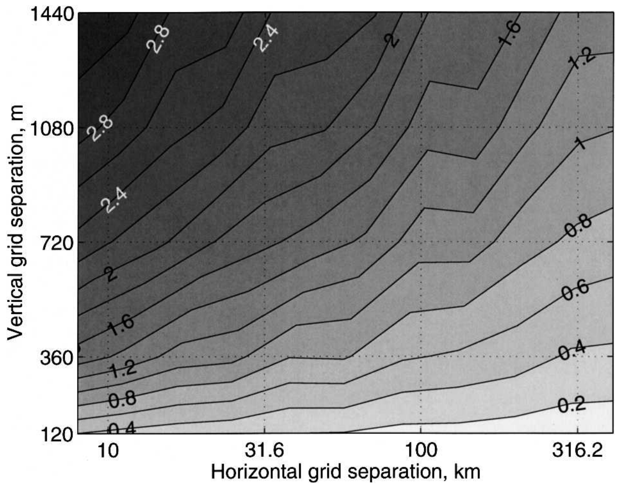

FIG. 9. Variation of the parameter f in Eq. (2) that best fits the observed relationship between C_a and C_v, averaged over bins of C_v, as a function of horizontal and vertical resolution. The analysis uses observations from the Chilbolton radar and lidar in odd numbered months between 1 May 1999 and 26 May 2000.

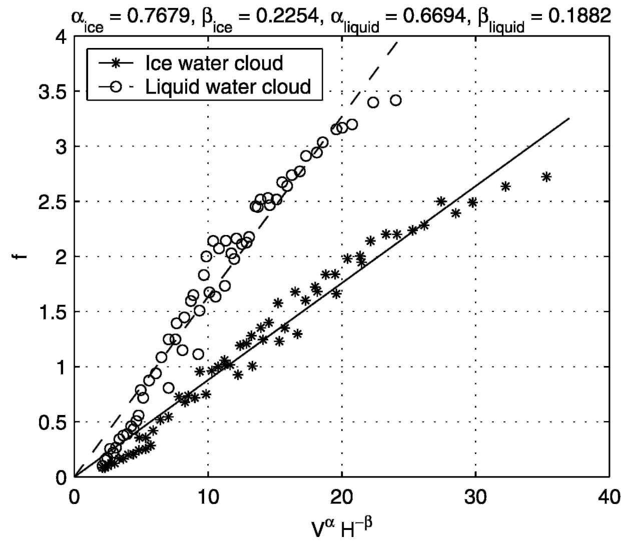

FIG. 10. Variation of the parameter f, in (2) which best fits the observed variation of mean C_a with C_v, against $V^{\alpha}\ H^{-\beta}$, where α and β are the parameter pair that gives the best fit of the observed f, in a least squares sense. The analysis uses observations from the Chilbolton radar and lidar in odd numbered months between 1 May 1999 and 26 May 2000, which are characterized as containing ice or liquid clouds using the model temperature profiles.

the exact relationship between f and the model resolution, we assume that as V tends to zero, or H tends to infinity, so C_a will tend to C_v, and therefore f tend to zero. We therefore assume that f may be given in (3),

$$f = AV^{\alpha}H^{-\beta}, \tag{3}$$

where A, α, and β are parameters to be fitted to the observations.

Based on the observation in section 3 that the underestimate of C_a by C_v was greater for liquid water clouds than for ice, the coefficients in Eq. (3) need to be determined separately. To determine the appropriate values for the coefficients in (3), α and β were allowed to vary, iteratively selecting the pair of values for which (3) gave the best fit to the observed values of f, in terms of least squares. The results of this analysis are shown in Fig. 10, where the analysis has been performed separately for ice and liquid water clouds, using the model temperature (T) profiles as in Fig. 5.

Given the dependence of $\overline{C_a - C_v}$, with wind shear (s) observed in Fig. 5, a logical extension to this work is to categorize clouds by wind shear as well as by cloud phase. For simplicity we assume that as wind shear varies, Eq. (3) still applies and that α and β remain constant while we allow A to vary.

The analysis shown in Fig. 10 was repeated for a number of wind shear bins; within each wind shear bin a shear-dependent value of A was computed as the gradient of f with $V^{\alpha}\ H^{-\beta}$. The variation of the mean observed $A(s)$ with wind shear is shown in Fig. 11a, and in order to represent this data, we define the limit of $A(s)$ as s tends to zero $[A(s = 0)]$. For the purposes of fitting a power law, Fig. 11b shows the variation of $A(s) - A(s = 0)$ against wind shear on logarithmic axis, and the resultant power laws are overplotted on Figs. 11a and 11b.

Therefore the f parameter in Eq. (2) can be parameterized as follows, with all terms in the equations in SI units. We propose that A can be determined either with or without a wind shear dependence as appropriate to the user requirements:

$$\text{for ice clouds:} \quad f = A_{\text{ice}}\, V^{0.7679}H^{-0.2254} \tag{4}$$

$$\text{ignoring shear:} \quad A_{\text{ice}} = 0.0880 \tag{5}$$

$$\text{and with shear:} \quad A_{\text{ice}}(s) = 0.0706 + 0.1274\, s^{0.3015} \tag{6}$$

$$\text{for liquid clouds:} \quad f = A_{\text{liq}}\, V^{0.6694}H^{-0.1882} \tag{7}$$

$$\text{ignoring shear:} \quad A_{\text{liq}} = 0.1635 \tag{8}$$

$$\text{and with shear:} \quad A_{\text{liq}}(s) = 0.1105 + 1.1906\, s^{0.5112}. \tag{9}$$

With this parameterization, the correction of C_v to C_a is increased with wind shear for both ice and liquid water cloud. However, ice cloud is less sensitive to wind shear, both in that the overall effect of wind shear is less, and that the shear term reaches an almost asymptotic value at lower wind shears than for liquid cloud (consistent with Fig. 5). One physical explanation for this would be that the liquid water cloud droplets fall more slowly than ice particles so that a given wind shear gives a larger difference in

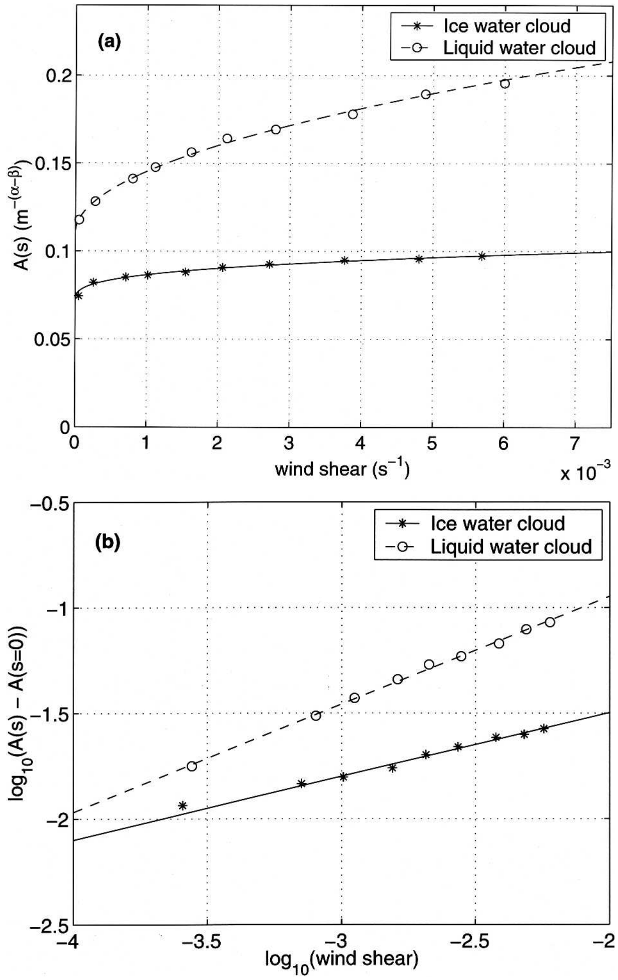

FIG. 11. Variation of the parameter A in Eq. (3) with (a) wind shear and (b) $\log_{10}$ (wind shear). The analysis uses observations from the Chilbolton radar and lidar in odd numbered months between 1 May 1999 and 26 May 2000, which are characterized using the model profiles as containing ice or liquid clouds by temperature, and subdivided into bins of wind shear, determined as the change in the horizontal wind velocity with height.

horizontal transport for the more slowly falling particles.

5. Evaluation of the C_v to C_a parameterizations

To evaluate the parameterizations given above, a test period consisting of the even numbered months over the period 1 May 1999 to 26 May 2000, was examined analyzed to produce an independent dataset of cloud fractions, C_a and C_v, using the same time and height resolutions as used in sections 3 and 4, and converting time to H using the ECMWF wind speed. The values of C_v were corrected using the symmetric parameterization, (2), and values of f obtained, using Eqs. (4)–(9), by first ignoring and then including the effects of wind shear and compared to the observed C_a.

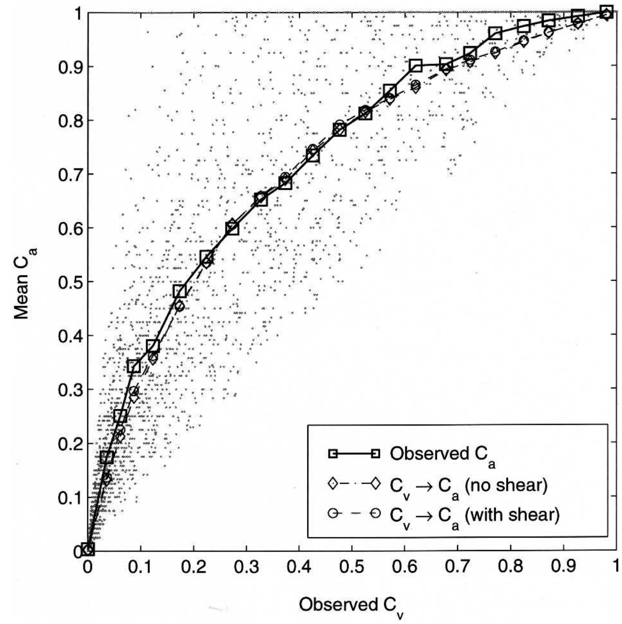

FIG. 12. Scatterplots of C_a vs C_v, as observed by radar and lidar, during the even numbered months between 1 May 1999 and 26 May 2000. Overlaid are the mean values, in bins of C_v, of the observed C_a and C_v corrected to represent C_a using the symmetric parameterization (2), with f found both without wind shear [(4), (5), (7), (8)] and with wind shear [(4), (6), (7), (9)].

Figure 12 shows, for an example resolution of 65 km in the horizontal by 720 m in the vertical, a scatterplot of C_a and C_v, and the mean relationship between C_a and C_v as observed and predicted by the symmetric parameterization, (2), obtaining f both with and without wind shear. It is clear that the use of C_v to represent C_a leads to a significant underestimate in C_a for partially cloudy grid boxes, and that averaging over a long period both methods of obtaining f give mean values of C_a that are in good agreement with the observed C_a across the full range of C_v.

Tables 1, 2, and 3 give the mean bias and rms error in the prediction of C_a by C_v during the test period, for uncorrected C_v and C_v and then corrected both without and with wind shear information, respectively. These statistics are reweighted to give equal weight to each grid resolution examined. For the uncorrected C_v, the mean bias is considerable, with the mean C_a during this period, and range of resolutions, is 0.126, and the mean C_v 0.079, which corresponds to an underestimate of 0.046, or 37% of the mean C_a. This is consistent with the mean underestimate from the odd numbered months as in Fig. 3. Ice clouds exhibit a less significant bias between C_a and C_v than liquid water clouds, with an underestimate of 27% compared to 48%, respectively. Both methods of correcting C_v (Tables 2 and 3) give good results, with the mean bias in C_a predicted from C_v

TABLE 1. Biases (both absolute, and in terms of the percentage of mean C_a), and rms errors, in the representation of C_a by C_v. The analysis is based on cloud masks of radar and lidar data over Chilbolton during even numbered months between 1 May 1999 and 26 May 2000.

	C_v (uncorrected)	
	Bias	Rms error
All events	−0.046 (−37%)	±0.139 (±110%)
$H < 20$ km	−0.061 (−47%)	±0.178 (±136%)
20 km $< H <$ 100 km	−0.045 (−39%)	±0.139 (±120%)
$H > 200$ km	−0.026 (−27%)	±0.071 (±74%)
$V < 500$ m	−0.014 (−15%)	±0.059 (±63%)
$V > 1000$ m	−0.078 (−50%)	±0.190 (±121%)
$T < -15$°C	−0.025 (−27%)	±0.087 (±93%)
$T > 0$°C	−0.103 (−48%)	±0.226 (±105%)
$s < 0.5$ m s^{-1} km^{-1}	−0.035 (−33%)	±0.109 (±102%)
$s > 3$ m s^{-1} km^{-1}	−0.080 (−46%)	±0.198 (±114%)

TABLE 3. Biases (both absolute, and in terms of the percentage of mean C_a), and rms errors, in the representation of C_a by C_v after correction using the symmetric parameterization (2), with f found including the effect of wind shear [(4), (6), (7), (9)]. The analysis is based on cloud masks of radar and lidar data over Chilbolton during even numbered months between 1 May 1999 and 26 May 2000.

	$C_v \rightarrow C_a$ (with shear)	
	Bias	Rms error
All events	-2.6×10^{-4} (−0.2%)	±0.055 (±33%)
$H < 20$ km	0.0017 (1.3%)	±0.062 (±28%)
20 km $< H <$ 100 km	−0.0031 (−2.7%)	±0.055 (±17%)
$H > 200$ km	2.0×10^{-4} (0.2%)	±0.034 (±3.3%)
$V < 500$ m	5.8×10^{-4} (0.6%)	±0.032 (±24%)
$V > 1000$ m	-3.1×10^{-5} (−0.0%)	±0.072 (±13%)
$T < -15$°C	-7.6×10^{-5} (−0.1%)	±0.037 (±25%)
$T > 0$°C	−0.0010 (−0.5%)	±0.089 (±17%)
$s < 0.5$ m s^{-1} km^{-1}	1.8×10^{-4} (0.2%)	±0.047 (±17%)
$s > 3$ m s^{-1} km^{-1}	−0.0039 (−2.3%)	±0.081 (±11%)

being less than 0.2% of the mean C_a, for both ice and liquid clouds.

In more detail, when correcting without wind shear, using Eqs. (4), (5), (7), (8) shown in Table 2, the parameterization performs well for all resolutions considered with remaining biases varying between +3% where H is less than 20 km, to a bias of −3% for grid boxes with H between 20 and 100 km although there remain biases of +5.7% in instances of low wind shear, and −9.0% for high wind shear. These are much less than the biases in the uncorrected C_v of −33% and −46%.

Correcting C_v using phase and wind shear (Table 3) performs well at all the vertical and horizontal resolutions examined, reducing underestimates ranging from 15%–50% to an underestimate of 2.7% in the very worst case. All of the resolutions considered show improved performance compared to excluding wind shear effects. Clouds in conditions of negligible wind shear exhibit a less significant bias than clouds in a more sheared environment, and when correcting C_v with wind shear the parameterization accounts for this effect.

TABLE 2. Biases (both absolute, and in terms of the percentage of mean C_a), and rms errors, in the representation of C_a by C_v after correction using the symmetric parameterization (2), with f found without wind shear [(4), (5), (7), (8)]. The analysis is based on cloud masks of radar and lidar data over Chilbolton during even numbered months between 1 May 1999 and 26 May 2000.

	$C_v \rightarrow C_a$ (no shear)	
	Bias	Rms error
All events	1.3×10^{-4} (0.1%)	±0.058 (±35%)
$H < 20$ km	0.0039 (2.9%)	±0.065 (±30%)
20 km $< H <$ 100 km	−0.0034 (−2.9%)	±0.056 (±18%)
$H > 200$ km	2.5×10^{-4} (0.3%)	±0.034 (±3.5%)
$V < 500$ m	−0.0012 (1.3%)	±0.035 (±26%)
$V > 1000$ m	6.5×10^{-4} (0.4%)	±0.074 (±14%)
$T < -15$°C	1.3×10^{-4} (0.1%)	±0.037 (±25%)
$T > 0$°C	1.6×10^{-4} (0.1%)	±0.09 (±18%)
$s < 0.5$ m s^{-1} km^{-1}	0.0061 (5.7%)	±0.052 (±19%)
$s > 3$ m s^{-1} km^{-1}	−0.0156 (−9.0%)	±0.086 (±11%)

Both parameterizations reduce the rms errors of the corrected C_v by a factor of 2 to 3, relative to the uncorrected C_v, but are still significant. This is to be expected given that the scheme provides an average correction to account for the random arrangement of cloud within the grid box, however by introducing the ability to both over- and underestimate C_a, the bias is all but eliminated. The inclusion of the effect of wind shear reduces the rms error by 2% (i.e., from 35% to 33%) relative to the parameterization that excludes the effect of wind shear. Combining the 15% variation in the bias of the parameterization between high and low wind shear environments, when wind shear is ignored, with the small improvements across the range of resolutions, when wind shear effects are included, indicates that including the effect of wind shear gives a discernable improvement to the performance of the parameterization, although the effect is small relative to magnitude of the correction that is being applied.

To quantify the impact of the C_a parameterization with a realistic GCM resolution, the radar and lidar data were analyzed as described in section 2 to obtain C_a and C_v on the same grid as the Met Office mesoscale model outputs in the vertical and where the time-averaging period was determined as the time taken to advect a 12-km grid box using the model wind speed. This is opposite approach to averaging for a set period of time and using the model winds to calculate H, which is the approach adopted in the previous sections (in practice, the results from the two approaches are indistinguishable). The observed mean profile of C_v is shown in Fig. 13 and is broadly comparable with those output by Met Of-

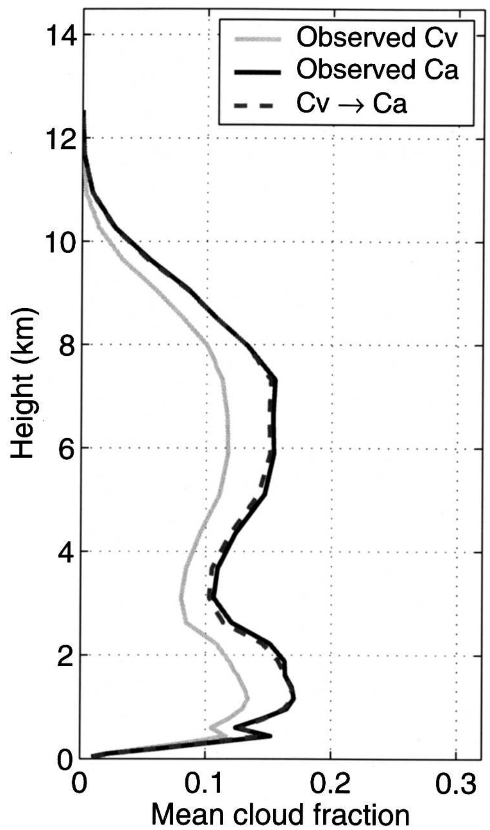

FIG. 13. Profiles of mean cloud fraction produced by averaging the observed occurrence of cloud on the radar and lidar grid onto the grid of Met Office mesoscale model. Profiles shown are the mean observed C_v and C_a, and the mean profile of C_a calculated from the observed C_v using Eq. (2).

fice mesoscale model, although a detailed evaluation of the differences between the observed cloud fractions and the model outputs is beyond the scope of this paper. Also shown is the mean profile of the observed C_v and the C_a produced by correcting the observed C_v using Eq. (2), with the f parameter determined using appropriate values of H and V, and values of wind shear taken directly from the model winds.

The mean profiles of C_a are approximately 30% higher than the mean profile of C_v, which is consistent with Fig. 3, while the profile of parameterized C_a agrees with the observed mean C_a to within 2% at all heights. There is no noticeable difference between the parameterization with and without wind shear when examining averages over long periods of time.

Last, we calculate the total cloud cover from the profiles of C_a and C_v, by applying the most commonly used maximum-random overlap assumption as defined in Geleyn and Hollingsworth (1979). The mean total cloud cover is found to be 53% when C_v is used, and 63% when using C_a (observed or parameterized). This increase is comparable to the 11% increase found by Morcrette and Jakob (2000) when changing from the maximum to the random overlap assumption in the ECMWF model, which in terms of global mean radiative fluxes at the top of the atmosphere reduced the absorbed shortwave radiation by 3.3 W m^{-2} and the outgoing longwave radiation by 8.3 W m^{-2}.

6. Conclusions

A quasi-continuous dataset of radar reflectivity and lidar backscatter over Chilbolton, in southern England, has been used to produce a climatology of cloud fractions, at a variety of horizontal and vertical resolutions appropriate for comparison with GCMs. By calculating cloud fractions by area and by volume on regular grid, a systematic underestimate has been found in which cloud fractions calculated by volume (C_v) are always lower than the corresponding cloud fraction by area (C_a), which is of more interest in the calculation of fluxes of radiation and precipitation. For example, at typical model vertical grid-box spacings of 500 to 1000 m, we observe that C_a is underestimated by C_v by 30% to 50%. This is caused by the stratification of cloud layers occurring with vertical scales that cannot be fully resolved at these resolutions, and by a degree of randomness with which cloud forms, so spreading cloud in the horizontal, rather than filling a grid box in the vertical. The underestimate of C_a by C_v has been shown to be approximately twice as large for liquid water clouds than for ice and this underestimate also increases with wind shear.

In the interest of expressing this effect in a manner that is easily applicable to most GCM cloud schemes, a parameterization has been developed that gives a correction to the C_v, for frontal/stratiform cloud. This parameterization is based on the observation that the mean relationship between C_a versus C_v is roughly symmetric about the line $C_a = 1 - C_v$, and varies with the horizontal and vertical grid-box dimensions (H and V, respectively), cloud phase, and to a lesser extent, wind shear. This approach performed better than the formulation of Del Genio et al. (1996), where $C_a = C_v^{2/3}$, even when the exponent was allowed to vary.

The resultant parameterizations, both excluding and including the effect of wind shear, have been tested against an independent dataset of Chilbolton radar and lidar data, and were found to give good agreement across the range C_v. On average, both correction schemes were found to reduce the mean underestimate of C_a by C_v from 37% (at the range of horizontal and vertical resolutions used) to a mean bias of less than 0.2% both with and without the use of wind shear information. The parameterizations cope with the variation in this underestimate with grid-box resolution, as well as phase. However when wind shear is ignored, the parameterization produces a variation of approximately 15% of the mean C_a, between the low and high wind shear environments. When wind shear is included, this variation is accounted for, and the overall performance of the parameterization is improved, although this effect is small relative to the magnitude of the correction that is applied.

When examining cloud fractions on a grid equivalent to that of the Met Office mesoscale model outputs, the mean profile of C_a is well captured by the parameterization. When using C_a rather than C_v, the mean total cloudiness of the column is increased from 53% to 63%. We would therefore expect the implementation of this parameterization of C_a would have significant radiative impacts, although the issue is complicated by the representation of the inhomogeneity of water content that could accompany a C_a parameterization.

Before applying the findings of this study it would be useful to study the subgrid geometry statistics from cloud observed in other locations, since the properties of convective cloud may differ significantly from the properties of stratiform and frontal cloud predominantly observed at Chilbolton, although we would expect C_a to be much closer to C_v for upright convective clouds.

Acknowledgments. We thank the Radio Communications Research Unit at the Rutherford Appleton Laboratory for providing the 94-GHz Galileo radar and lidar ceilometer data. This research received funding from the European Union CloudNET project (Grant EVK2-CT-2000-00065) and NERC Grant NER/T/S/1999/00105. We are also grateful to Pete Clark and Richard Forbes at the Met Office for their help and for providing the Unified Model output and to Christian Jakob for providing the ECMWF model data. The first author was also supported by an NERC studentship.

REFERENCES

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner, 2000: Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteor.,* **39,** 645–665.

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

Edwards, J. M., and A. Slingo, 1996: Studies with a flexible new radiation code. I: Choosing a configuration for a large scale model. *Quart. J. Roy. Meteor. Soc.,* **122,** 689–719.

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

Goddard, J. W. F., J. Tan, and M. Thurai, 1994: Technique for calibration of meteorological radars using differential phase. *Electron. Lett.,* **31**, 166–167.

Hogan, R. J., and A. J. Illingworth, 2000: Deriving cloud overlap statistics from radar observations. *Quart. J. Roy. Meteor. Soc.,* **126,** 2903–2909.

——, and ——, 2003: Parameterizing ice cloud inhomogeneities using cloud radar data. *J. Atmos. Sci.,* **60**, 756–767.

——, C. Jakob, and A. J. Illingworth, 2001: Comparison of ECMWF winter season cloud fraction with radar derived values. *J. Appl. Meteor.,* **40,** 513–525.

——, D. Bouniol, D. N. Ladd, E. J. O'Connor, and A. J. Illingworth, 2003a: Absolute calibration of 94/95-GHz radars using rain. *J. Atmos. Oceanic Technol.,* **20**, 572–580.

——, A. J. Illingworth, J. P. V. Poiares Baptista, and E. J. O'Connor, 2003b: Characteristics of mixed-phase clouds: Part II: A climatology from ground-based lidar. *Quart. J. Roy. Meteor. Soc.,* **129,** 2117–2134.

——, M. D. Behera, E. J. O'Connor, and A. J. Illingworth, 2004: Estimating the global distribution of supercooled liquid water clouds using spaceborne lidar. *Geophys. Res. Lett.,* **32,** L05106, doi:10.1029/2003GL018977.

Jakob, C., and S. A. Klein, 1999: The role of varying cloud fraction in the parameterization of microphysical processes in the ECMWF model. *Quart. J. Roy. Meteor. Soc.,* **125,** 941–965.

Mace, G. G., T. P. Ackerman, E. E. Clothiaux, and B. A. Albrecht, 1997: A study of composite cirrus morphology using data from a 94-GHz radar and correlations with temperature and large-scale vertical motion. *J. Geophys. Res.,* **102**, 13 581–13 593.

——, C. Jakob, and K. P. Moran, 1998: Validation of hydrometeor occurrence predicted by the ECMWF model using millimeter wave radar data. *Geophys. Res. Lett.,* **25**, 1645–1648.

Martin, G. M., M. R. Bush, A. R. Brown, A. P. Lock, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part II: Tests in climate and mesoscale models. *Mon. Wea. Rev.,* **128,** 3200–3217.

Mitchell, J. F. B., 2000: Modelling cloud-climate feedbacks in predictions of human-induced climate change. World Climate Research Program Workshop on Cloud Processes and Cloud Feedbacks in Large-Scale Models, WCRP-110, WMO/TD 993.

Mittermaier, M. P., and A. J. Illingworth, 2003: Comparison of model-derived and radar-observed freezing level heights: Implications for vertical reflectivity profile correction schemes. *Quart. J. Roy. Meteor. Soc.,* **129,** 83–96.

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

——, and C. Jakob, 2000: The response of the ECMWF model to changes in the cloud overlap assumption. *Mon. Wea. Rev.,* **128**, 1707–1732.

Smith, R. N. B., 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Quart. J. Roy. Meteor. Soc.,* **116,** 435–460.

Stephens, G. L., 1984: Review: The parameterization of radiation for numerical weather prediction and climate models. *Mon. Wea. Rev.,* **112,** 826–867.

Stocker, T. F., 2001: Physical climate processes and feedbacks. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 944 pp.

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

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Mon. Wea. Rev.,* **121,** 3040–3061.

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

Wilson, D. R., and S. P. Ballard, 1999: A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Quart. J. Roy. Meteor. Soc.,* **125,** 1607–1636.