Objective Determination of Cloud Heights and Radar Reflectivities Using a Combination of Active Remote Sensors at the ARM CART Sites

EUGENE E. CLOTHIAUX,* THOMAS P. ACKERMAN,* GERALD G. MACE,+ KENNETH P. MORAN,#
ROGER T. MARCHAND,* MARK A. MILLER,@ AND BROOKS E. MARTNER#

* *Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania*
+ *Department of Meteorology, University of Utah, Salt Lake City, Utah*
# *NOAA Environmental Technology Laboratory, Boulder, Colorado*
@ *Division of Applied Sciences, Brookhaven National Laboratory, Upton, New York*

(Manuscript received 2 November 1998, in final form 30 April 1999)

ABSTRACT

The U.S. Department of Energy's Atmospheric Radiation Measurement (ARM) Program is deploying sensitive, millimeter-wave cloud radars at its Cloud and Radiation Test Bed (CART) sites in Oklahoma, Alaska, and the tropical western Pacific Ocean. The radars complement optical devices, including a Belfort or Vaisala laser ceilometer and a micropulse lidar, in providing a comprehensive source of information on the vertical distribution of hydrometeors overhead at the sites. An algorithm is described that combines data from these active remote sensors to produce an objective determination of hydrometeor height distributions and estimates of their radar reflectivities, vertical velocities, and Doppler spectral widths, which are optimized for accuracy. These data provide fundamental information for retrieving cloud microphysical properties and assessing the radiative effects of clouds on climate. The algorithm is applied to nine months of data from the CART site in Oklahoma for initial evaluation. Much of the algorithm's calculations deal with merging and optimizing data from the radar's four sequential operating modes, which have differing advantages and limitations, including problems resulting from range sidelobes, range aliasing, and coherent averaging. Two of the modes use advanced phase-coded pulse compression techniques to yield approximately 10 and 15 dB more sensitivity than is available from the two conventional pulse modes. Comparison of cloud-base heights from the Belfort ceilometer and the micropulse lidar confirms small biases found in earlier studies, but recent information about the ceilometer brings the agreement to within 20–30 m. Merged data of the radar's modes were found to miss approximately 5.9% of the clouds detected by the laser systems. Using data from only the radar's two less-sensitive conventional pulse modes would increase the missed detections to 22%–34%. A significant remaining problem is that the radar's lower-altitude data are often contaminated with echoes from nonhydrometeor targets, such as insects.

1. Introduction

A primary objective of the U.S. Department of Energy's Atmospheric Radiation Measurement (ARM) Program is to assess how clouds affect global climate and climate change through their radiative interactions with the earth's surface and atmosphere (Stokes and Schwartz 1994). Accurate knowledge of cloud-layer altitudes, together with the layer temperatures, is fundamental to these assessments. Whereas crude estimates of cloud heights from the ground may be a simple endeavor, obtaining accurate measurements is surprisingly difficult. No single ground-based sensor has proven to be capable of doing the job for all of the wide variety of atmospheric cloud situations. The measurement approach in the ARM program relies on both optical and microwave active remote sensors to establish accurate cloud heights, and on a suite of radiometric instruments to assess their radiative impacts.

The ARM instruments are densely concentrated at the ARM Cloud and Radiation Test Bed (CART) sites, which include the Southern Great Plains (SGP) site near Lamont, Oklahoma; the North Slope of Alaska (NSA) site at Barrow, Alaska; and the tropical western Pacific (TWP) sites on the islands of Nauru and Manus, which are located near the equator. Instruments for profiling cloud heights include the commercial Belfort laser ceilometer (BLC) or Vaisala laser ceilometer (VCEIL) and a micropulse lidar (MPL) described by Spinhirne (1993). The ARM program is complementing these optical devices at the CART sites with the deployment of a high-sensitivity millimeter-wave cloud radar (MMCR) described by Moran et al. (1998). The first MMCRs began operation at the SGP site in November 1996, the NSA site in March 1998, and the TWP-Nauru site in

Corresponding author address: Dr. Eugene E. Clothiaux, Department of Meteorology, 503 Walker Building, The Pennsylvania State University, University Park, PA 16802.
E-mail: cloth@essc.psu.edu

December 1998. In theory, collocated, simultaneous measurements from the laser and radar instruments allow for a more accurate and comprehensive assessment of hydrometeors overhead because each individual instrument has measurement limitations that are circumvented by using them collectively. This article addresses the procedure for putting this synergism into practice.

In general, the laser devices are excellent for detecting essentially all clouds that are visible from the ground and are within the instruments' height measurement ranges. Although its sensitivity is impressive, the millimeter-wave radar fails to detect some of these clouds, especially if the clouds are composed of small hydrometeors, are thin, or are located at high altitudes. The laser instruments are required for these situations. The radar estimation of cloud-base height is often obscured or complicated by the presence of large precipitation and virga particles. Cloud radars also readily detect nonhydrometeor particulates, such as insects and bits of vegetation, which are commonly suspended in the atmospheric boundary layer and constitute airborne clutter that may be mistaken for echoes from stratus clouds. The ceilometer and lidar are usually insensitive to these large, but sparse, precipitation, virga, and insect targets, and therefore provide more accurate cloud-base measurements in these circumstances. However, if the concentration of precipitation particles becomes sufficiently large, then even ceilometer and micropulse lidar determinations of cloud-base height become problematic.

The great strength of radar is its ability to penetrate clouds and reveal multiple layers aloft. The ceilometers and micropulse lidar are unable to provide any information about higher cloud layers when lower liquid-water layers, such as stratus or fog, are present, because the laser beams are readily extinguished by the cloud water and are unable to penetrate much beyond the base of the lowest layer. In these situations, which include periods of drizzle, light rainfall, and most snowfall situations, the radar measurements are essential. When rainfall reaches moderate or heavier intensity, the radar signal is also severely attenuated and not useful.

The MMCR is a 35-GHz (8.7-mm-wavelength, K_a-band) system designed to maximize radar detection of a wide range of cloud conditions by providing excellent sensitivity, resolution, and flexibility of operating options. At the CART sites the radar cycles through a sequence of four operating modes that have differing sensitivities, height coverage, resolutions, and other properties. Two of the modes use a phase-coded pulse compression technique that boosts the radar's sensitivity by 10–15 dB over the other two conventional pulse modes. This allows the detection of weakly reflecting clouds, but it comes at the expense of range sidelobe artifacts, which may contaminate data in the vicinity of stronger echoes, and a higher minimum useable altitude because of incomplete phase decoding in the first several (8–32) range gates. Additionally, some modes have more desirable characteristics than others in terms of unambiguous range (second-trip echoes or range aliasing) and Nyquist unambiguous velocity (velocity folding). In the case of velocity folding, the accuracy of the Doppler moment (reflectivity, mean vertical velocity, and velocity spectral width) calculations is degraded by filtering associated with the coherent averaging used in three of the MMCR modes. The reflectivity data, in particular, are fundamental to various ARM techniques for retrieving ice- and liquid-cloud microphysical properties, such as water–ice mass content and particle sizes. Detailed descriptions of the MMCR radars and their operating characteristics can be found in Moran et al. (1998), Clothiaux et al. (1999a), and Clothiaux et al. (1999b).

The goal of this study was to develop and to test an objective scheme for making optimal use of the four different radar operating modes and the simultaneous, collocated, ceilometer and lidar measurements in order to provide the most accurate available estimates of cloud and precipitation particle heights and Doppler moments over the CART sites. The procedure, in terms of an automated algorithm, and the initial results using many months of CART data are described in the following sections.

2. Methods

Clothiaux et al. (1999a) argue that, when used in combination, the MMCR's four routine operating modes are sufficient to detect most hydrometeors overhead and to accurately measure their reflectivities, vertical velocities, and Doppler spectral widths. Hydrometeors with reflectivities in the range of approximately -50 to $+20$ dBZ are detectable with good accuracy up to heights of 10 km AGL or higher. The four routine modes employ differing hardware operating characteristics that result in different sensitivities and height resolutions. Some modes are also more subject than others to artifacts or accuracy degradations from velocity folding, second-trip echoes, range sidelobes, and coherent averaging filters. Together, however, they cover the wide range of expected tropospheric cloud conditions with good measurement accuracy and resolution.

The individual modes (1–4) may be approximately described as having the following basic observational goals and attributes. [See Clothiaux et al. (1999a) for detailed descriptions of the operational characteristics of the ARM MMCRs, as well as for tables that specify the properties of each mode.] The stratus mode (mode 1) has high sensitivity and vertical resolution for observing boundary layer clouds, which are often composed of small droplets. The cirrus mode (mode 2) uses the highest available sensitivity, enabling it to observe weakly reflecting, high-altitude clouds. The general mode (mode 3) has good sensitivity for detecting a variety of clouds and with relatively few processing artifacts. Last, the robust mode (mode 4) has somewhat-reduced sensitivity, but it is virtually free of the artifact

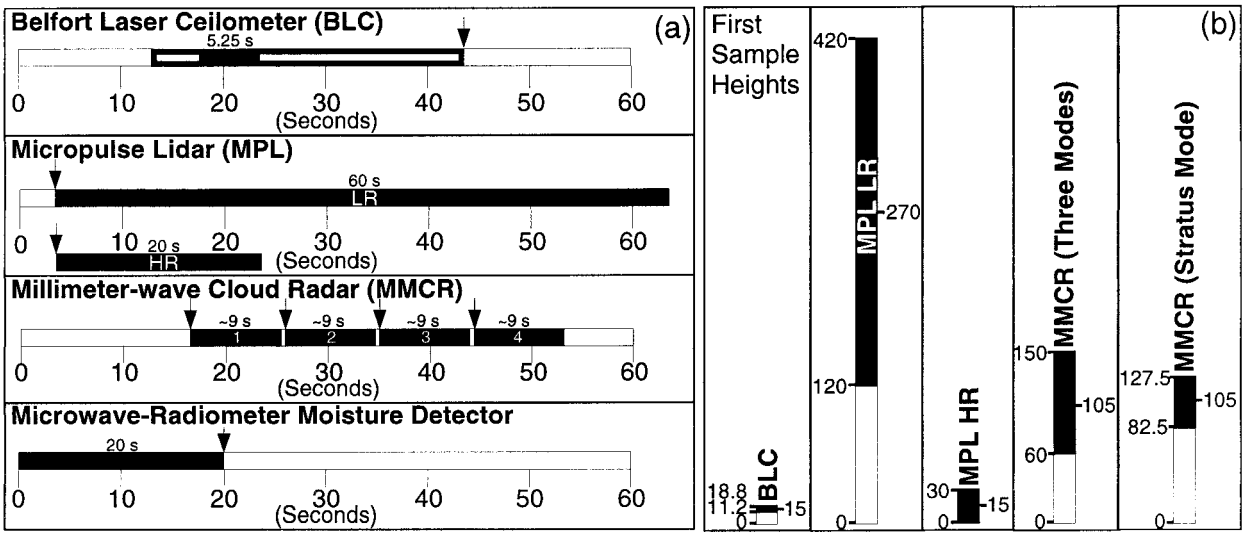

FIG. 1. (a) Temporal and (b) spatial resolutions of the active remote sensors used in this study. The BLC, low-resolution MPL, high-resolution MPL, a single MMCR mode, and the moisture detector have temporal resolutions of 30, 60, 20, 9, and 20 s, respectively. Note that the BLC collects data from the atmosphere during only a 5.25-s interval. Furthermore, only the moisture detector time stamp (time stamps are indicated by arrows) is referenced exactly to the beginning of 0000 UTC for each day; the first time stamp of each day for the other sensors is variable from day to day. The BLC, low-resolution MPL, high-resolution MPL, MMCR stratus mode, and MMCR general mode spatial resolutions [given by the difference of the two numbers to the left of each vertical column in (b)] are 7.6, 300, 30, 45, and 90 m, respectively, and their first sample volume heights, referenced to the middle of the sample volume [as illustrated by the numbers to the right of each vertical column in (b)] are 15, 270, 15, 105, and 105 m, respectively.

and accuracy limitations that may affect the other modes because it has a wide Nyquist range and does not use either pulse coding or coherent averaging. Thus the robust mode provides the most accurate uncorrected measurements of the hydrometeors that it can detect, including weak precipitation cases. Note that modes 1 and 2 utilize a phase-coded pulse compression method to achieve heightened sensitivity, while modes 3 and 4 use conventional pulses.

To develop a complete picture of the vertical distribution of hydrometeors, with an emphasis on cloud particles, at the ARM CART sites we must integrate the data from the four modes together with the ceilometer and lidar data. The temporal and spatial resolutions of the instruments used in this study are presented in Fig. 1. A flow chart of our approach to this integration is presented in Fig. 2 and summarized in Table 1. To illustrate the function of the elements in Fig. 2 we discuss each step in the context of processing radar data that were collected on 7 and 8 May 1997 at the SGP CART site (Fig. 3). The synoptic situation during these two days, which we now describe, is illustrated in Fig. 4.

At 0000 UTC 7 May (Fig. 4a), convection was present in the western parts of Oklahoma and Kansas along a developing cold front. Upper-level westerly winds subsequently advected the cirrus in this region over the SGP central facility between 0000 and 0600 UTC 7 May as seen in Fig. $3K_2$. Also during the period between 0000 and 0600 UTC convective precipitation developed within the warm sector in western and central Oklahoma, reaching the SGP central facility around 0600 UTC. The precipitation at the SGP site ceased around 1200 UTC at which time the site and surrounding area were enveloped in fog and low-level stratus (Fig. $3K_2$). (Dewpoint temperatures were within a few degrees Fahrenheit of the air temperatures at this time.) As the daylight hours progressed from 1200 to 2100 UTC, the surface air temperature and dewpoint at the site increased by 12°F and 4°F, respectively, a combination that caused the lifting condensation and cloud-base heights to rise. By 2100 UTC 7 May the stratus were evaporating and cirrus associated with showers and thunderstorms along a cold front approaching from the northwest reached the site (Fig. $3K_2$, Fig. 4b). Rain showers and thunderstorms occurred throughout the SGP central facility area from 0200 to 0900 UTC 8 May in association with the cold front. The surface front passed the SGP central facility between 0900 and 1200 UTC. From 1200 UTC 8 May to 0000 UTC 9 May low clouds were present throughout the central facility area as a result of residual moisture. During the first part of this period, that is, from 1200 to 1800 UTC, a layer of fog lifted and merged with stratocumulus above. By 2100 UTC 8 May (Fig. 4c) the cold front was into southern Oklahoma.

The fundamental data produced by the MMCR are the backscattered power returns from each range sample volume. Power density spectra are computed from the power returns for each sample volume height, together with the first three spectral moments that correspond to the reflectivity (dBZ), Doppler velocity (m s^{-1}), and Doppler spectral width (m s^{-1}). An example of the reflectivities from the four MMCR modes on 7 and 8

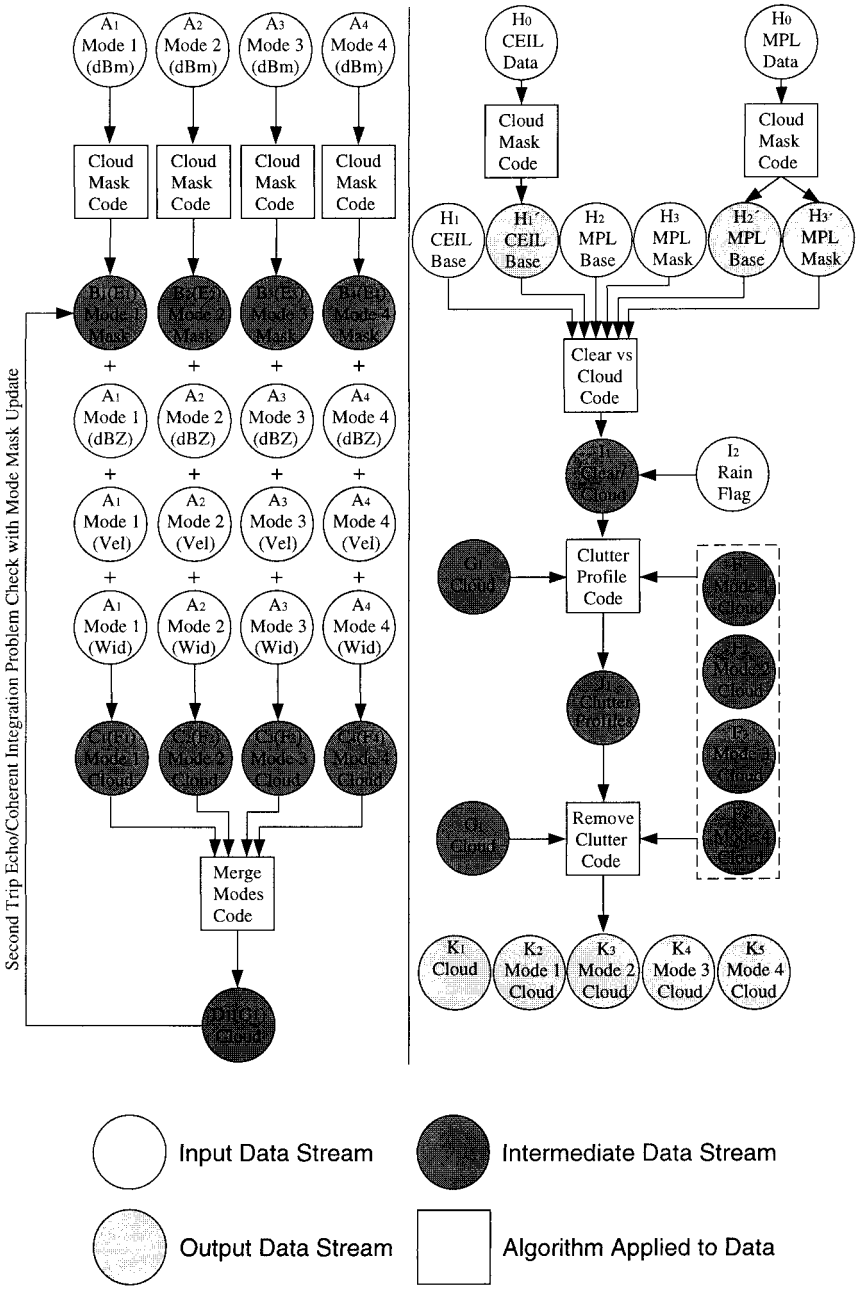

FIG. 2. Schematic diagram of the algorithm described in this paper. Unshaded circles represent the input data and lightly shaded circles represent the output data that are saved, whereas darkly shaded circles represent intermediate data products that are not saved. Squares represent algorithm processing steps that are applied to the data stream. The labels in parentheses, e.g., E_1 in the label $B_1(E_1)$, indicate datasets identical in form to the unparenthesized labels but with additional processing applied to them.

May 1997 is presented in Fig. 3A. After converting the backscattered power from units of dBm to mW, we apply an adaptation of the cloud detection algorithm developed by Clothiaux et al. (1995) (appendix A) to produce a mask of significant detections (Fig. 3B). A significant detection for a radar sample volume indicates that the power backscattered from it is distinguishable from the noise power produced by the radar receiver. The source of the significant backscattered power may be hydrometeors, or it may be airborne clutter, such as insects, seeds, and other nonhydrometeor particulates, collectively referred to as ''atmospheric plankton'' by Lhermitte (1966). In order to remove the effects of range sidelobes, we currently apply the threshold test de-

TABLE 1. Step-by-step description of the algorithm described in this paper.

Step	Processing/problems addressed
A → B	Search radar data for significant detections (modes 1–4)
	Eliminate data from lowest *N* gates (modes 1, 2), where *N* is the number of bits used in the pulse coding
	Eliminate data with range sidelobe contamination (modes 1, 2)
B → D	Merge data (modes 1–4) using specified criteria
D → E	Identify and flag aliased velocities in merged data
	Identify and flag second-trip echoes in merged data
E → G	Remerge data (modes 1–4) using only unflagged data
G → I	Use ceilometer/lidar data to establish better cloud-base heights
	Use ceilometer/lidar data to flag noncloud clutter such as insects
I → K	Produce final merged ''best'' cloud-boundary heights
	Produce final merged radar reflectivity, velocity, and width values
	Identify which of the four radar modes was retained in final data

scribed by Moran et al. (1998) to the data obtained from the pulse compression modes. The Doppler moments for the significant detections in each mode make up the cloud mode data represented by the circles labeled with a ''C'' in Fig. 2.

To merge the mode data a grid in space and time to which the mode data can be interpolated is required. To avoid complications involved in reducing the temporal resolution of the laser-derived cloud-base heights and to keep the radar data on a time grid close to a 9-s spacing, we choose a temporal resolution of 10 s. The vertical resolution of the grid is set to 45 m, as 45 m represents the spatial resolution of the mode with the highest resolution, and the lowest gridpoint height is 105 m, which is the first sample volume height for each radar mode. We linearly interpolate the radar reflectivity, Doppler velocity, and Doppler spectral width data for each mode to each time–height point on the grid and then apply a set of tests to the interpolated values to decide which mode set is used in the final result. We emphasize that at each grid point all three Doppler moments (i.e., reflectivity, velocity, and width) are derived from the same mode.

Our criteria for deciding which mode data are used at a grid point attempt to 1) eliminate data with coherent averaging or second-trip echo problems, 2) minimize the effects of pulse coding while keeping data with relatively large signal-to-noise ratios, 3) use data from the modes with comparable Doppler velocity resolutions, and 4) produce contiguous regions in the grid from a single mode. Therefore, we use the mode 3 data [see Clothiaux et al. (1999a) for the characteristics of each radar mode before and after 15 September 1997] at each grid point if the Doppler velocity inferred from the mode 4 data does not exceed the mode 3 Nyquist velocity and the signal-to-noise ratio of the mode 3 data is greater than 5 dB. The mode 4 velocity data are not used in this test if their signal-to-noise ratios are less than 10 dB. If the mode 4 velocity estimate does exceed the mode 3 Nyquist velocity and the mode 4 data signal-to-noise ratio is greater than 10 dB, we replace the mode 3 data at the grid point by the mode 4 data. If the signal-to-noise ratio of the mode 3 data is less than 5 dB, we use either mode 1 or mode 2 data, selecting the mode with the largest signal-to-noise ratio at grid points where the two modes overlap. In summary, we use the mode 3 data everywhere except in regions where there are strong radar return powers from fast-moving hydrometeors, in which case we use mode 4 data, and in regions where there are weak radar return powers, in which case we use either mode 1 or mode 2 data. The poorest velocity estimates are in the regions of weakest radar power returns, and we attempt to minimize the errors in the velocity estimates by using the data, that is, mode 1 and 2 data, with the best signal-to-noise ratios. The output of this data-merging process is illustrated in Fig. 3D.

Potential problems in the merged data may result from coherent averaging and second-trip echoes (Moran et al. 1998; Clothiaux et al. 1999a,b). To remove these problems from the merged data, as well as to assess the quality of the data from each individual radar mode, we use the reflectivity and Doppler velocity data illustrated in Fig. 3D to check the individual mode data for problems (appendix B). The results of the process are represented by the mode data in Fig. 3E. (Note that in Fig. 3E coherent averaging problems are indicated by the muddy red color, e.g., Fig. $3E_3$, whereas range aliasing, or second-trip echo, problems are represented by the light brown color, e.g., Fig. $3E_4$. Data with both kinds of problems are represented by magenta, as the small streaks embedded in the light brown areas of Fig. $3E_4$ illustrate.) As the data-quality flags generated by this process illustrate (Fig. 3E), mode 4 data in this instance suffer from second-trip echoes, as well as coherent averaging problems; the data-quality flags for SGP modes 1–3 (Figs. $3E_1$, $3E_2$, and $3E_3$, respectively) indicate coherent averaging problems at a range of heights. Using the quality flags to exclude mode data with second-trip echo or coherent averaging problems, we again merge the mode data to produce the dataset represented by Fig. $3G_1$.

On 15 September 1997 SGP mode 4 was slightly adjusted in order to remove all coherent averaging problems (Clothiaux et al. 1999a). For the adjusted mode 4 the coherent averaging problems evident in Fig. $3E_4$ would no longer be present. The adjustment to mode 4 also raised its maximum height to a value comparable with modes 2 and 3. Consequently, after 15 September 1997 there are few instances of coherent averaging problems in the final merged data.

To aid in the identification of hydrometeor contributions to the final merged data we generate a best-estimate cloud-base height from the ceilometer and lidar

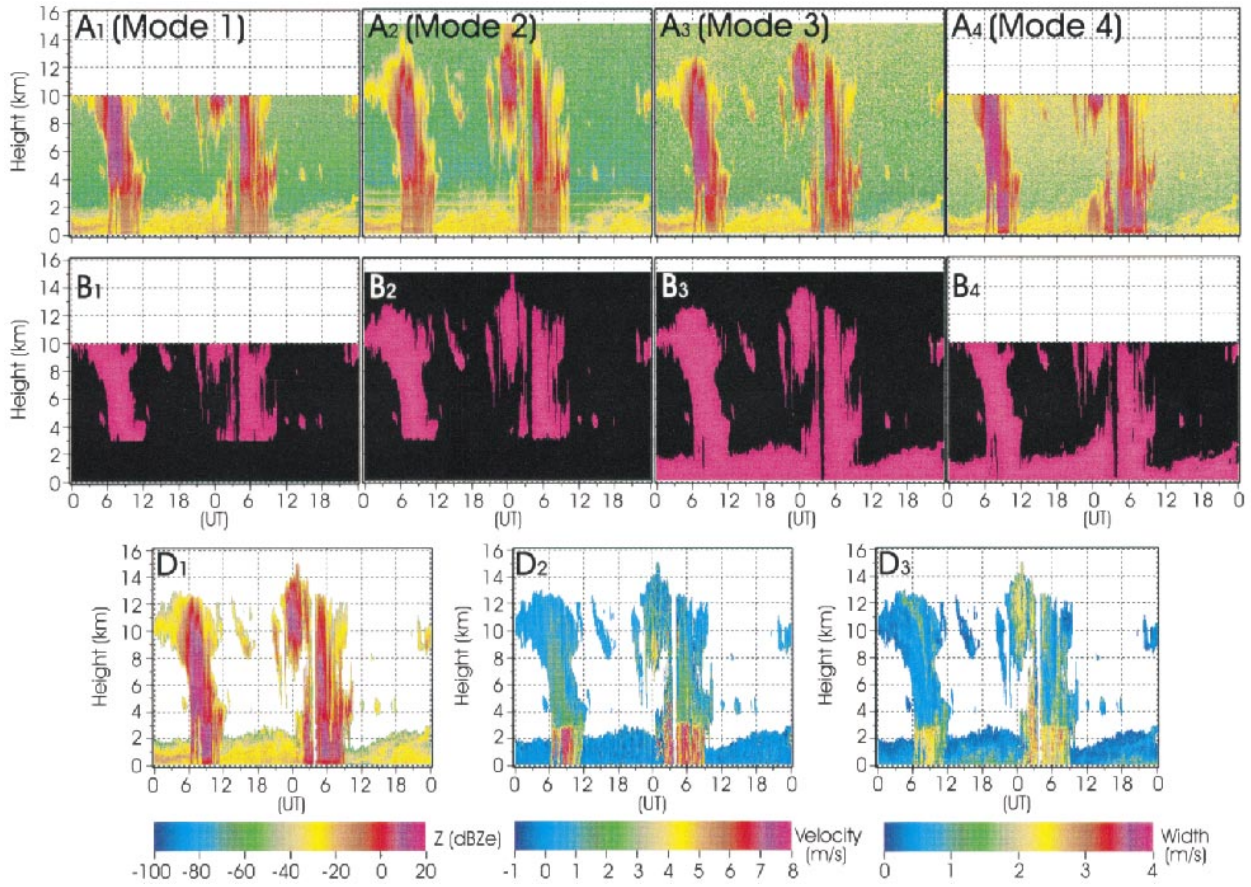

FIG. 3. Visual representation of the data products at each stage of processing illustrated in Fig. 2. (A) The input reflectivity data for each mode; (B) the initial mask results illustrating the significant radar return powers; (D) the initial merge results of radar reflectivity (D_1), Doppler velocity (D_2), and Doppler width (D_3); (E) the final mask results with labels identifying coherent integration problems (red), second-trip echo problems (light brown), and both coherent and second-trip echo problems (magenta); (G) the final merged radar reflectivity data (G_1), together with the laser-derived cloud-base heights (black dots), with a blowup of the clutter in the lower radar sample volumes (G_2); (H) the MPL photoelectron count data on a logarithmic scale (H_1), the mask of significant MPL detections (H_2), where light brown indicates a significant detection that is not labeled as cloud, and magenta indicates a significant detection that is labeled as cloud, and a time–height

data. To this end the micropulse lidar photon count data (Fig. $3H_1$) are processed for significant detections (Fig. $3H_2$) using the algorithms described in Clothiaux et al. (1998) and one recently developed by Campbell et al. (1998). One algorithm described by Clothiaux et al. (1998) is called the Scott and Spinhirne, or SS, algorithm after its developers, while a second algorithm described by Clothiaux et al. (1998) is called the CA algorithm. (Note that an additional test now included in the CA algorithm is the one developed by Pal et al. 1992.) We subsequently combine these cloud mask data with the cloud-base heights that are generated by the Belfort laser ceilometer system (Fig. $3H_3$). The laser-derived cloud-base heights are then analyzed (appendix C) to produce a best-estimate cloud-base height (Fig. $3I_1$). In the final cloud-base height estimates the Belfort ceilometer data provide the largest contribution below 3 km, whereas the results of either the Campbell et al. (1998) or SS micropulse lidar algorithm provide the largest contribution above 3 km. If the rain flag (Fig. $3I_2$) indicates rain, and all of the laser-dependent retrievals fail, the cloud-base height is set to the surface. In the application of the algorithm and throughout the following discussion, we assume that the ceilometer and micropulse lidar are able to accurately detect the cloud base. However, this assumption becomes increasingly poor as the number of precipitation particles near and below cloud base increases. This point must be kept in mind when interpreting these data.

The final laser-derived cloud bases for 7 and 8 May 1997 are illustrated in relation to the significant radar reflectivities in Fig. $3G_2$. From 0000 to 0600 UTC and 2100 UTC 7 May to 0000 8 May there are significant radar reflectivities below 2 km with no corresponding laser-dependent cloud detections in this region. Furthermore, from 1500 to 2100 UTC on 7 May and 1800 UTC 8 May to 0000 9 May there are strong radar reflectivities below the laser-derived cloud-base heights. These four periods provide examples of atmospheric clutter contributions to the radar reflectivities and we

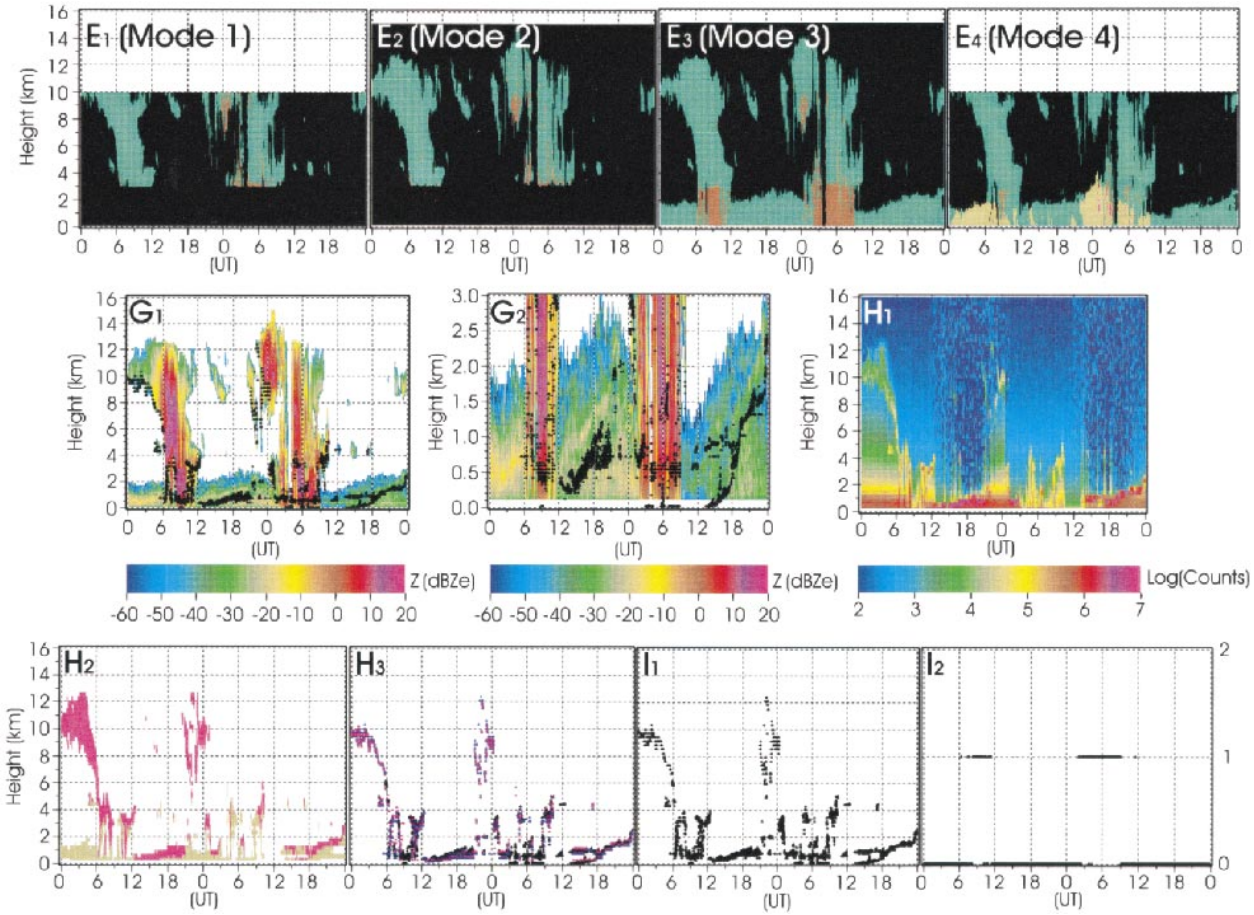

FIG. 3. (*Continued*) plot of all of the laser-derived cloud-base heights (H_3), where the BLC heights are in black, the SS algorithm MPL heights (offset by −150 m) are in blue, and the CA algorithm MPL heights are in magenta; (I) the laser-derived, best-estimate cloud-base heights (I_1) generated from the different heights illustrated in (H_3) together with the rain flag (I_2), where a 1 indicates precipitation at the surface; and (K) the final set of cloud location products output by the algorithm, including the reflectivity with no clutter (K_1), the best-estimate reflectivity field indicating the location of hydrometeors (K_2), flags identifying the mode contributing to the final data product (K_3), flags identifying uncontaminated returns (1), pure clutter returns (3), and a mixture of cloud and clutter returns (2; K_4), the Doppler velocities (K_5), and widths (K_6), for all significant returns, and the signal-to-noise ratio (K_7) for all significant returns.

must attempt to label them as such. To identify the clutter in the radar returns we assume that the laser-dependent cloud detections are accurate. Consequently, all significant radar returns that occur during a period when the lasers fail to detect a cloud are classified as clutter. Furthermore, all significant radar returns that are below the laser-derived cloud-base heights and that do not extend continuously from the surface up to cloud base are classified as clutter. An example of such clutter reflectivities occurs from 0000 to 0300 UTC 7 May 1997 below 3 km. Periods when the significant radar reflectivities extend from cloud base to the surface can occur during precipitation (e.g., 0600–1200 UTC 7 May and 0300–0900 UTC 8 May) or significant amounts of clutter (e.g., 1500–2100 UTC 7 May and 1800 UTC 8 May–0000 UTC 9 May). Below the laser-derived cloud-base heights we attempt to distinguish precipitation from clutter in each radar profile using ''clear-sky'' radar profiles generated from periods such as 0000–0300 UTC 7 May.

To characterize the clutter we use a running 20-min time window to search the combined radar reflectivity–laser-derived cloud-height data (Fig. $3G_2$) for periods when the laser cloud-base height estimates lie completely above the surface layer of significant radar reflectivities. When such a period is identified, as from 0000 to 0020 UTC 7 May, we compute the maximum value of reflectivity across the 20-min interval for each range gate and store the results as a clutter profile (circle labeled J_1 in Fig. 2) with a time stamp at the center of the interval, for example, 0010 UTC 7 May for the 0000 to 0020 UTC 7 May ''clear-sky'' interval. The search for the next profile commences at the end of the interval. For the period of data illustrated in Fig. $3G_2$ 16 clutter profiles were created at 20-min intervals from 0010 to 0510 UTC 7 May and clutter profiles were also gen-

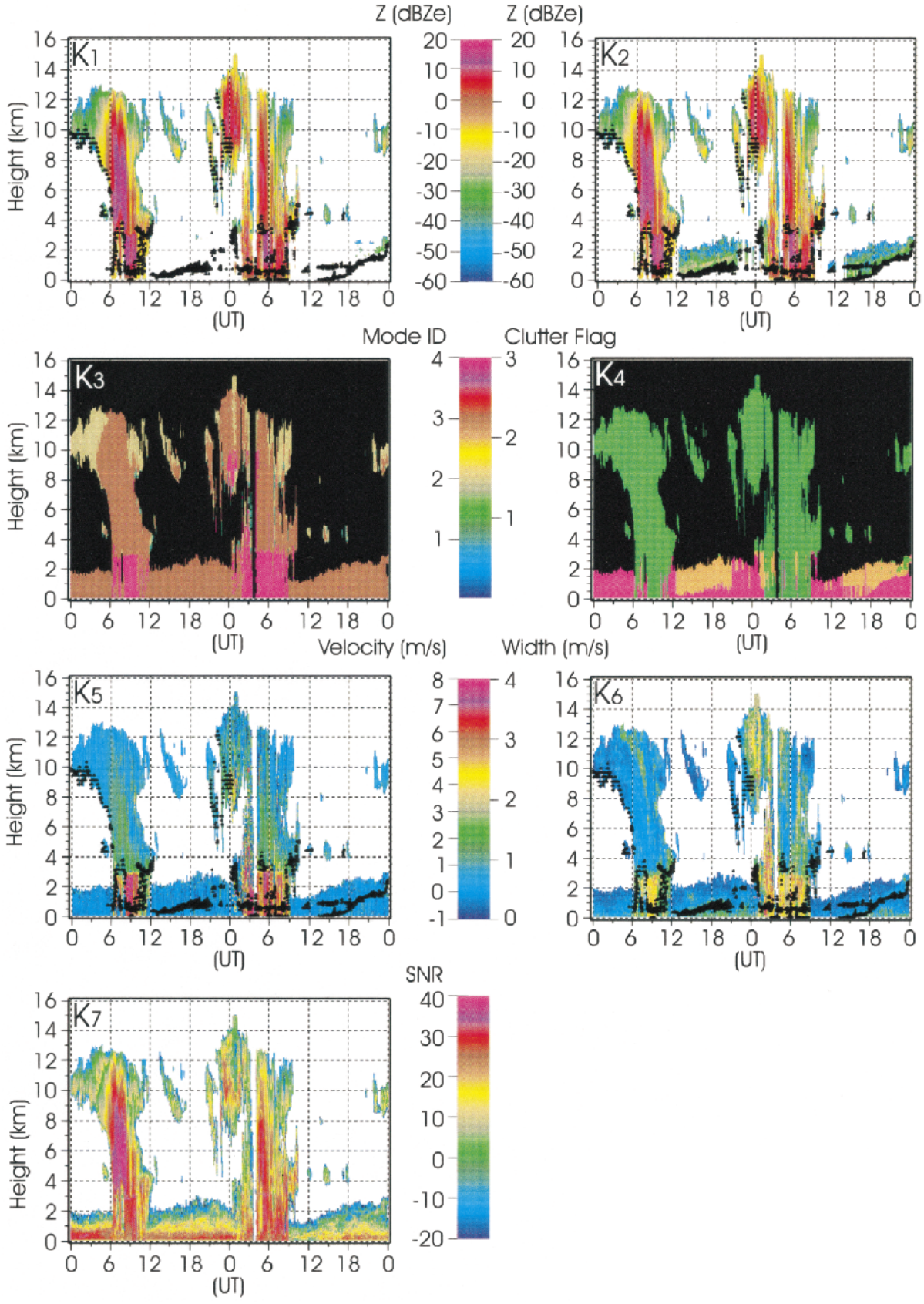

FIG. 3. (*Continued*)

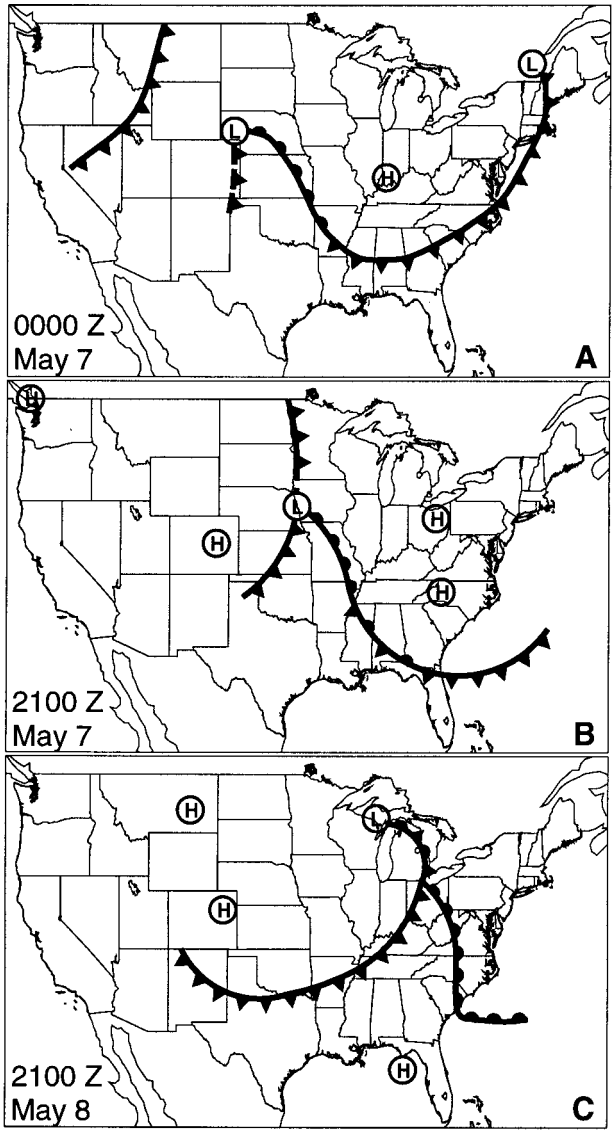

FIG. 4. Illustration of the synoptic situation on 7 and 8 May 1997.

erated at 2144, 2204, 2244, 2305, 2325, and 2345 UTC 7 May, as well as 1035, 1246, and 1306 UTC 8 May. Our procedure for identifying the clutter is to compare each profile of significant radar return with the two clutter profiles that surround it in time. If the radar reflectivity for a radar sample volume is less than either of the two clutter profile reflectivities for the same sample volume, the reflectivity is classified as clutter. In Fig. $3K_1$ all of the radar reflectivities identified as clutter-contaminated are removed, leaving those reflectivities that result primarily from hydrometeors with negligible clutter contamination.

The results illustrated in Fig. $3K_1$ make up one part of the final set of cloud-height products generated by the current algorithm. Unfortunately, from 1200 UTC 7 May to 0000 UTC 8 May and 1100 to 2100 UTC 8 May in Fig. $3K_1$ the laser-dependent cloud detections, which are included in the final data product, have no corresponding uncontaminated radar reflectivities. In an attempt to incorporate some information on the cloud-top heights during these periods the contaminated radar reflectivities that occur above a laser-dependent cloud detection and that are continuous in height with it are reinserted into the dataset (Fig. $3K_2$). These reflectivities are subsequently labeled as clutter contaminated (Fig. $3K_4$). The top of the contaminated layer (Fig. $3K_4$) serves as an upper bound on the cloud-top height.

To illustrate which radar modes are contributing to the final set of radar data we incorporate flags into the dataset that identify the mode from which the data are derived (Fig. $3K_3$). As expected, the general mode (mode 3) provides the dominant contributions to the final cloud-height data product, with the cirrus mode (mode 2) providing significant contributions above 3 km and the robust mode (mode 4) contributing data when the particle fall speeds are relatively large, such as during precipitation events. The mean Doppler velocities (Fig. $3K_5$) and widths (Fig. $3K_6$) are provided for each radar sample volume for which there is a significant radar return; the data fields illustrated in Fig. $3K_1$ and Fig. $3K_2$, together with the radar reflectivity field illustrated in Fig. $3G_1$, which is also included in the cloud-location data product, can be used to interpret the significance of the velocity and width data. Since the Doppler velocity and width data become more reliable as the signal-to-noise ratios of the data increase, we include the signal-to-noise ratio of each significant detection into the final data product (Fig. $3K_7$). These data products, considered together, represent our best determination of the vertical distribution of hydrometeors above the ARM SGP site.

3. Results

In the current study we attempt to ascertain the vertical distribution of hydrometeors in the column above the ARM SGP site. To assess the performance of the current detection approach we first compare the cloud-base heights reported by the Belfort ceilometer with the cloud-base heights generated by applying the SS and CA algorithms to the micropulse lidar data. We then compare the laser-derived, best-estimate cloud-base heights with the corresponding profiles of significant radar returns to assess the performance of the millimeter-wave cloud radar in detecting all hydrometeors. In this second comparison the important assumption is that the lasers always accurately detect the bottom of the lowest cloud layer, although problems are known to exist during periods of precipitation. In the laser cloud-base height comparisons we use data that were collected at the ARM SGP site from 7 November 1996 to 30 June 1998. Whenever radar data are used in an analysis we restrict the data to the time period 1 October 1997–30

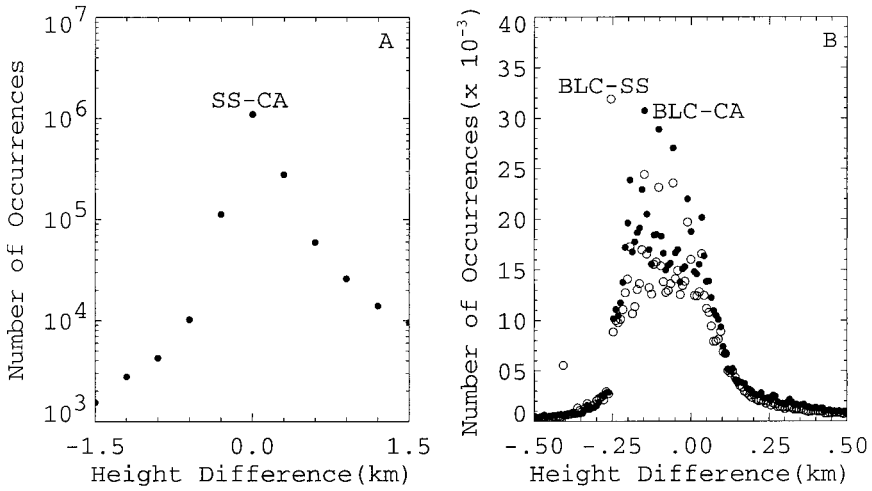

FIG. 5. Number-of-occurrence histograms of the cloud-base height differences between the various laser-dependent systems. (a) The SS algorithm cloud-base heights minus the CA algorithm cloud-base heights when both algorithms are applied to micropulse lidar data; the bin size is 300 m. (b) The Belfort ceilometer cloud-base heights minus the cloud-base heights obtained by applying the SS (open circles) and CA (solid circles) algorithms to the micropulse lidar data; the bin size is 7.4 m.

June 1998, because the current set of operational modes were implemented on 15 September 1997.

Before discussing an analysis of the hydrometeor properties derived by the lasers and radar we first address the reliability of these instruments. This is an important issue, as these types of instruments have only recently been developed and they have never before been operated for a long period of time in an automated fashion. The period from 7 November 1996 to 30 June 1998 consists of 601 days. For each day there are 8640 profiles, so that across the whole dataset there are 5 192 640 possible sample profiles. Out of this pool of profiles the Belfort ceilometer, micropulse lidar, and millimeter-wave radar produced data 93.8%, 94.9%, and 96.5% of the time, respectively. Inspection of the data indicates that dropouts for each instrument generally occur in contiguous blocks of time, which could be expected to result from either instrument problems or site-wide power losses. Inspection of the data indicates that the former problem, that is, instrument problems, accounts for the majority of missing data. The large percentage of reported data for all instruments indicates that the current dataset is fairly complete across the whole time period.

To compare the laser-derived cloud-base heights for the time period of 7 November 1996–30 June 1998 we subtract the height reported by one algorithm, or system, from the heights reported by each of the other two algorithms on a profile-by-profile basis. We then bin the height differences to create number-of-occurrence histograms (Fig. 5). The two algorithms applied to the micropulse lidar data produce identical heights in most cases (Fig. 5a), and the average height difference between them is 106 m with the SS alogrithm heights being higher on average. Relative to the SS algorithm the numbers of false positives and negatives produced by the CA algorithm are 18% and 2%, respectively. With respect to the Belfort laser ceilometer heights the SS and CA algorithm cloud-base height estimates are higher on average by 65 and 69 m, respectively (Fig. 5b). The Belfort ceilometer system fails to detect clouds at heights over approximately 5 km, while its higher temporal and spatial resolution lead to substantially more detections below 5 km.

These results are nearly identical to those obtained by Clothiaux et al. (1998), who used limited data from the autumn of 1995. The Clothiaux et al. (1998) results suggest that by doubling the threshold used to identify cloud in the CA algorithm from an optical depth of 0.03 to 0.06 per 300-m-resolution volume, the bias between the SS and CA algorithms can be substantially reduced. These earlier results also suggest that the micropulse lidar cloud-base height estimates produced by the CA algorithm would then be higher than the Belfort ceilometer heights by as much as 170 m. At the time of this earlier study the source of this bias between the Belfort ceilometer system and CA algorithm was unknown. Recent analyses of the Belfort ceilometer backscatter profiles have indicated a timing offset in the Belfort profiles that amounts to a height offset of 193 m (C. J. Flynn 1998, personal communication). By adding 193 m to the Belfort ceilometer cloud-base heights the bias between the Belfort ceilometer and CA algorithm cloud-base heights would then be reduced from 170 m to approximately 23 m with the Belfort ceilometer heights being higher. By doubling the threshold in the earlier study the bias between the SS and CA algorithms was reduced by 106 m. If such a reduction

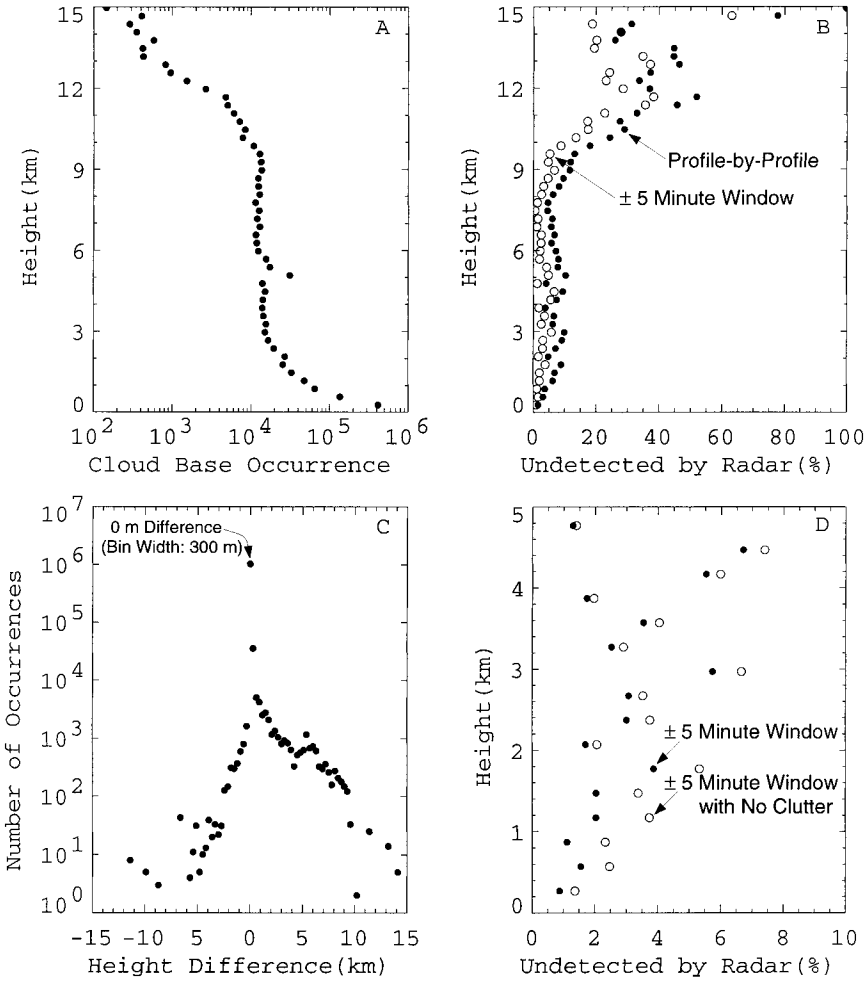

FIG. 6. (a) Number of occurrences of the laser-derived, best-estimate cloud-base heights. The bump just above 5 km is a result of a change in the cloud detection threshold of the CA algorithm at 5 km. (b) Percentage of the laser-derived cloud-base heights that go undetected by the radar on a profile-by-profile basis (solid circles). The percentage of cloud detections missed by the radar decreases if all of the radar profiles within 5 min of the laser profile are incorporated into the analysis (open circles). (c) The number of occurrences of the height difference between the laser-derived cloud-base heights and the nearest significant radar detection; the bin size is 300 m. (d) The solid circles are the same as the open circles in (b) and the open circles represent the percentage of undetected clouds by the radar for the population of profiles when the laser-derived cloud-base heights are clearly above the clutter in the radar returns.

were to hold in the current dataset as well, the bias between the two algorithms in this study would be eliminated. These results suggest that by implementing the correction to the Belfort ceilometer heights and reprocessing the micropulse lidar data with the CA algorithm threshold set to a cloud optical depth of 0.06 $(300\ \text{m})^{-1}$ the cloud-base heights generated by the Belfort system and the CA algorithm can be brought to within 20–30 m of each other on average. Furthermore, the SS and CA algorithm cloud-base height bias would be eliminated.

The best-estimate cloud-base height, which is based on an analysis of the Belfort ceilometer and micropulse lidar data, contained missing data only 1.6% of the time. Of the remaining profiles (98.4%) for which there are data, a cloud was detected 56.0% of the time. The number of occurrences of the laser-derived cloud-base detections as a function of height is illustrated in Fig. 6a. For each laser-derived cloud-base height obtained during the time period from 1 October 1997 to 30 June 1998 we computed the vertical distance to the nearest MMCR significant detection. If there was no significant MMCR return, we categorized this profile as one for which the MMCR failed to detect a cloud and we counted the number of times this occurred as a function of the laser-derived cloud-base height (Fig. 6b). Relative to the laser-derived cloud detections, the MMCR failed to detect a hydrometeor 5.9% of the time. For those

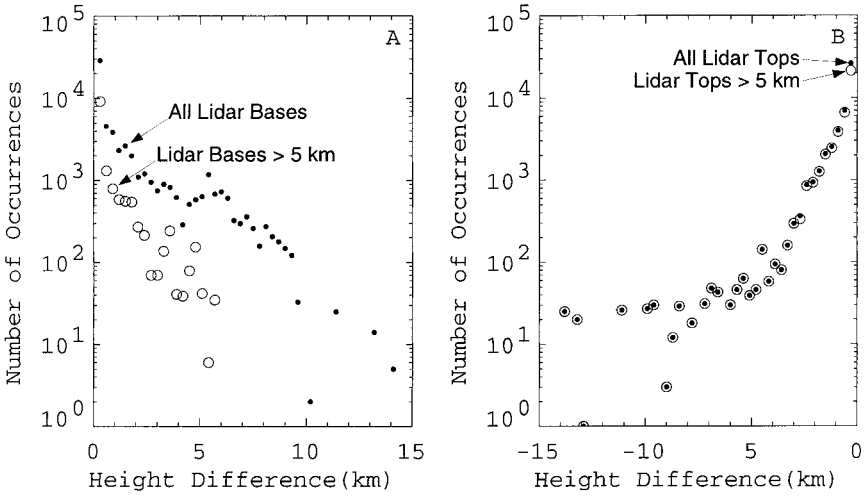

FIG. 7. (a) Number of occurrences of the difference in height of the lowest-altitude MMCR significant detection minus the laser-derived, best-estimate cloud-base height for those times when the lowest-altitude MMCR detection is above the laser-derived cloud-base height. (b) Number of occurrences of the difference in height of the highest-altitude MMCR significant detection minus the highest-altitude micropulse lidar significant detection for those times when the highest-altitude MMCR detection is below the highest-altitude micropulse lidar detection.

profiles where there is a significant MMCR detection we binned the height difference between the laser-derived, best-estimate, cloud-base height and the nearest MMCR detection to form number-of-occurrence histograms of the height differences (Fig. 6c). As Fig. 6c illustrates, when the MMCR detects a hydrometeor in a profile for which there is a laser-dependent cloud detection, the difference in height is generally close to 0 m. To support our subjective impression that the MMCR generally detected some part of each cloud element detected by the lasers we again computed the vertical height difference between each laser-dependent cloud detection and the nearest MMCR significant detection, but this time we incorporated all of the MMCR returns within 5 min of the laser-dependent cloud detection in the search for the nearest height. (We chose a 5-min interval as a trade-off between staying close in time to the current laser-derived detection and incorporating more MMCR time records into the comparison.) In this second comparison the number of cloud detections missed by the MMCR dropped to 2.9%.

As Fig. 6b illustrates, the MMCR detected most clouds below 1 km and between 6 and 8 km. The most difficult clouds to detect occurred between 1 and 6 km and above 8 km. The clouds between 1 and 6 km not detected by the MMCR are most likely thin stratus, thin altostratus, patchy fair-weather cumulus, or altocumulus. These clouds may consist of extremely small drops that are weakly reflecting at millimeter-wavelengths, or the clouds may be thinner than the radar sample volume depth resulting in partial beam filling and reduced reflectivity. As the height of the laser-derived cloud detection increases from 8 to 15 km, the percentage of thin cirrus missed by the MMCR increases; however, above 12 km the absolute number of missed detections is relatively small, because the number of laser-derived cloud detections at these altitudes is small.

The performance of the MMCR in detecting boundary layer clouds is actually slightly worse on a percentage basis than Fig. 6b indicates. This arises from the prevalence of clutter in the lower atmosphere that can lead to significant MMCR detections in the vicinity of a boundary layer cloud, which by itself, might go undetected. To illustrate this point we again compared the laser-derived cloud detections with the MMCR returns in a 5-min interval surrounding the laser detection, this time using only those laser-derived cloud-base heights that are completely above the radar clutter. As Fig. 6d shows, the percentage of low-level clouds missed by the MMCR goes up as clutter-contaminated MMCR detections in the vicinity of the laser detection are eliminated from the comparison.

In comparisons between lidars and radars important questions are how well does the radar detect the bases of all clouds and the tops of optically thin clouds and how often is the lidar beam completely attenuated before reaching cloud top. To address the first question we formed two histograms. In the first histogram (Fig. 7a) we show the number of occurrences of the difference in height between the lowest-altitude MMCR significant detection and the laser-derived best-estimate cloud-base height for those times when the lowest-altitude MMCR detection was above the laser-derived cloud-base height. The second histogram (Fig. 7b) is of the number of occurrences of the difference in height between the highest-altitude MMCR significant detection and the highest-altitude micropulse lidar significant detection for those times when the highest-altitude MMCR de-

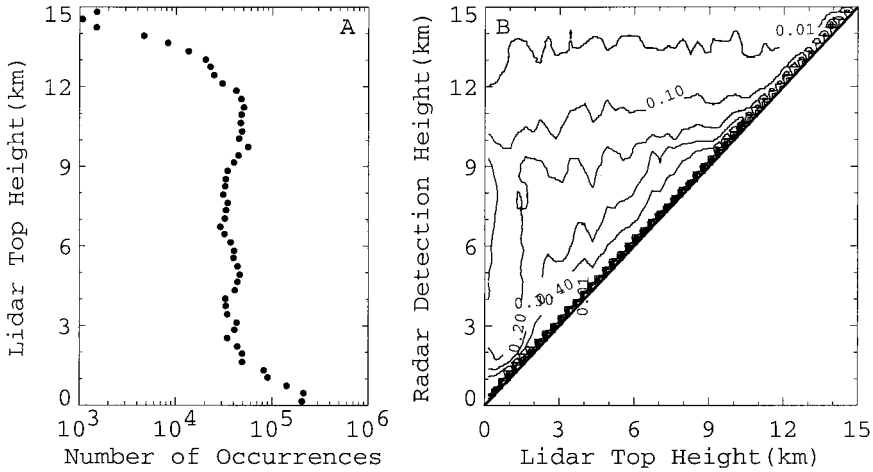

FIG. 8. (a) Number of occurrences of the highest-altitude significant detection by the micropulse lidar for each 10 s interval. (b) Frequency of occurrence of significant hydrometeor detections by the MMCR that occur above the micropulse lidar–top heights illustrated in (a).

tection was below the highest-altitude micropulse lidar detection. Only data from altitudes above the atmospheric clutter were used in the analysis. Of the 869 844 laser-derived cloud-base detections clearly above the radar clutter the MMCR failed to detect a hydrometeor in the column in 7.9% of the cases. The total number of samples in Fig. 7a is 58 344, which represents 6.7% of the total number of laser cloud-base detections. Of the 750 507 laser-derived cloud-top detections the MMCR failed to detect a hydrometeor in the column in 7.5% of the cases. The total number of samples in Fig. 7b is 46 514, which represents 6.2% of the total number of lidar-derived cloud-top detections. Separating the laser-derived bases and tops above 5 km from the total population (Fig. 7) indicates that the MMCR more often fails to detect those hydrometeors near the base of low-altitude clouds, as opposed to high-altitude clouds, while almost always having a detection in the vicinity of the lidar tops for the low-level clouds that the radar can detect. These results may be due in part to complete attenuation of the laser beams in thick low-level clouds.

To illustrate those hydrometeors that go undetected by the lasers above optically thick lower clouds consider Fig. 8. In Fig. 8a we plot the number of occurrences of the highest-altitude significant detection by the micropulse lidar for each 10-s interval over the period from 7 November 1996 to 30 June 1998. (If the micropulse lidar beam was never completely attenuated, this histogram would represent the vertical distribution of cloud-top heights.) For each 10-s time interval represented by a height in Fig. 8a we counted the number of MMCR range gates above this height that contained a significant hydrometeor detection to produce the frequency-of-occurrence contour plot illustrated in Fig. 8b. Two particularly interesting features present in Fig. 8b are the dip in the 20% contour for lidar-top heights near 1 km and the lack of radar detections above the lidar at altitudes of 14 km. The dip in the 20% contour is due to convective periods when clouds and precipitation reach from near the surface to 9–10 km of altitude. The second result, considered together with the results illustrated in Fig. 6, indicates that the micropulse lidar, as compared with the MMCR, detects cloud particles to higher altitudes during periods of isolated cirrus. Overall, however, Fig. 8b illustrates that it is not uncommon for the lidar beam to be completely attenuated before reaching cloud top.

The discussion in the above two paragraphs raises a number of issues in the interpretation of laser and radar backscatter data. For example, we call the lowest-altitude laser-derived detection cloud base although this detection is not always at cloud base. If the density of precipitation particles is sufficiently high below cloud base, the algorithms implemented on the Belfort and micropulse lidar systems will produce cloud-base heights that are too low in altitude. Also, in the comparison of the MMCR detections with the laser-derived cloud-base heights the effects of precipitation were ignored. If the cloud particles at the bases of the clouds were considered in isolation from any precipitation particles that might be falling through them, the radar might not be able to detect them. The results illustrated in Fig. 7a are for clouds that are not drizzling, and we did not attempt to ascertain the number of significant radar detections at the laser-derived cloud bases that resulted from precipitating, as opposed to cloud, particles.

The problems associated with identifying the heights of cloud tops from surface-based lasers and radars are just as significant. The only unambiguous method for identifying cloud top from the ground is the presence of molecular, or Rayleigh, backscatter in the lidar returns above the highest detected cloud particles. If the particles at cloud top are sufficiently small, a radar will not be able to detect them. Therefore, if the radar detections

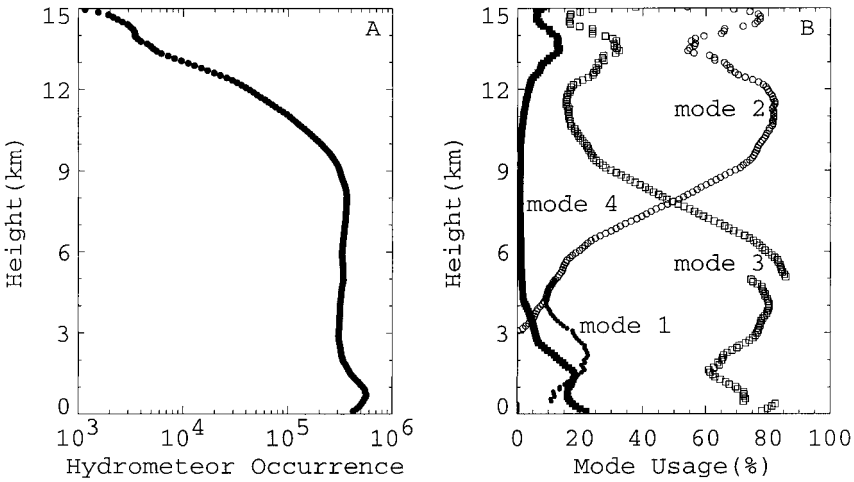

FIG. 9. The number of occurrences of radar detections of hydrometeors or mixtures of hydrometeors and clutter as a function of height. (b) The percentage contributions of mode 1 (solid circles), mode 2 (open circles), mode 3 (open squares), and mode 4 (solid squares) data to the merged cloud mask.

go to a higher altitude than the laser detections, one knows that the laser beam has completely attenuated and the true cloud top is either at or above the highest radar detection. The current comparisons of the MMCR and micropulse lidar detections indicate that the MMCR is able to detect most of the particles at the highest altitudes detected by the lidar. With the deployment of the high-resolution lidars at the ARM sites, together with improvement of the algorithms that are applied to their data, identification of the times when the lidar beam reaches cloud top will be possible and analyses of the MMCR data at these times will better characterize the capability of the MMCRs in detecting the cloud tops present over the ARM sites.

For the period of 7 November 1996–30 June 1998 there are 601 × 8640 × 353 range gate samples in the final merged product between altitudes of 105 and 16 000 m. Of these range gate samples, only 10.4% contained significant power returns due either to hydrometeors or to a mixture of hydrometeors and clutter (i.e., flag values 1 and 2 in Fig. $3K_4$). The vertical distribution of hydrometeor and mixed hydrometeor and clutter detections for the time period 1 October 1997–30 June 1998 is illustrated in Fig. 9a, and the contribution of each mode to the significant detections at each height is illustrated in Fig. 9b. As Fig. 9a illustrates, low-level clouds, combined with precipitation events that reached the surface, occurred the most frequently. In the current scheme for combining the mode data each of the three modes made contributions to the low-level cloud detections. The mode 4 data were used during the precipitation events, the mode 1 data detected the weaker reflecting clouds between 1 and 3 km, and the mode 3 data were used for the remaining, moderately reflecting hydrometeors. The emphasis on using the mode 3 data is evident, as this mode accounts for 60%–80% of all detections from the surface up to 7 km. As expected, in the upper troposphere between 9 and 15 km, the cirrus mode data (from mode 2, which is the most sensitive of the four modes) contributes the most detections since the majority of detections in this region are from weakly reflecting, thin cirrus. However, the occurrence of deep convection in Oklahoma also leads to a local maximum in the percent usage of mode 4 between 12 and 15 km. This last result attests to the importance of having a mode with a wide Nyquist velocity and no coherent integration that covers these high altitudes.

To assess the importance of the enhanced sensitivity of modes 1 and 2 (the pulse compression modes) relative to modes 3 and 4 we compared the hydrometeor detections produced by modes 3 and 4 with the corresponding hydrometeor detections of the final merged cloud product. For this exercise we used the returns from 1 October 1997 to 30 June 1998 that were classified as originating from clutter-free hydrometeors (i.e., flag value of 1 in Fig. $3K_4$). We interpolated the mode 3 and 4 data to the merged time–height grid and counted, as a function of height, the number of times there was a significant detection in the merged mask with no corresponding detection in the mode 3 and 4 data (Fig. 10). Overall, modes 3 and 4 failed to detect 14.3% and 26.7%, respectively, of the hydrometeors that are present in the merged mask. The vertical distributions of the missed hydrometeor detections (Fig. 10b) illustrate that, on a percentage basis, the hydrometeors below approximately 3 km and above 9 km create the largest problems for these two less sensitive modes. (Note that the discontinuity at 5.5 km results from more stringent tests of clutter occurrence in the mode data. Consequently, more significant detections are classified as clutter in the mode data as compared to the merged data, increasing the number of missed detections in the mode data below

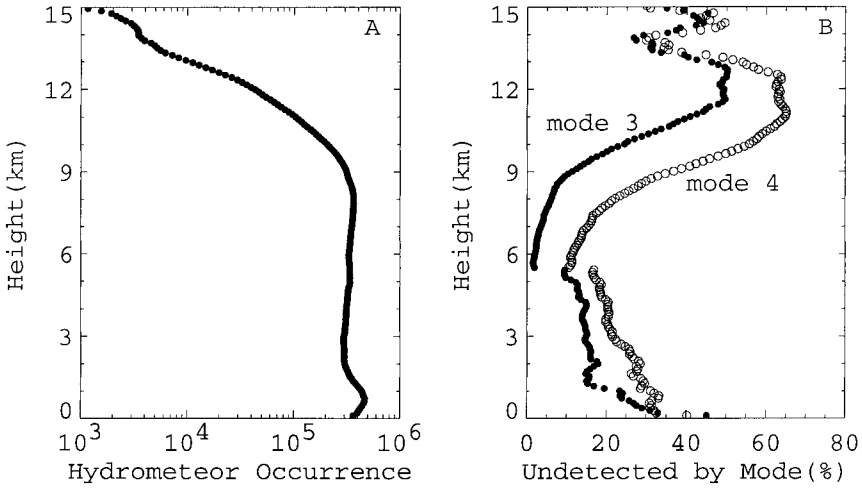

FIG. 10. (a) The number of occurrences of radar detections of hydrometeors with insignificant clutter contamination as a function of height. (b) The percentage of the detections illustrated in (a) that go undetected by mode 3 (solid circles) and mode 4 (open circles). The discontinuity of the data at 5.5 km is due to slightly different treatments of the clutter in the merged and mode data.

5.5 km.) These findings, together with the results illustrated in Fig. 6, indicate that high radar sensitivity, in this case provided by the pulse compression technique used in modes 1 and 2, is critical to developing a comprehensively accurate depiction of the hydrometeors in a vertical column of the atmosphere.

The four different modes currently in operation are a complementary set. Furthermore, each mode is calibrated independently of the other modes in the set. To quantify the impact of merging the mode data into a single reflectivity product we compared the reflectivities in the final merged product with coincident mode 4 cloud detections. Since mode 4 does not use pulse coding or coherent averaging, the moments data from this mode should be accurate when there is a significant detection. We interpolated the mode 4 data to the merged time-height grid for the period from 1 October 1997 to 30 June 1998. For each grid point, if there is a significant mode 4 detection, we bin the mode 4 reflectivity together with the corresponding merged product reflectivity to form a two-dimensional number-of-occurrence histogram. The data in the histogram are actually segregated by the mode identification of the merged reflectivity, thereby leading to three distinct histograms (Fig. 11a). We form two one-dimensional number-of-occurrence histograms for each two-dimensional histogram by summing all of the merged reflectivities within each 1-dB wide mode 4 reflectivity bin and vice versa (Fig. 11b).

The mean values of the mode 4 reflectivities in Figs. $11b_1$ (mode 1), $11b_2$ (mode 2), and $11b_3$ (mode 3) are -28.3, -21.6, and 0.4 dBZ, respectively. The corresponding mean values of the mode 1, mode 2, and mode 3 reflectivities are -25.6, -21.7, and -0.1 dBZ, respectively. Note that the low mean values for the mode 1 and 2 data are expected since the data from these modes are primarily from weakly reflecting clouds. The mean values for the mode 2 and mode 3 results differ by 0.1 dB and 0.5 dB, respectively, from the mode 4 values with the mode 4 reflectivity means being slightly larger. However, the mean for the mode 1 analysis differs by 2.7 dB from the mode 4 value with the merged mode 1 mean being higher. A direct comparison (i.e., without interpolation) of the mode 1 and mode 4 reflectivities for one day with a mixture of different cloud types did not exhibit such a difference in the mean reflectivity when only reflectivities with low signal-to-noise ratio were incorporated into the means. Rather, the mean differed on the order of the mode 2 and mode 3 results. This led us to inspect the interpolation routine and we discovered an error in it that leads to a factor of two increase in the interpolated mode 1 reflectivities for the day under investigation. (This error will be corrected during the first reprocessing of the data when calibration corrections will also be applied to the data.) The impact of the interpolation error on the 90-m-resolution mode 2 and mode 3 data for this day appeared to be minimal. Subtracting 3 dB from the mode 1 mean reflectivity value generated from the data illustrated in Figs. 11a and 11b, we find that on average all of the merged reflectivity data is within 0.5 dB of the mode 4 data. The slightly smaller values of the mode 1, 2, and 3 mean reflectivities as compared to the mode 4 value most likely results from coherent integration, as we now describe.

Although on average the mode 4 reflectivity is approximately the same as the average reflectivity of each mode within the merged cloud product, errors are apparent in the contour plots of Fig. 11a, especially for the mode 2 data. As Fig. $11a_2$ illustrates, the merged

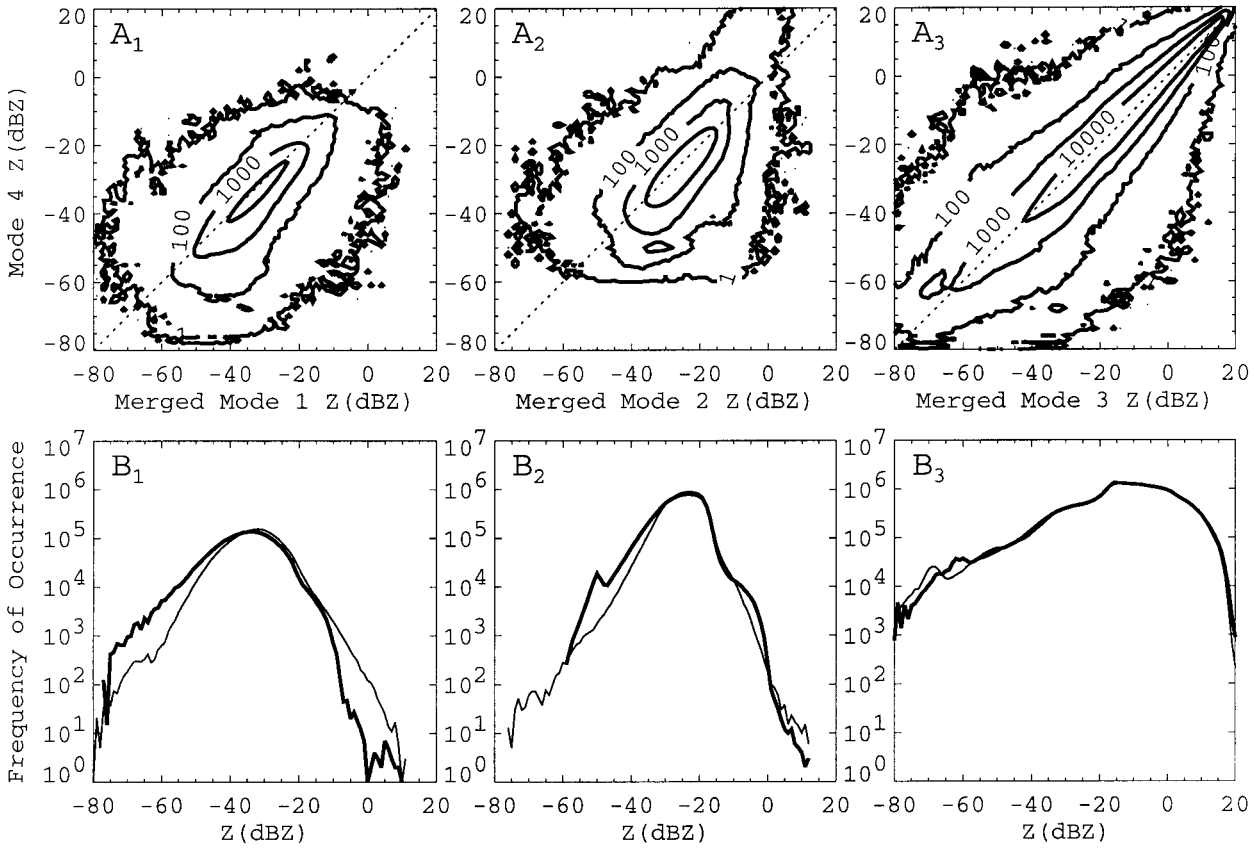

FIG. 11. (a) Contour plots of the two-dimensional number-of-occurrence histograms of coincident mode 4 reflectivities and merged mode 1 (a_1), merged mode 2 (a_2), and merged mode 3 (a_3) reflectivities. (b) Number-of-occurrence histograms of the mode 4 (thick solid lines), merged mode 1 (thin solid line, b_1), merged mode 2 (thin solid line, b_2), and merged mode 3 (thin solid line, b_3) reflectivities obtained by row and column summing the histograms in (a).

mode 2 reflectivities are biased high compared to the mode 4 reflectivities below approximately −40 dB*Z* and they are biased low for mode 4 reflectivities above approximately −20 dB*Z*. These biases at low and high values of the mode 4 reflectivities are the result of pulse coding and coherent integration, respectively. Since the data used in the comparison include only times and altitudes for which mode 4 detected the hydrometeors, we know that the signal returns are significant. Hence for the mode 2 range gates we can expect on occasion to have bleeding of the power returns from nearby sample volumes into the one under consideration (range sidelobe contamination), leading to a slight enhancement of its power. This would lead to the high bias of the mode 2 reflectivities relative to the mode 4 reflectivities at low values of the mode 4 reflectivities, as is evident in Fig. 11a_2. For values of the mode 4 reflectivities above −20 dB*Z* the hydrometeor velocities may begin to be large enough to exceed the mode 2 Nyquist velocity, leading to a loss of power in the mode 2 data because of folded velocities and coherent integration. We attempted to screen this possibility by thresholding on the mode 3 signal-to-noise ratio. However, since mode 2 contributes to the merged reflectivity product primarily above 7 km (Fig. 9b), it is possible that the mode 3 signal-to-noise ratios at these altitudes dropped at times below the threshold that we set on them while the particle fall speeds remained outside of the mode 2 Nyquist velocity. The local maximum in the mode 4 usage between 12 and 15 km (Fig. 9b) lends further support to this argument, as mode 4 is utilized only when the mode 4 velocity estimates exceed all of the other mode Nyquist velocities. Coherent integration artifacts are not apparent in the mode 1 contour plot, while the effects of pulse coding are present but limited to a small number of samples (Fig. 11a_1). Since mode 3 does not use pulse coding, the agreement of mode 3 and mode 4 reflectivities below 0 dB*Z* is expected. The effects of coherent integration on the mode 3 reflectivities are only apparent at reflectivities above 10 dB*Z*.

To illustrate the present methodology for inferring the vertical distribution of hydrometeors above the ARM SGP site we used data from 7 and 8 May 1997. On this day there was significant atmospheric clutter that severely contaminated the returns from boundary layer stratus (Fig. 3K_4). To illustrate the long-term severity of the clutter problem consider the 8640 × 64 range gate sample volumes per day that are from altitudes

below 3 km. We computed the percentage of these sample volumes that contained only clutter or a mixture of clutter and hydrometeor returns (i.e., flag values 2 and 3 in Fig. $3K_4$) for each day from 7 November 1996 to 30 June 1998 (Fig. 12). As the results illustrate, there is a cycle in the clutter associated with the changing seasons (Fig. 12a). During the months of June, July, and August the clutter is at a maximum and most range gate samples below 3 km are contaminated with clutter returns (Fig. 12b). Only during the months of November–February and part of March are the reflectivities between 1 and 3 km relatively clutter-free. This seasonal variation supports the assumption that the targets that produce this clutter are predominately insects because it matches their seasonal presence. Additional sources are likely to include seeds, bits of vegetation, and other small particles that become airborne into the convective boundary layer.

4. Discussion and conclusions

The millimeter-wave cloud radar at the ARM SGP site has operated continuously since mid-November 1996. Apart from a two-month lag period necessary to acquire all of the instrument data, we have now processed the data from this time to the present using the algorithms described above. The final output product includes the variable fields illustrated in Figs. $3G_1$, $3H_2$, $3H_3$, $3I_1$, and $3K_1$–$3K_7$. These variables provide our best estimates of the heights of hydrometeors and their Doppler moments in the vertical column above the ARM SGP site. We anticipate this data product to be one starting point for studies on the distribution of hydrometeors and their properties as a function of time and height at the ARM sites. As the millimeter-wave cloud radars come online at the ARM NSA and TWP sites, we plan to process these radar data with the same algorithms outlined in this paper.

Our current methodology for combining the mode data is to some extent arbitrary. We attempt to maximize the use of data from a single mode (i.e., mode 3) in order to prevent artifacts that may result from interleaving data from different modes. When data from modes other than mode 3 are used, these data are generally concentrated in contiguous regions in time and space. This is an important attribute of the data, as the cloud reflectivities can change dramatically over 10-s timescales and meshing data from different modes on this timescale could be problematic. Nonetheless, having the data from all of the modes enhances the final hydrometeor products by removing potentially unreliable mode 3 data associated both with large rapidly moving hydrometeors and weakly reflecting hydrometeors.

The time and space grids of 10 s and 45 m, respectively, attempt to keep the data close to its original resolution while allowing us to avoid the issue of how best to combine ceilometer and lidar cloud-base height estimates. Detailed comparisons of the merged cloud product with the original time series data of mode 4 demonstrated that small residual effects caused by coherent integration and pulse coding are entering into the final product. However, the number of samples for which these problems are evident is relatively small and confined primarily to the mode 2 data. The current results are consistent with a relative calibration between the four modes of better than 1 dB. Although not discussed in detail in the current study, recent millimeter-wave cloud radar intercomparisons also indicate that the absolute calibration of the SGP MMCR is accurate to within 1 dB.

The number of laser-derived cloud detections missed by the radar amounted to 5.9%. The radar, however, generally detected some part of each cloud element. The number of missed detections dropped to 2.9% when several radar profiles surrounding each laser-derived sample were incorporated into the analysis. Analysis of the mode reflectivities indicates that the performance of the radar would drop significantly if the pulse-coded mode data were unavailable. Exclusion of the pulse coded mode data would lead to a drop of radar sensitivity of approximately 10–15 dB from its optimal state and the number of missed detections would increase to 22%–34%. These results emphasize the importance of the coded modes for a radar attempting to produce a reliable cloud climatology within a single atmospheric column.

As the airborne clutter or ''atmospheric plankton'' flags in Fig. $3K_4$ illustrate, the radar reflectivities at cirrus altitudes are free from clutter. The sensitivity of mode 2 to small particles also ensures that the MMCR at the ARM SGP site detects most cirrus. Therefore, all of the MMCR data are readily usable in cirrus cloud studies. However, the presence of large amounts of clutter in the lower troposphere (Fig. $3K_4$), at least for the ARM SGP site, significantly reduces the quality of the Doppler moments obtained from boundary layer stratus. The clutter was sufficiently severe during 1997 at the ARM SGP site that it contaminated approximately 70%–80% of the boundary layer stratus cases. The uncontaminated stratus cases occurred during the late fall and winter months, presumably when the abundance of insects in the atmosphere was at a minimum.

Our current approach for identifying clutter-contaminated radar reflectivities is limited by the use of lidars for identifying the clutter. For example, in the case of optically thick stratus layers embedded within a plankton layer the lidar beams are completely attenuated in the lower regions of the stratus cloud. Consequently, there is no direct information available above cloud base that enables the clutter contributions to be separated from the hydrometeor contributions to the radar returns. If the characteristics of the clutter at any point in time are not significantly different from the clutter during one of the two surrounding clear-sky periods, the current algorithm correctly identifies the clutter. However, if the characteristics of the clutter do change during a cloudy

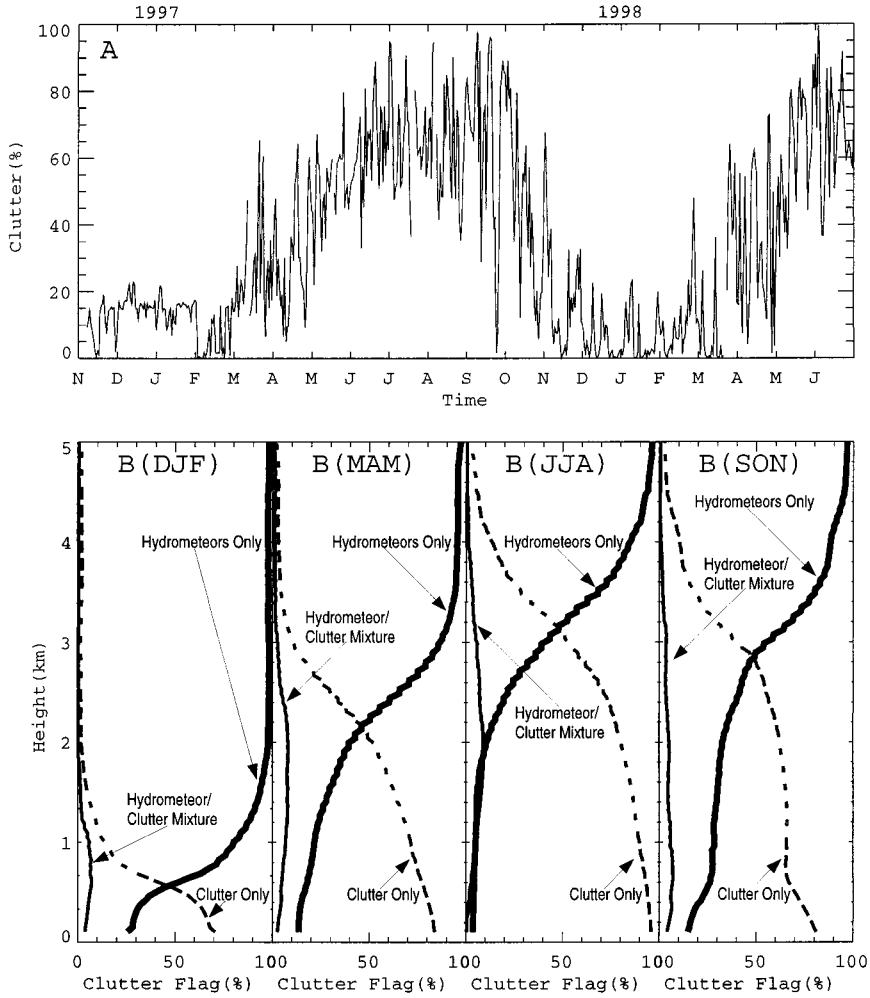

FIG. 12. (a) The percentage of the radar range gates below 3 km for each day that is contaminated with clutter. (b) The frequency of occurrence by season of hydrometeors with insignificant clutter contamination (thick solid line, flag value of 1 in Fig. $3K_4$), of hydrometeors with significant clutter contamination (thin solid line, flag value of 2 in Fig. $3K_4$), and of significant clutter contamination in the absence of any hydrometeors (dashed line, flag value of 3 in Fig. $3K_4$) as a function of height.

period, such as the appearance and disappearance of a lofted insect layer above a stratus cloud, the current algorithm may fail to properly identify the presence of all of the clutter. Therefore, hydrometeor returns that are labeled as uncontaminated should be treated as suspect when they occur at times and heights close to those that have been flagged as clutter contaminated.

For the current hardware configuration of the MMCR one potential solution of the clutter problem is to introduce an additional data-processing step that filters the clutter contributions to the raw voltage time series generated by the radar. Jordan et al. (1997) have developed an algorithm for removing bird contamination from wind-profiler data, and such an algorithm, in principle, is applicable to plankton clutter in millimeter-wave radar returns. However, such a processing step will be successful only if the plankton contribution to the voltage time series has a signature distinct from the cloud hydrometeor contribution. If the plankton in the radar sample volume contains more than a few point targets, the statistics of the signal resulting from it can become indistinguishable from the statistics of the cloud hydrometeor returns. Under these conditions the plankton contributions to the radar returns may not be removable or they may be removable only at certain heights where the plankton is not too abundant.

Polarimetric measurements are an effective tool for identifying insects and other nonspherical plankton throughout the vertical extent of a liquid-water cloud (Mueller and Larkin 1985; Lohmeier et al. 1997; Sekelsky et al. 1998). The ARM program is adding dual-polarization capability to a future MMCR unit to test this solution. However, apart from identifying the presence of the plankton, polarimetric measurements by

themselves do not provide a mechanism for removing it. Interestingly, in comparisons of collocated 35- and 94-GHz returns with nearly identical beamwidths the presence of the atmospheric plankton is much less conspicuous in the 94-GHz data. This result illustrates that 94-GHz may be a useful radar operating frequency at heights in the atmosphere where plankton is abundant.

The current output data of the MMCR, that is, the three Doppler moments, are not sufficient for separating cloud and precipitation particles throughout the vertical column. Consequently, we labeled the particles detected by the radar as hydrometeors instead of cloud particles. We assumed that the ceilometer and micropulse lidar accurately detect cloud base under all conditions, although precipitation can lead to an underestimate of the cloud-base height by these systems. These kinds of shortcomings in the current datasets are being rectified. In order to begin to separate precipitation from cloud particles throughout the vertical columns above the ARM sites an effort is currently underway within the ARM program to routinely archive the Doppler spectra, as earlier studies by Gossard et al. (1997) and Babb et al. (1999) indicate that spectral analyses might be capable of identifying cloud particles in the presence of precipitation. Recent analyses of Raman lidar data by Demoz et al. (1999) demonstrate that these data can be used to accurately identify the cloud-base height under all circumstances. Therefore, the coincident Raman lidar, micropulse lidar, and ceilometer data at the SGP central facility might enable algorithms to be developed for the micropulse lidar and ceilometer that are accurate during periods of precipitation. Last, the deployment of the high-resolution lidars at the ARM CART sites, together with improvement of the algorithms applied to their data, should enable the unambiguous identification of cloud top in those cases where the lidar beam is not completely attenuated by underlying clouds.

Acknowledgments. Support for this research was funded by the Environmental Sciences Division of the U.S. Department of Energy (under Grant DE-FG02-90ER61071) and Battelle Pacific Northwest Laboratory (Subcontract 091572-A-Q1). Data were obtained from the Atmospheric Radiation Measurement (ARM) Program sponsored by the U.S. Department of Energy, Office of Energy Research, Office of Health and Environmental Research, Environmental Sciences Division. We would like to thank Taneil Uttal of NOAA ETL for many useful comments in regard to this study. Stephen Sekelsky has provided valuable data and insights into possible techniques for attempting to mitigate the effects of atmospheric clutter. William Syrett provided us with the data to make the charts in Fig. 4 and assisted us in the interpretation of these data. This study would not have been possible without the support of members of the ARM infrastructure. We especially wish to note Connor Flynn's assistance in interpreting the ceilometer and micropulse lidar data, Krista Gaustad's unrelenting pursuit of producing SGP MMCR data in good form, Robin Perez's efficiency in providing us all of the needed data, David Turner's software support, and Daniel Rodriguez's logistical support.

APPENDIX A

The Cloud Mask Algorithm

For a vertically pointing millimeter-wave radar that is not viewing the sun the only significant source of noise that contaminates the radar signal returns $P_{a,ij}$ from the atmosphere is the noise power $P_{n,ij}$ from the mixer in the radar receiver. In this notation the subscripts i and j refer to the time of day the radar power was collected and the radar range gate corresponding to a particular radar sample volume height, respectively. The P stands for the power (mW) and the subscripts a and n represent atmospheric and noise contributions, respectively. For a single pulse, therefore, the total output power at the radar receiver is $P_{t,ij} = P_{a,ij} + P_{n,ij}$.

The mixer noise power $P_{n,ij}$ is a random sample from a ''Gaussian-like'' process with a mean power of $\overline{P}_n$ and a standard deviation of $\sigma_{n,}$. Since an estimate of the average radar return power $\overline{P}_{a,ij}(\Delta t)$, where Δt represents some time interval about the ith time of the day, is the meaningful quantity, we average N_s consecutive power samples at each height j to form

$$P_{t,ij}(N_s) = P_{a,ij}(N_s) + P_{n,ij}(N_s). \tag{A1}$$

Since the standard deviation of the noise $\sigma_{n,ij}(N_s)$ about the average noise power $P_{n,ij}(N_s)$ is approximately

$$\sigma_{n,ij}(N_s) = \frac{\sigma_n}{(N_s)^{1/2}}, \tag{A2}$$

large values of N_s lead to values of $P_{n,ij}(N_s)$ with smaller uncertainties about $\overline{P}_n$. Large values of N_s also ensure that the average power $P_{a,ij}(N_s)$ becomes a better estimate of the total backscatter cross section of the particles in the radar sample volume. For each radar sample volume, therefore, our first criterion for a possible significant detection is

$$\overline{P}_n + \frac{\sigma_n}{(N_s)^{1/2}} < P_{t,ij}(N_s). \tag{A3}$$

Since we do not know $\overline{P}_n$ and $\sigma_n(N_s)^{-1/2}$ a priori, we estimate them from the average power returns $P_{t,ij}(N_s)$ for resolution volumes j such that $N_g - N_{\text{trop}} + 1 \leq j \leq N_g$, where N_g is the number of resolution volumes along the vertical atmospheric column that the radar samples and N_{trop} is the number of resolution volumes above the tropopause. For those resolution volumes that are above the tropopause, we assume that the atmospheric power return $P_{a,ij}(N_s)$ is zero and $P_{t,ij}(N_s) = P_{n,ij}(N_s)$. Therefore, we use the average

$$P_{n,i}(N_s) = \frac{1}{N_{\text{trop}}} \sum_{j=N_g-N_{\text{trop}}+1}^{j=N_g} P_{t,ij}(N_s) \qquad \text{(A4a)}$$

as an estimate of $\overline{P}_n$ and we use the standard deviation

$$\sigma_{n,i}(N_s) = \sqrt{\frac{1}{N_{\text{trop}} - 1} \sum_{j=N_g-N_{\text{trop}}+1}^{j=N_g} [P_{t,ij}(N_s) - P_{n,i}(N_s)]^2} \qquad \text{(A4b)}$$

as an estimate of $\sigma_n(N_s)^{-1/2}$. We apply checks to the estimates generated in this fashion and if the estimates are unreasonable for a particular time i, we use the noise estimates generated at an earlier neighboring time.

Applying the criterion in (A3) to N_t temporally consecutive profiles, each containing N_g radar sample volumes, we produce a $N_t \times N_g$ ''cloud mask binary image,'' where values corresponding to radar sample volumes that satisfy (A3) are labeled with a 1 and all other resolution volumes are labeled with a 0. Because of the receiver mixer noise some values labeled 1 actually contain pure noise and some pixels labeled by 0 contain a significant atmospheric power return. Following the earlier work of Clothiaux et al. (1995), we analyze each pixel in the binary mask in the context of the box of 25 pixels centered on it. If the value

$$p = (0.84)^{n_0}(0.16)^{n_1}, \qquad \text{(A5)}$$

where n_0 and n_1 are the number of pixel values in the box labeled 0 and 1, respectively, and 0.84 and 0.16 are the areas under a standard normal distribution for values less than and greater than one standard deviation, respectively, is less than the threshold $p_{\text{thresh}} = 5 \times 10^{-12}$, then the box center pixel is set to 1; otherwise, it is set to 0. The threshold p_{thresh} and box size are set by maximizing the performance of the algorithm on a simulated radar dataset with realistic noise and signal return properties. The order in which the pixels are tested is random, that is, an asynchronous update scheme, since we changed the pixel values in the image itself and all of the $N_t \times N_g$ pixels must be updated n times before any pixel value is updated $n + 1$ times. We currently update each pixel two to five times to produce the final cloud mask. Updating each pixel only once often leaves noise speckle in the image, whereas updating each pixel more than five times leads to a degradation of resolution in the final cloud mask.

To enhance the detection of thin cloud layers in the lower troposphere we complement the above tests with an additional test that takes into account the magnitude of the signal return with respect to the noise. In this second test we form the sum

$$p = \sum_{i'=i-1}^{i+1} \sum_{j'=j-1}^{j+1} \left\{ \frac{P_{t,ij}(N_s) - [P_{n,i}(N_s) + \sigma_{n,i}(N_s)]}{\sigma_{n,i}(N_s)} \right\}^2, \qquad \text{(A6)}$$

where only the terms that satisfy (A3) are included. If the value of p fails to reach a threshold value $p_{\text{thresh}} = 10^5$, the power $P_{a,ij}(N_s)$ is considered insignificant and the corresponding pixel value in the ''cloud mask image'' is set to 0; otherwise, the pixel value is kept at 1. In this second test we use only 9 pixel values in order to maintain higher spatial resolution in the final cloud detection mask. Furthermore, pixel values of 0 are never set to 1, again in an attempt to keep the cloud edges from smearing.

APPENDIX B

Assessing the Quality of Individual Mode Data

Three mechanisms that can compromise accurate moment estimates are range sidelobes as a result of pulse coding, range aliasing (or second-trip echo returns), and coherent averaging of strong power returns from rapidly moving hydrometeors. Range sidelobes from pulse coding and second-trip echo returns either cause apparent power returns from radar sample volumes containing no hydrometeors or contaminate the power returns from radar sample volumes that do contain hydrometeors. Coherent averaging of the power returns from large, rapidly moving hydrometeors leads to inaccurate Doppler moments. To remove these problems from the merged data illustrated in Fig. 3D we use the merged data in Fig. 3D to help to identify problems in the original mode data illustrated in Fig. 3A.

To prevent range sidelobe artifacts from entering into the initial cloud masks Moran et al. (1998) developed a simple threshold test for identifying radar sample volume power returns that are potentially contaminated by range sidelobe effects. Let $P_{t',ij}(N_s) = 10 \log[P_{t,ij}(N_s)]$ be the return power to the radar in dBm from the jth range gate at the ith time of the day. The power return at the $j = n$ range gate height is considered to be free of range sidelobe effects if

$$P_{t',ik}(N_s) < P_{\text{thresh}} + P_{t',in}(N_s) \qquad \text{(B1)}$$

for all k such that $n - N_{\text{bits}} \leq k \leq n + N_{\text{bits}}$. The parameter N_{bits} is the number of coded bits in the pulse-coding sequence (Schmidt et al. 1979; Moran et al. 1998). We initially set $P_{\text{thresh}} = 25$ dB to test all of the power returns. This value of the threshold eliminated most, but not all, noticeable occurrences of range sidelobe effects. For those few times when range sidelobe effects were present with the 25-dB threshold we reprocessed just these time periods with a 15-dB threshold, thereby eliminating all of the remaining artifacts. The only noticeable loss of cloud that resulted from using this simple threshold test occurred when distinct, weak returns from cirrus were located within N_{bits} sample volumes of large reflectivity, high Doppler velocity radar returns. Detailed numeric simulations of the range sidelobe effects using the radar ambiguity function to derive the effects may eventually enable these weak cirrus returns to be recovered (R. Lataitis 1997, personal communication).

To identify second-trip echoes the merged reflectivity field (e.g., Fig. $3D_1$), which accurately portrays the location of high-altitude cloud elements, is used to identify high-altitude clouds that can lead to second-trip echo problems at lower elevations in each of the modes.

For the case study period illustrated in Fig. 3 the tops of the deep convective clouds are producing second-trip echoes in the mode 4 data as evidenced by the appearance of these cloud tops in the lowest elevations of the mode 4 data (Fig. $3E_4$). These regions in the mode 4 data, as well as any corresponding regions in the other mode data, are flagged and subsequently discarded in the second merge process.

Coherent averaging of the voltage series generated by a mode can lead to artifacts in the data when the speeds of the particles approach the Nyquist velocity of the mode. The Nyquist velocity of mode 4 since September 1997 is approximately 20 m s^{-1}. Therefore, during periods when there is not heavy rainfall, which is most of the time, the velocity estimates derived from this mode are generally accurate to within the velocity resolution of the mode data, that is, 0.32 m s^{-1}. Recall that to produce the velocity field illustrated in Fig. $3D_2$ we use the mode 4 velocity estimates to ensure that the velocity estimates of the mode actually utilized in the merged product are within their Nyquist velocity. Hence the velocity field shown in Fig. $3D_2$ should be accurate. Each point in the time–height Doppler moment fields generated from a single mode is flagged as suspect if the velocity recorded in the merged velocity field (e.g., Fig. $3D_2$) exceeds the Nyquist velocity of the mode data. Again, the Doppler moments that are flagged as suspect as a result of coherent averaging are not utilized in the second merge process.

APPENDIX C

The Laser-Dependent Best-Estimate Cloud-Base Height Algorithm

To produce a single estimate of the cloud-base height from the ceilometer and lidar instruments we attempt to use the instrument cloud-base height that is the most accurate in a particular region of the atmosphere. The high resolution of the Belfort laser ceilometer data makes these data attractive to use in the lower troposphere. Consequently, if the Belfort laser ceilometer detects a cloud between the surface and 3 km we use this value for the cloud-base height. We also use the Belfort laser ceilometer cloud-base height estimates that are above 3 km, as long as they are within a specified distance, which we take to be 600 m, of a cloud detection by one of the two micropulse lidar cloud-detection algorithms. If the Belfort laser ceilometer either fails to detect a cloud or produces a cloud-base height that does not satisfy the above criteria, the micropulse lidar cloud-base heights are subsequently checked. If the Campbell et al. (1998) algorithm detects a cloud contribution to the micropulse lidar returns, its estimate of cloud-base height is utilized. Otherwise, if the Clothiaux et al. (1998) algorithm detects a cloud, its value is utilized. If there is no cloud detection by any of the systems and one of the systems reports no obscurations as a result of dew, or perhaps fog, then a ''clear-sky'' value is reported. If all of the laser systems produce no data, then a ''no-data'' value is reported. Otherwise, if the laser systems produce data but it is flagged as suspect, then an ''invalid-data'' value is reported.

REFERENCES

Babb, D. M., J. Verlinde, and B. A. Albrecht, 1999: Retrieval of cloud microphysical parameters from 94-GHz radar Doppler power spectra. *J. Atmos. Oceanic Technol.,* **16,** 489–503.

Campbell, J. R., D. L. Hlavka, J. D. Spinhirne, D. D. Turner, and C. J. Flynn, 1998: Operational cloud boundary detection and analysis from micropulse lidar data. *Proc. Eighth ARM Science Team Meeting,* Tucson, AZ, U.S. Department of Energy, 119–122.

Clothiaux, E. E., M. A. Miller, B. A. Albrecht, T. P. Ackerman, J. Verlinde, D. M. Babb, R. M. Peters, and W. J. Syrett, 1995: An evaluation of a 94-GHz radar for remote sensing of cloud properties. *J. Atmos. Oceanic Technol.,* **12,** 201–229.

——, G. G. Mace, T. P. Ackerman, T. J. Kane, J. D. Spinhirne, and V. S. Scott, 1998: An automated algorithm for detection of hydrometeor returns in micropulse lidar data. *J. Atmos. Oceanic Technol.,* **15,** 1035–1042.

——, and Coauthors, 1999a: The Atmospheric Radiation Measurement Program cloud radars: Operational modes. *J. Atmos. Oceanic Technol.,* **16,** 819–827.

——, and Coauthors, 1999b: The ARM millimeter-wave cloud radars (MMCRs) and the active remote sensing of clouds (ARSCL) value added procedure (VAP). Department of Energy Tech. Memo. ARM VAP-001, 58 pp. [Available from Battelle, Pacific Northwest National Laboratory, 101 Battelle Boulevard, Richland, WA 99352.]

Demoz, B. B., and Coauthors, 1999: Determination of cloud base height using the GSFC Raman lidar. *Proc. Ninth ARM Science Team Meeting,* San Antonio, TX, U.S. Department of Energy. [Available online at http://www.arm.gov./docs/documents/technical/conference.html]

Gossard, E. E., J. B. Snider, E. E. Clothiaux, B. Martner, J. S. Gibson, R. A. Kropfli, and A. S. Frisch, 1997: The potential of 8-mm radars for remotely sensing cloud drop size distributions. *J. Atmos. Oceanic Technol.,* **14,** 76–87.

Jordan, J. R., R. J. Lataitis, and D. A. Carter, 1997: Removing ground and intermittent clutter contamination from wind profiler signals using wavelet transforms. *J. Atmos. Oceanic Technol.,* **14,** 1280–1297.

Lhermitte, R. M., 1966: Probing air motion by Doppler analysis of radar clear air returns. *J. Atmos. Sci.,* **23,** 575–591.

Lohmeier, S. P., S. M. Sekelsky, J. M. Firda, G. A. Sadowy, and R. E. McIntosh, 1997: Classification of particles in stratiform clouds using the 33 and 95 GHz polarimetric cloud profiling radar system (CPRS). *IEEE Trans. Geosci. Remote Sens.,* **35,** 256–270.

Moran, K. P., B. E. Martner, M. J. Post, R. A. Kropfli, D. C. Welsh, and K. B. Widener, 1998: An unattended cloud-profiling radar for use in climate research. *Bull. Amer. Meteor. Soc.,* **79,** 443–455.

Mueller, E. A., and R. P. Larkin, 1985: Insects observed with dual-polarization radar. *J. Atmos. Oceanic Technol.,* **2,** 49–54.

Pal, S. R., W. Steinbrecht, and A. I. Carswell, 1992: Automated method for lidar determination of cloud-base height and vertical extent. *Appl. Opt.,* **31,** 1488–1494.

Schmidt, G., R. Ruster, and P. Czechowsky, 1979: Complementary code and digital filtering for detection of weak VHF radar signals from the mesosphere. *IEEE Trans. Geosci. Electron.,* **17,** 154–161.

Sekelsky, S. M., and Coauthors, 1998: Comparison of millimeter-wave cloud radar measurements for the Fall 1997 cloud IOP. *Proc. Eighth ARM Science Team Meeting,* Tucson, AZ, U.S. Department of Energy, 671–675.

Spinhirne, J. D., 1993: Micro pulse lidar. *IEEE Trans. Geosci. Remote Sens.,* **31,** 48–55.

Stokes, G. M., and S. E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the Cloud and Radiation Test Bed. *Bull. Amer. Meteor. Soc.,* **75,** 1201–1221.