

# An Automated Cloud-Edge Detection Algorithm Using Cloud Physics and Radar Data

JENNIFER G. WARD AND FRANCIS J. MERCERET

*NASA, Kennedy Space Center, Florida*

(Manuscript received 10 August 2003, in final form 20 December 2003)

## ABSTRACT

An automated cloud-edge detection algorithm was developed and extensively tested. The algorithm uses in situ cloud physics data measured by a research aircraft coupled with ground-based weather radar measurements to determine whether the aircraft is in or out of cloud. Cloud edges are determined when the in/out state changes, subject to a hysteresis constraint. The hysteresis constraint prevents isolated transient cloud puffs or data dropouts from being identified as cloud boundaries. The algorithm was verified by detailed manual examination of the dataset in comparison to the results from application of the automated algorithm.

## 1. Introduction

Since 1999, the Kennedy Space Center has sponsored an Airborne Field Mill (ABFM) experiment in support of its Lightning Launch Commit Criteria (LLCC) project. The LLCC project aims to improve weather constraints (launch commit criteria) designed to protect space launch vehicles, including the space shuttle, from natural and triggered lightning. If these constraints are violated, launch is delayed or scrubbed until the weather improves. The first ABFM field campaign took place in June 2000. Another was conducted in May–June 2001 for a total of 30 flight days. A key component of the experimental design was to couple ground-based weather radar measurements with in situ cloud physics and electric field measurements from an instrumented aircraft. Details are presented in Merceret and Christian (2000).

The ABFM measurements are being used to learn enough about the behavior of electric charge in and near clouds to safely relax the LLCC. Although the current constraints are safe, they have a high false alarm rate (rule violated when it would actually be safe to fly). This is due primarily to our ignorance of how charge behaves in the atmosphere compounded by the need for large margins to ensure safety because life and expensive property are at risk. The LLCC project is directed at reducing our ignorance level so that less restrictive yet even safer rules may be developed.

A first step in understanding charge behavior is collecting accurate estimates of electric field decay as a function of distance from cloud boundaries. The current

LLCC presume flight within 5 n mi (9.3 km) of any anvil may be dangerous for up to 3 h after any electrical activity in the anvil. Measurements of the spatial behavior of the electric field with respect to the edge of the anvil as a function of time may permit refinement of these rules. Because of the massive amount of data collected by the project, an automated system for identifying cloud edges is essential. A search of the literature found abundant algorithms for identifying cloud boundaries in satellite photographs (e.g., Rossow and Gardner 1993; Solvsteen 1995; Fejit et al. 2000). There were also algorithms for radiosonde data (Naud et al. 2003), lidar data (Clothiaux et al. 1998), and clear-air Doppler radar (Gossard et al. 1992). Unfortunately, nothing was found for automated cloud-edge detection in in situ cloud physics data with or without associated radar support. Several members of the ABFM project team with extensive experience in cloud physics were consulted, but only one had a possible cloud boundary detection algorithm. It proved unsatisfactory, as noted below, so this algorithm was developed and is being reported here.

The new cloud-edge detection algorithm has two components: an in-cloud detection component and a boundary detection component. The in-cloud component relies on cloud physics data from the research aircraft as well as ground-based weather radar data. Details of the instrumentation are given in Ward et al. (2003). The boundary detection component examines the output of the in-cloud algorithm and applies a hysteresis test to avoid false boundary detections due to momentary fluctuations in the data. This article describes the development and testing of the algorithm.

## 2. Methodology

The National Center for Atmospheric Research (NCAR) provided ASCII format files containing time-

---

*Corresponding author address:* Francis J. Merceret, NASA/ YA-D, Kennedy Space Center, FL 32899.  
E-mail: Francis.J.Merceret@nasa.gov

synchronized and quality controlled values at 10-s intervals for the following variables used to develop this algorithm.

- Cloud particle concentrations (per liter) from a Particle Measuring System (PMS) Forward Scattering Spectrometer Probe (FSSP), and PMS 1D and 2D cloud probes.
- King liquid water content (LWC) and Rosemount ice detector probes.
- Radar reflectivity at the aircraft position (dBZ) measured by ground-based weather radar.

Other variables are in these files, but they were not used to develop the cloud-edge detection algorithm and will not be discussed here. In addition to the ASCII files, NCAR provided time-synchronized constant altitude radar maps (CAPPI) with the aircraft track overlaid and simultaneous time series plots [i.e., microphysics–field–radar (MER)] of all of the above variables.

The MER and CAPPI plots were manually examined for each day. A list of each entry into or exit from cloud was compiled with the time of the transition estimated to the nearest 10 s, and the behavior of the cloud physics measurements was noted. A tentative relationship between these variables and the analyst's judgment regarding the presence or absence of cloud was formed. This judgment was refined by a more detailed examination of each cloud boundary transition until an objective process for determining whether cloud was present was formulated.

Initially, an algorithm provided by the University of North Dakota (UND) using only the FSSP data was tested. UND has advised caution in the use of the algorithm because noise in the FSSP would give false cloud indications if the threshold of the algorithm was set low. If the threshold was set high, there would be false indications of no-cloud (C. Grainger 2003, personal communication). Our analysis confirmed these weaknesses. Ultimately, we found that every algorithm depending on a single instrument, whether in situ or radar, was subject to similar problems. Radar data were sometimes contaminated by sidelobes, anomalous propagation, and data dropouts. The in situ instruments all were subject to intermittent dropouts or noise. This was overcome by incorporating data from multiple sensors into the algorithm presented here.

Once developed, the algorithm was coded and run without manual intervention on just the ASCII data. The results were compared with the manual analysis. In most cases, the results were identical. In those cases where there were discrepancies, further analysis proved the automated algorithm to be correct. This will be discussed in section 5.

Once the reliable method of determining the presence of cloud was complete, the remaining task for automated boundary detection was to incorporate some way of handling fluctuations at cloud edges to avoid rapid cycling in wispy cloud fragments at a cloud boundary.

This was done with the hysteresis check described in section 4. These two elements, cloud detection and hysteresis, compose the cloud-edge detection algorithm. Comparison of the manual and automated cloud edges was used to select the appropriate hysteresis threshold.

### 3. Cloud presence component

Upon examining the MER plots, it became apparent that if the 2D cloud probe total was greater than or equal to  $0.1 \text{ L}^{-1}$ , then the aircraft was in-cloud. There were frequent cases where the 2D total was less than  $0.1 \text{ L}^{-1}$ , but there was cloud present. In order to diagnose these cases, the 1D total was examined along with the radar reflectivity at the aircraft. If the 1D total was greater or equal to  $1 \text{ L}^{-1}$  and the radar reflectivity was  $>0 \text{ dBZ}$ , then the aircraft was in-cloud. If neither of the above indicated the presence of cloud, the presence of any large particles on the 2D probe, "2D  $> 1 \text{ mm}$ ," would indicate that the aircraft was in-cloud. Otherwise, the aircraft was out-of-cloud. Because this algorithm was created for anvil and mid- to high-level clouds, it is possible for a false "in-cloud" reading to occur in some circumstances such as low-level flight in precipitation. A flowchart of the cloud presence detection component is presented in Fig. 1. The FSSP was not used because it was too noisy resulting in too many false in-cloud indications. The LWC was not used because it did not give a reliable in-cloud indication in anvils, which were our primary focus.

### 4. Hysteresis component

Because the goal of locating cloud boundaries for this project is to examine the variation of the electric field with distance from cloud edge, it is essential to isolate true boundaries of significant clouds. Unfortunately, small wisps of cloud in otherwise clear air will be designated as in-cloud, and small gaps in otherwise solid cloud masses will be designated as "clear" by any local automated cloud detection algorithm. These designations are not erroneous, but neither are they desirable for finding the true edge of nearly continuous cloud masses. In addition, momentary data dropouts can be falsely interpreted as cloud boundaries.

The solution we have adopted is to only use "clean" cloud boundaries in our dataset. A clean boundary is a cloud boundary with two additional "hysteresis" constraints. There are four steps in the process. Unless all four steps are satisfied, there is no cloud edge as defined by this algorithm. In the steps listed below, a "record" refers to one line of 10-s data in a file. Each line contains the 10-s average of each of the measured variables along with the position and attitude of the aircraft and the time of day.

- Examine the current record for a transition from cloud to clear or clear to cloud.

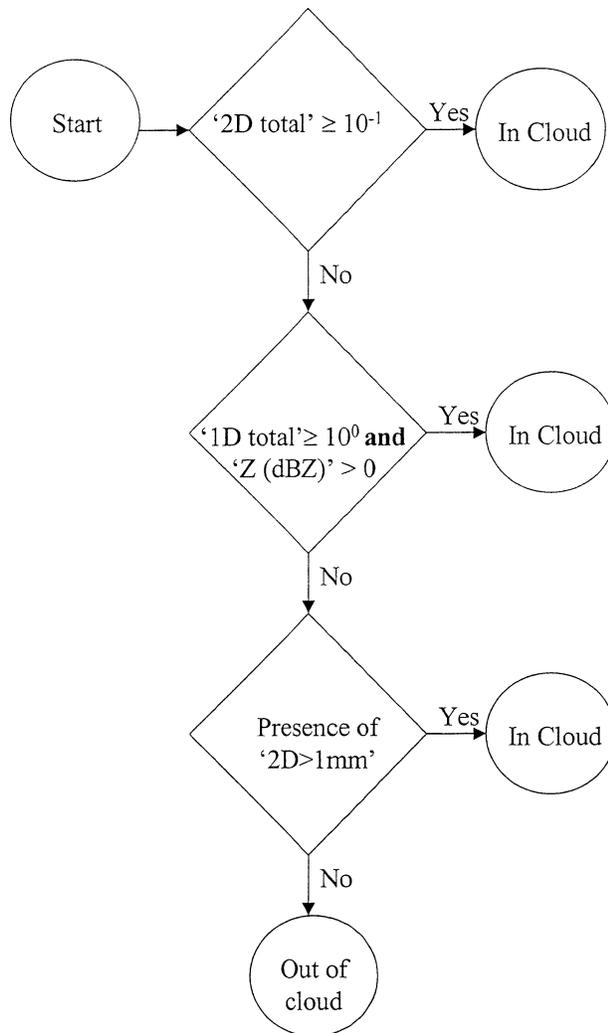


FIG. 1. Flowchart of the cloud presence component of the cloud-edge detection algorithm.

- If a transition has occurred, examine the previous 20 records to locate how many records back,  $J_{\text{Minus}}$ , the immediately previous transition occurred. If no transition is found,  $J_{\text{Minus}}$  is set to 20.
- If a transition has occurred, examine the next 20 records to locate how many records ahead,  $J_{\text{Plus}}$ , the next transition occurs. If no transition is found,  $J_{\text{Plus}}$  is set to 20.
- Both  $J_{\text{Plus}}$  and  $J_{\text{Minus}}$  must be greater than or equal to a user-selected value,  $H$ , between 0 and 10.

Selecting  $H = 0$  turns off all hysteresis testing and locates all boundaries, however evanescent. Setting  $H = N$  ensures that at least  $N$  continuous records of the same kind (in-cloud or clear) exist on each side of the boundary.

For the ABFM program, the records are spaced 10 s apart. The true airspeed of the research aircraft ranged from 100 to 130  $\text{m s}^{-1}$ . Thus, the value of  $H$  is approximately the length in kilometers of cloud/clear continuity required on each side of the cloud edge for that transition to be included in the analysis dataset. Values of  $H$  ranging from 0 to 10 were tried on sample days. Here  $H = 2$  most closely matched the manual analysis of cloud boundaries, and  $H < 2$  included transitions due to data dropouts and small puffs of cloud undetectable on radar. Data dropouts can occur for a variety of reasons including instrument anomalies, recording system failures, power bus transients, and operator error. Note that  $H > 2$  eliminated transitions significant enough for the analyst to list them.

## 5. Verification

Once the in-cloud rule was devised and the hysteresis concept developed, code was created and the dataset was processed. A sample of the output is shown in Table 1. Columns labeled  $C(N)$  contain the algorithm's evaluation of whether the aircraft was in-cloud or in the clear at time  $N$  from the cloud boundary detected by the algorithm. Time  $N$  ranges from  $-5$  to  $5$  where each unit corresponds to 10 s of flight. This unit was selected for two reasons. First, the data were available at 10-s intervals, so  $N$  corresponds to the number of records from the boundary. Second, the aircraft speed was about 100  $\text{m s}^{-1}$  so each unit is approximately 1 km of the distance. If the data required to determine the presence of clouds were flagged by the automated QC process as suspect, the designation "?Clear" appears.

This was compared to the manual cloud detection spreadsheets completed beforehand. The results showed all of the manual entry/exit points had been picked up by the software as well as some additional points. These other points were examined more closely and determined to be correct. The reason they were overlooked in the manual analysis was that they were less than 20 long, mostly near the edge of the MER plots, and looked initially like artifacts of the plotting process. A hysteresis of 2 was chosen as the optimum one because it agreed most closely with the manual evaluations. Small-

TABLE 1. Example of a few lines of output from the automated cloud-edge detection algorithm with hysteresis,  $H$ , set to 0.

Type	HHMMSS	Lat	Lon	Alt (m)	C(-4)	C(-3)	C(-2)	C(-1)	C(0)	C(1)	C(2)	C(3)	C(4)
Exit	215830	27.472°	-79.993°	6990	Cloud	Cloud	Cloud	Cloud	Clear	Cloud	Clear	Clear	Clear
Entry	215840	27.462°	-79.998°	6938	Cloud	Cloud	Cloud	Clear	Cloud	Clear	Clear	Clear	Cloud
Exit	215850	27.455°	-80.007°	6879	Cloud	Cloud	Clear	Cloud	Clear	Clear	Clear	Cloud	Cloud
Entry	215920	27.45°	-80.042°	6724	Cloud	Clear	Clear	Clear	Cloud	Cloud	Cloud	?Clear	Cloud
Exit	220540	27.394°	-80.495°	6701	Cloud	Cloud	Cloud	Cloud	Clear	Clear	Clear	Cloud	Cloud

er values included minor data dropouts or cloud wisps as transitions, while large values excluded some valid transitions. In the full dataset, the manual process found 1014 entry/exit transitions, while the automated algorithm found 1269.

## 6. Conclusions

An automated process for identifying cloud boundaries in airborne cloud physics data with accompanying ground-based radar was developed and tested. For a fully automated procedure in an application as sensitive as revising the LLCC, we needed to be absolutely confident of the result. By using multiple instruments we have achieved a robustness not achievable using data from any one instrument.

The algorithm performed slightly better than manual analysis on an extensive dataset from the Airborne Field Mill Program. It permits automated analysis of the variation of electric field and radar reflectivity with distance from cloud edge. It can also be used to automate stratification of data depending on cloud presence for statistical analysis. Both of these functions are extremely labor intensive when performed manually. The automated algorithm is expected to reduce the labor required for the target analyses by more than 75%. Although the algorithm was developed and verified with specific instrumentation, the concept should be applicable with minor modifications to any project where airborne cloud physics data with or without simultaneous ground-based radar are available—even if the instrumentation is not exactly the same. For example, FSSP or LWC data could be used depending on the noise level of the specific instrument and the intended application. Selection of an in situ instrument sensitive only to smaller drop sizes could eliminate subcloud precipitation from being detected as cloud.

*Acknowledgments.* The authors acknowledge the ABFM team and especially Jim Dye, Sharon Lewis, and Mike Dye of the National Center for Atmospheric Research for providing the calibrated, synchronized data used in this project. We also thank Tony Grainger of the University of North Dakota for detailed information about the cloud physics instrumentation used to collect the data and for providing the first algorithm we tested. Mention of a proprietary product or service does not constitute an endorsement thereof by the authors, their employers, or the American Meteorological Society.

## REFERENCES

- Clothiaux, E. E., 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.
- Feijt, A., P. de Valk, and S. van der Veen, 2000: Cloud detection using METEOSAT imagery and numerical weather prediction model data. *J. Appl. Meteor.*, **39**, 1017–1030.
- Gossard, E. E., R. G. Strauch, D. C. Welsh, and S. Y. Matrosov, 1992: Cloud layers, particle identification, and rain-rate profiles from ZRV<sub>p</sub> measurements by clear-air Doppler radars. *J. Atmos. Oceanic Technol.*, **9**, 108–119.
- Merceret, F. J., and H. Christian, 2000: KSC ABFM 2000—A field program to facilitate safe relaxation of the lightning launch commit criteria for the American space program. Preprints, *Ninth Conf. on Aviation and Range Meteorology*, Orlando, FL, Amer. Meteor. Soc., 6.4.
- Naud, C. M., J. P. Muller, and E. E. Clothiaux, 2003: Comparison between active sensor and radiosonde cloud boundaries over the ARM Southern Great Plains site. *J. Geophys. Res.*, **108**, 4140, doi:10.1029/2002JD002887.
- Rossov, W. B., and L. C. Gardner, 1993: Cloud detection using satellite measurements of infrared and visible radiances for ISCCP. *J. Climate*, **6**, 2341–2369.
- Solvsteen, C., 1995: A correlation based cloud-detection algorithm combined with an examination of some split-window assumptions. *Int. J. Remote Sens.*, **16**, 2875–2901.
- Ward, J. G., F. J. Merceret, and C. A. Grainger, 2003: An automated cloud-edge detection algorithm using cloud physics and radar data. NASA Tech. Memo. TM-2003-211189, 20 pp.