



U.S. DEPARTMENT OF
ENERGY

Office of
Science

DOE/SC-ARM/TR-098

Micropulse Lidar Cloud Mask Value-Added Product Technical Report

C Sivaraman
J Comstock

July 2011



DISCLAIMER

This report was prepared as an account of work sponsored by the U.S. Government. Neither the United States nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

Micropulse Lidar Cloud Mask Value-Added Product Technical Report

C Sivaraman
J Comstock

July 2011

Work supported by the U.S. Department of Energy,
Office of Science, Office of Biological and Environmental Research

Contents

1.0	Introduction	1
2.0	Algorithm and Methodology	1
3.0	The Input Data.....	3
3.1	Flow chart.....	3
4.0	Output Data	6
5.0	Summary.....	7
6.0	References	7

Figures

1	Flow diagram of MPLCMASK VAP.	5
2	An example of a plot generated by the VAP.	6

1.0 Introduction

Lidar backscattered signal is a useful tool for identifying vertical cloud structure in the atmosphere in optically thin clouds. Cloud boundaries derived from lidar signals are a necessary input for popular ARM data products, such as the Active Remote Sensing of Clouds (ARSCL) product. An operational cloud boundary algorithm (Wang and Sassen 2001) has been implemented for use with the ARM Micropulse Lidar (MPL) systems. In addition to retrieving cloud boundaries above 500 m, the value-added product (VAP) named Micropulse Lidar Cloud Mask (MPLCMASK) applies lidar-specific corrections (i.e., range-square, background, deadtime, and overlap) as described in Campbell et al. (2002) to the measured backscattered lidar. Depolarization ratio is computed using the methodology developed by Flynn et al. (2007) for polarization-capable MPL systems. The cloud boundaries output from MPLCMASK will be the primary lidar cloud mask for input to the ARSCL product and will be applied to all MPL systems, including historical data sets.

2.0 Algorithm and Methodology

The primary derived quantity determined in MPLCMASK is the vertical location of cloud layers as clouds pass over the lidar system. In order to retrieve accurate cloud layers, we must apply several corrections to the lidar signal. MPLCMASK applies corrections for molecular background, range-square, deadtime, and overlap according to methods described in Campbell et al. (2002). The deadtime and overlap corrections are measured by the vendor and are system-dependent. The molecular background is computed by either averaging the backscatter signal from the highest altitude bins (i.e. bins 44–54) or using the first 203 bins for instruments when the background signal is stored there (i.e., for “fast-switching” MPLs). All corrections are applied to the returned lidar signal except the range-square correction. However, range-square is applied to the final backscattered signal that is output to the netCDF file.

Cloud Boundary Detection

The cloud detection algorithm, based on Wang and Sassen (2001), uses the change in the slope of the backscattered signal as a function of height to identify the presence of cloud layers in the atmosphere, and uses several techniques to distinguish aerosol layers from cloud layers. The algorithm is applied between 500 m and 20 km above ground level. There are five basic steps to this algorithm as detailed in Wang and Sassen (2001):

1. The slope, standard deviation (σ) of the background noise, and signal noise are computed.
2. Starting near the ground, the lidar signal is examined upward to identify and record possible layers. Layers are identifying by marking the height of the base, peak, and top for each layer. These points are determined by examining the signal slope. The layer base corresponds to the location when there is a positive signal slope; layer top corresponds to the location where the signal returns to the molecular backscatter level or noise level.
3. The cloud layer is distinguished from an aerosol layer or random noise by examining the physical characteristics of the signal. Aerosol layers are distinguished from clouds primarily because aerosol layers have smaller extinction or backscatter coefficients than clouds. The algorithm uses the ratio of

the peak backscattered signal in the layer over the backscattered signal at the layer base to distinguish these layers.

4. The algorithm examines the cloud-top signal to determine if the layer top is the actual cloud top or an effective cloud top. If the lidar signal becomes attenuation-limited in optically thick clouds, the cloud top is defined as the “effective” top. The effective top is identified when the signal above cloud top is below the minimum reliable signal and has a zero mean slope, and strong attenuation (according to the negative slope) is detected within the cloud.
5. Cloud base is distinguished from virga or drizzle below clouds by examining the slope below the layer base. The signal slope in virga is usually much smaller in magnitude than the slope in cloud. Dense water clouds, which often have drizzle or virga below, tend to have a large positive slope above the virga layer, followed by a strong negative slope as the signal becomes attenuated in the cloud.

Interested readers should refer to Wang and Sassen (2001), which provides figures and text describing the algorithm in detail. We currently do not apply this algorithm to lidar profiles below 500 m due to the difficulty in determining the overlap function in the near-field of the lidar. We recommend that the data user obtain cloud base height derived from ARM ceilometer datastreams (i.e., vceil25k) for heights below 500 m. Future versions of MPLCMASK will incorporate the ceilometer data into this product for a complete profile.

Depolarization Ratio

The polarization-sensitive MPLs measure a version of the depolarization ratio δ_{MPL} by alternately transmitting linearly and circularly polarized light using an actively controlled liquid crystal. Most polarization-capable lidar systems transmit strictly linear or circular polarized light. Flynn et al. (2007) describes the design and measurement techniques of the polarization MPL systems.

The MPL-measured quantities $P_{\perp}(0)$, $P_{\perp}(\pi/2)$, $P_{\parallel}(0)$, and $P_{\parallel}(\pi/2)$ are related to the diagonal elements of the scattering matrix. The parameter δ_{MPL} is related to these polarized vector elements as:

$$\delta_{MPL} = \frac{|P_{\perp}(0)|}{|P_{\perp}(\pi/2)|}, \quad (1)$$

which can be related to the linear and circular equivalents as:

$$\delta_{MPL} = \frac{1}{2} \delta_{circ} = \frac{\delta_{linear}}{(1-\delta_{linear})}. \quad (2)$$

Further details regarding the depolarization measurement are explained in Flynn et al. (2007).

3.0 The Input Data

The primary measurement required for MPLCMASK is attenuated backscatter measured by the MPL. Historically, the ARM Climate Research Facility operated single wavelength non-polarized MPLs at all fixed sites and mobile facilities. In 2004, polarized MPL systems were installed at all sites and mobile facilities. In addition, polarized MPL systems were upgraded starting in June 2010 to “fast-switching” systems to allow improved switching capability between the linear and circular polarization channels. All MPL data for all years will be processed using MPLCMASK. Depolarization ratio will only be reported when polarization-capable MPL data are available. The variables from input data that are used in this product are listed in Appendix A.

Micropulse Lidar Instruments

sssmplXX.b1 or sssmplpolavgXX.b1, sssmplpolfsC1.b1

Where:

- sss is the site and XX is the facility
- sgpmp1C1.b1 is the ingested data produced before 04/01/2004
- sgpmp1polavgC1.c1 is the VAP data produced after 04/01/2004 and before the Recovery Act instrument fast-switching upgrade to MPL (date varies)
- sgpmp1polfsC1.b1 is the ingested data produced after the upgrade to fast-switching MPL system.

NOTE: For details of the input variables see Appendix A.

Balloon-Borne Sounding System (SONDE) instruments:

ssssondewnpnXX.b1, sgpsondewnpnXX.a1

Where:

- sss is the site and XX is the facility.

NOTE: For details of the input variables see Appendix A, Table 1.

If the SONDE datastream is not available, a static molecular backscatter profile is used based on the atmospheric state variables for a standard atmosphere (see atm.dat in Appendix E).

3.1 Flow chart

This VAP has several source files and a single header file. The purpose of the header file is to define all the input fields and the output fields in a central location. This header file is included in the source file – mplcmask1zwang.c – which makes it easier to access all the fields defined in the header file. Similarly, all the external and internal functions are centrally defined in the header file.

The following steps are a brief description of the logic used in the VAP:

1. Read the input data based on the retriever file. The datastreams and the associated fields are listed in Appendix A.
2. If the MPL data are not available, the VAP will abort and no output will be generated.
3. A 30-second interval sample set is established based on the command line input.
4. Subtract background signals.
5. Establish the maximum height for the output to 20 km.
6. Compute the total backscatter by combining the co-polarized and cross-polarized channels.
i.e. Total $\beta_{\text{scat}} = \text{copol} + 2 * \text{crosspol}$
7. Apply the overlap corrections. For the fast-switching MPL systems, the corrections reside in the input file. For systems prior to 2010, the corrections reside in the \$CONF_DATA directory. The corrections are based on the serial number of the instrument (Appendix C). For MPL data prior to the polarized systems, we use static overlap corrections established as part of the MPLNOR code, which is a predecessor to this VAP.
8. Subtract energy output. These data are read from the input file.
9. Apply deadtime corrections. For the fast-switching MPL systems, the corrections reside in the input file. For systems prior to 2010, the corrections reside in the \$CONF_DATA directory. The corrections are based on the serial number of the instrument (Appendix B).
10. Calculate “beam-blocked” conditions, which occur at sites where the solar zenith angle is small enough near solar noon that the disk of the sun viewed by the detecting telescope. The telescope is covered during these times to avoid damaging the instrument. This procedure is performed primarily for MPLs located in the tropics. If the beam is blocked, the corrected MPL signals are set to missing value.
11. To compute when the MPLs are covered (beam blocked condition), the solar zenith angle, the background signal at noontime, and the average background signal at night time are calculated. If the background signal during a 4-hour window around noon is less than the average background signal at nighttime, the corrected MPL signals are set to missing. The quality check (qc) flag for linear depolarization ratio, backscatter, and cloud mask were set to the appropriate bits.
12. Calculate linear depolarization ratio (see discussion in Section 2).
13. Degrade resolution of the input time and height sample resolution to 30-second and 30-meter resolution.
14. Calculate backscatter profiles for the output resolution.
15. Calculate the signal-to-noise ratio of backscatter using a rolling window of 3 minutes without averaging the vertical. The average divided by the standard deviation of this window is the resulting signal-to-noise ratio.
16. Smooth the backscatter profiles.
17. If the sonde data are available, the Rayleigh backscatter is computed using the sonde temperature and pressure. If not, static molecular backscatter is used from standard atmosphere profiles (atm.dat file in Appendix E).
18. The cloud top, cloud base, and cloud layer are calculated using the cloud boundary algorithm (Wang and Sassen 2001).

19. A cluster test is applied to remove false identification of cloud above 10 km that occurs when the backscatter is noisy.
20. Bad lidar times are flagged as missing based on input from the mentor or translator that identifies bad lidar profiles in the raw datastream if reprocessing is necessary.
21. The cloud boundaries are then recalculated after the above tests are performed.
22. If at any time the backscatter signal is not valid, it is replaced with a missing value (-9999.0), and the QC bit is set.
23. If the backscatter signal-to-noise ratio is below 1.0 between cloud top and 30 meters below a cloud top, then the attenuation flag is set to true.
24. As a final step, a single file for each day is written as output.
25. A quicklook plot is generated for backscatter profile, linear depolarization ratio, and cloud mask.

For more details on the C and Fortran code, please see the webpage for documentation:

<https://engineering.arm.gov/~sivaraman/docs/vap/mplcmask1zwang/html/>

Quality checks are applied for five variables. Since raw MPL signals can have a valid value of zero, the calculated backscatter or the linear depolarization ratio could result in NaNs due to division by zero. The NaNs are set to missing, and the appropriate QC bits are set.

Primary Output Variables:

1. linear depolarization ratio
2. cloud_mask
3. linear depolarization signal-to-noise ratio
4. backscatter
5. backscatter signal-to-noise ratio.

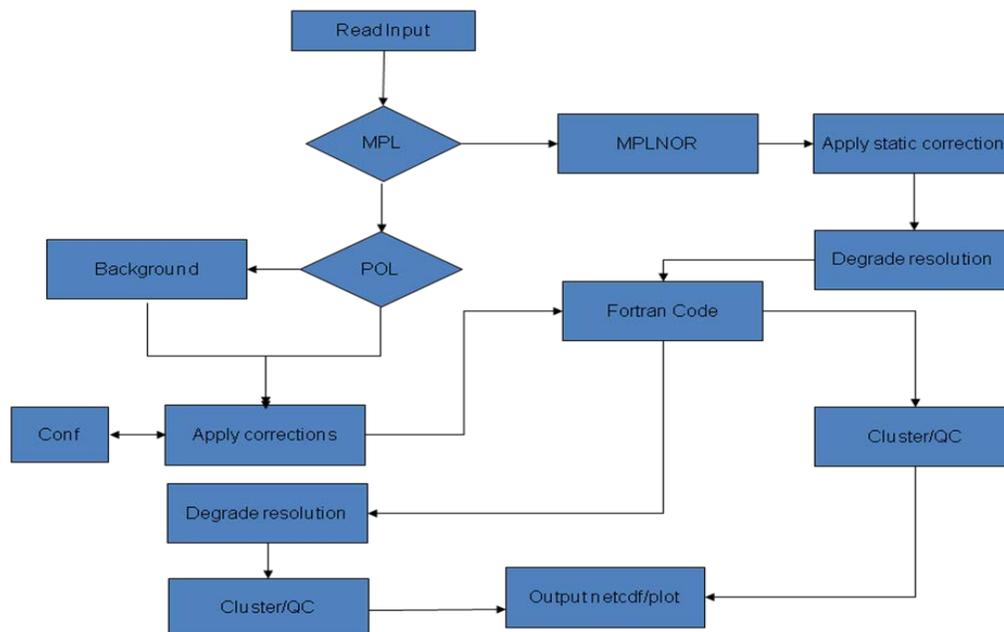


Figure 1. Flow diagram of MPLCMASK VAP.

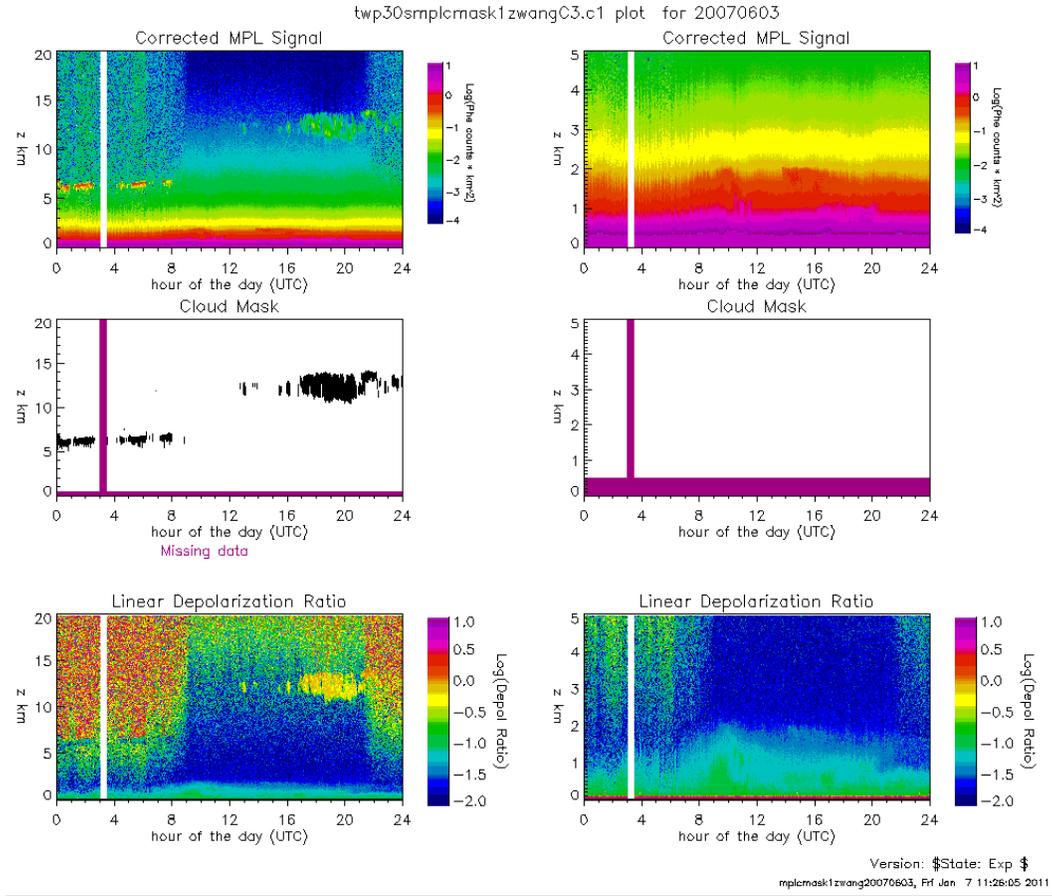


Figure 2. An example of a plot generated by the VAP.

4.0 Output Data

The MPLCMASK VAP produces one output file. The header information is available in Appendix D. The name of the output file is:

SSS30sMPLCMASK1ZWANGXX.c1.YYYYMMDD.hhmmss

Where:

- SSS is the site of the instrument (e.g., sgp)
- XX is the facility (i.e. C1, C2, etc.)
- YYYY is the year
- MM is the month of the year
- DD is the day of the month
- hh is the hour of the day
- mm is the minute of the hour
- ss is the second of the minute of data start.

5.0 Summary

This release of the MPLCMASK1ZWANG VAP includes a robust algorithm to provide cloud boundaries for cloud detected above 500 m. Since boundary-layer clouds exist below this height, we anticipate that a future update to this VAP would include additional observations (such as ceilometer data) to improve the detection of low clouds. Currently, this VAP is used as input to the ARSCL VAP, which further combines the output of MPLCMASK1ZWANG with ARM vertically pointing radar reflectivity and ceilometer cloud base heights to provide a more complete cloud mask.

6.0 References

- Wang, Z, and K Sassen. 2001. "Cloud Type and Macrophysical Property Retrieval Using Multiple Remote Sensors." *Journal of Applied Meteorology* 40: 1665–1682.
- Campbell, JR, DL Hlavka, EJ Welton, CJ Flynn, DD Turner, JD Spinhirne, VS Scott, and IH Hwang. 2002. "Full-Time, Eye-Safe Cloud and Aerosol Lidar Observation at Atmospheric Radiation Measurement Program Sites: Instruments and Data Processing." *Journal of Atmospheric and Oceanic Technology* 19: 431–442.
- Clothiaux, EE, GG Mace, TP Ackerman, TJ Kane, JD Spinhirne, and VS Scott. 1998. "An Automated Algorithm for Detection of Hydrometeor Returns in Micropulse Lidar Data." *Journal of Atmospheric and Oceanic Technology* 15: 1035–1042.



U.S. DEPARTMENT OF
ENERGY

Office of Science