THE CALIPSO LIDAR CLOUD AND AEROSOL DISCRIMINATION: VERSION 2 ALGORITHM, PERFORMANCE AND VALIDATION STRATEGIES

Zhaoyan Liu[1], Mark Vaughan[2], Chieko Kittaka[3], Ralph Kuehn[4], Ali Omar[5], and Kathleen Powell[6]

[1]*National Institute of Aerospace, 100 Exploration Way, Hampton, VA 23666, USA , zhaoyan.liu-1@nasa.gov*
[2]*NASA Langley Research Center, MS 420, Hampton, VA 23681, USA, mark.a.vaughan@nasa.gov*
[3]*SSAI/NASA Langley Research Center, MS 420, Hampton, VA 23681, USA, chieko.kittaka-1@nasa.gov*
[4]*SSAI/NASA Langley Research Center, MS 420, Hampton, VA 23681, USA, ralph.e.kuehn@nasa.gov*
[5]*NASA Langley Research Center, MS 420, Hampton, VA 23681, USA, ali.h.omar@nasa.gov*
[6]*SSAI/NASA Langley Research Center, MS 420, Hampton, VA 23681, USA, Kathleen.a.powell@nasa.gov*

ABSTRACT

The CALIPSO lidar version 2 data products have been released recently. The cloud and aerosol discrimination (CAD) is a critical step in selecting appropriate lidar ratios for the extinction retrieval. In this paper, we present an overview of the CAD algorithm, describe recent updates incorporated to increase its accuracy, and evaluate its overall performance. We also discuss the strategies for the CAD validation.

1. INTRODUCTION

The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite was launched in April 2006 [1]. The primary instrument onboard the CALIPSO payload is a two-wavelength, polarization-sensitive lidar, called the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP). CALIOP provides measurements of the atmospheric backscatter profiles at 532 nm and 1064 nm and linear depolarization ratios at 532 nm. These measurements can help greatly in understanding the role of clouds and aerosols in climate change.

Version 2 of the CALIOP data products has been released recently. Accurate cloud and aerosol discrimination (CAD) is a critical step in selecting appropriate lidar ratios for the lidar retrieval to produce extinction profiles, a key Level 2 data product. In this paper we describe the recently updated CAD algorithm and evaluate the performance of the CAD results in the version 2 data products.

2. ALGORITHM OUTLINE

The CALIOP data products include Level 1B and Level 2 products. The principal Level 1B data products are calibrated profiles of total attenuated backscatter at 532 nm and 1064 nm and the perpendicular component of the 532 nm backscatter. In the Level 2 data processing [2], features are identified in the calibrated backscatter profiles where the return signals are enhanced. These features are then classified. The CAD is the first step for the tropospheric feature classification [2]. The CAD algorithm is based on the confidence function [3]

$$f = \frac{P_{cloud}\left(\beta', \chi', z\right) - P_{aerosol}\left(\beta', \chi', z\right)}{P_{cloud}\left(\beta', \chi', z\right) + P_{aerosol}\left(\beta', \chi', z\right)}, \quad (1)$$

where P_{cloud} and $P_{aerosol}$ are, respectively, the cloud and aerosol probability density function (PDF) as a function of layer-averaged attenuated backscatter β' at 532 nm, total color ratio χ' defined as the ratio of the layer-averaged attenuated backscatters at 1064 nm and 532 nm, and altitude z. If $f > 0$, the feature is classified as cloud and, if $f < 0$, the feature is classified as aerosol. A confidence level, called the "CAD score" in the CALIOP layer products, is assigned based on the magnitude of the confidence function. The standard CAD scores reported in the CALIOP layer products range between -100 and 100. The absolute value of the CAD score provides a confidence level for the classification. The larger the magnitude of the CAD score, the higher our confidence that the classification is correct. An absolute value of 100 therefore indicates complete confidence. Absolute values less that 100 indicate some ambiguity in the classification; that is, the scattering properties of the feature are represented to some degree in both the cloud PDF and in the aerosol PDF. In this case, a definitive classification cannot be made; that is, although we can provide a "best guess" classification, this guess could be wrong, with a probability of error related to the absolute value of the CAD score. A value of 0 indicates that a feature has an equal likelihood of being a cloud and an aerosol.

The PDFs incorporated into version 2.01 of the CAD algorithm were developed based on expert manual classification of all layers detected during one full day of data acquired by CALIOP during August 2006. From these results, a single set of cloud and aerosol PDFs was constructed. This set of PDFs is applied globally for all seasons and at all latitudes. The PDFs however were binned by altitude for every kilometer from the surface to 20 km.

In the version 2 CAD algorithm, an additional test on layer volume depolarization ratio (defined as the ratio of perpendicular and parallel components of the attenuated backscatter including molecular scattering) is also conducted to help reduce misclassifications of cloud as aerosol. This is based on the fact that cirrus and dense water clouds normally have a depolarization ratio larger than aerosols. For dense water clouds, multiple scattering is significant (particularly for space lidars), and can produce large depolarization ratios. The additional CAD test is only applied to features that have been initially classified as aerosol by the *f* function. That is, if a feature is classified as aerosol by *f*, the volume depolarization ratio of this layer, δ, is then checked. If $\delta > \delta_{th}$, (δ_{th} is the threshold as defined in Figure 1), the layer is reclassified as cloud, and a value of 101 is assigned to the CAD score for this reclassified feature.

The threshold of the volume depolarization ratio is selected so that no aerosol layer is misclassified by the depolarization ratio test. Dust aerosols have a large particulate depolarization ratio due to the nonsphericity of dust particles. The value of the dust depolarization ratio measured by CALIOP is normally smaller than 0.4. The volume depolarization ratio will decrease in the presence of molecular scattering which has a very small depolarization ratio (~0.0036 for CALIOP). The smaller the dust concentration, the smaller the volume depolarization ratio. When the dust concentration is very high (e.g., over dust source regions), the volume depolarization ratio is close to the particulate depolarization ratio. For this reason, a value of 0.4 has been selected for a "dust belt", defined for the CALIOP analyses as the region between 0˚-50˚N and 40˚W-130˚E. Outside of the "dust belt", smaller threshold values have been selected. As seen in Figure 1, the threshold decreases when approaching the North and South Poles.

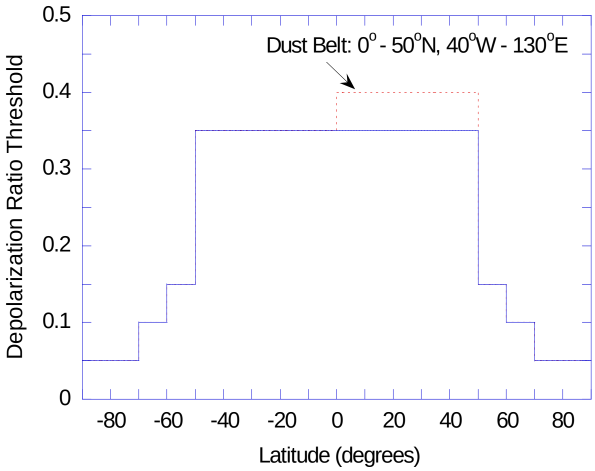


Figure 1. Latitude dependent threshold for the depolarization ratio test.

3. PERFORMANCE

Overall, the CAD algorithm works well in most cases. Manual verification of the classifications for a full day of data suggests that the success rate is in the neighborhood of 90% or better. An example of the CALIOP cloud and aerosol discrimination is given in Figure 2. The data was acquired by CALIOP on July 15, 2006, across the eastern North Atlantic along the coast of North Africa. Saharan dusts that were transported from North Africa were observed. These dust layers were easily identified from the depolarization ratio measurement (green-yellow-orange colored features lower than ~ 6km in the middle panel). On the right-hand-side of the scene, several small scale clouds are observed on the top of the dust layer, and a stratocumulus with a very large horizontal extent is seen below. In addition, high cirrus clouds, low water clouds, and marine aerosols are also present. It is seen that the CAD algorithm has identified most features correctly. In this scene, misclassifications occur mainly in the relatively dense part of the dust layer on the right-hand-side, where its optical properties are similar to what would be expected for optically thin clouds. We note that, when a cloud is adjacent to or embedded in an aerosol layer, combined ensemble can be identified as one layer. The features at the right-hand-side around 12˚N in Fig. 2 are a typical example, where the dust layer and the cloud layer beneath it are found a single layer. By convention, these mixed layers should be classified as cloud.

In general, several types of misclassifications occur with some frequency. Among these, the most prevalent are:

a) Dense aerosol layers (primarily very dense dust and smoke over and close to the source regions) are sometimes labeled as cloud, as demonstrated by the example in Figure 2. Because the CAD algorithm operates on individual layers, without a contextual awareness of any surrounding features, it can happen that small but strongly scattering regions within an extended aerosol layer will occasionally be labeled as cloud. This occurs because the optical properties (backscatter and color ratio) within the region are similar to what would be expected for the relatively faint clouds that fall within the overlap region of the PDFs. These misclassifications are often apparent from studying the Level 1 browse images. Based on the initial analysis of the CALIOP measurements, the cloud and aerosol distributions show variabilities that depend on season and on geophysical location. The globally averaged PDFs used in the current release will have a larger overlap between the cloud and aerosol than would occur for more regionally specific statistics. For future versions of the CAD algorithm, we expect to develop and deploy PDFs that will correctly reflect both seasonal and latitudinal variations.

b) Many optically thin clouds, both ice and water, are encountered in the polar regions. The current CAD PDFs do not work as well in the polar regions as they do at lower latitudes, and thus misclassifications of clouds as aerosol are more common. In particular, thin ice clouds, which can extend from the surface to several kilometers in altitude in the polar regions, are sometimes misclassified as aerosol.

c) Correct classification of heterogeneous layers is always difficult, and the process can easily go awry. An example of a heterogeneous layer would be an aerosol layer that is vertically adjacent to a cloud or contains an embedded cloud, but which is nonetheless detected by the feature finder as a single entity. We note that, by convention, heterogeneous layers should be classified as clouds. However, depending on the relative strengths of the components, these layers are sometimes erroneously identified as aerosol.

d) Some so-called features identified by the layer detection scheme are not legitimate layers, but instead are artifacts due to the noise in the signal, multiple scattering effects, or to artificial signal enhancements caused by non-ideal detector transient response or an over estimate of the attenuation due to overlying layers. These erroneous “pseudo-features” are neither cloud nor aerosol; however, because they are not properly interdicted in the processing stream, the CAD algorithm nonetheless attempts to assign them to one class or the other. Very frequently these layers can be identified by their very low CAD scores (typically less than 20, depicted by the red color in the lower panel in Fig. 2).

4. VALIDATION STRATAGIES

Although attempts have been made to compare the CALIOP measurements with model studies (e.g., [4]), which are useful to help interpret/validate the CALIOP scene classification, a systematic plan to validate the CALIOP scene classification still needs to be developed. It is now time to develop specific strategies for validating the CALIOP cloud and aerosol discrimination. Below we present several initial ideas that we hope will provoke an open discussion to that will quickly yield a mature and feasible CAD validation plan.

a) Model analyses have shown the ability to help interpret the CALIOP measurement [4] and [5]. As mentioned earlier, dust is an aerosol type that can be misclassified as cloud when the layer is especially dense. The identification of dusts is further complicated by the presence of clouds. Asian dust events are normally transported by cold fronts, and transported dusts are frequently adjacent to, embedded or mixed with clouds [5]. When dust plumes are very dense, it is hard to discriminate dust from cloud based solely on the CALIOP measurement. In such case, the dust model analysis can help to identify dusts.

b) There are numerous ground-based lidar networks located in the US, Europe and Asia. These lidar networks have acquired large amounts of long term observations. Statistical comparisons of the CALIOP measurements with PDFs developed from these ground-based lidar observations will provide a great opportunity to validate the CALIOP cloud and aerosol classification. Parameters that are desired for comparison include integrated attenuated backscatter at 532 nm and 1064 nm, color ratio, depolarization ratio, layer top and base altitudes, layer thickness, top, base and mid-layer temperature, etc. All of this information would also useful in modifying and improving the CALIOP CAD algorithm.

5. SUMMARY

The CALIOP version 2 data products have been released recently. Modifications have been made to the CAD algorithm. These modifications include mainly the development of a new set of PDFs based on one full day of the CALIOP measurements and the inclusion of depolarization ratio test.

The performance of the revised CAD algorithm has been presented. Overall, the CAD algorithm works quite well in most cases. Manual verification of the classifications for a full day of data suggests that the success rate is in the neighborhood of 90% or better. Nevertheless, several types of misclassifications occur with some frequency, and these are enumerated and explained.

We have also presented preliminary ideas for validating the CALOP cloud and aerosol classification to initiate an open discussion to yield a systematic validation plan.

REFERENCES

[1] Winker, D., W. Hunt, and M. McGill, 2007: Initial performance assessment of CALIOP, *Geophys. Res. Lett.*, **34**, L19803, doi:10.1029/2007GL030135.

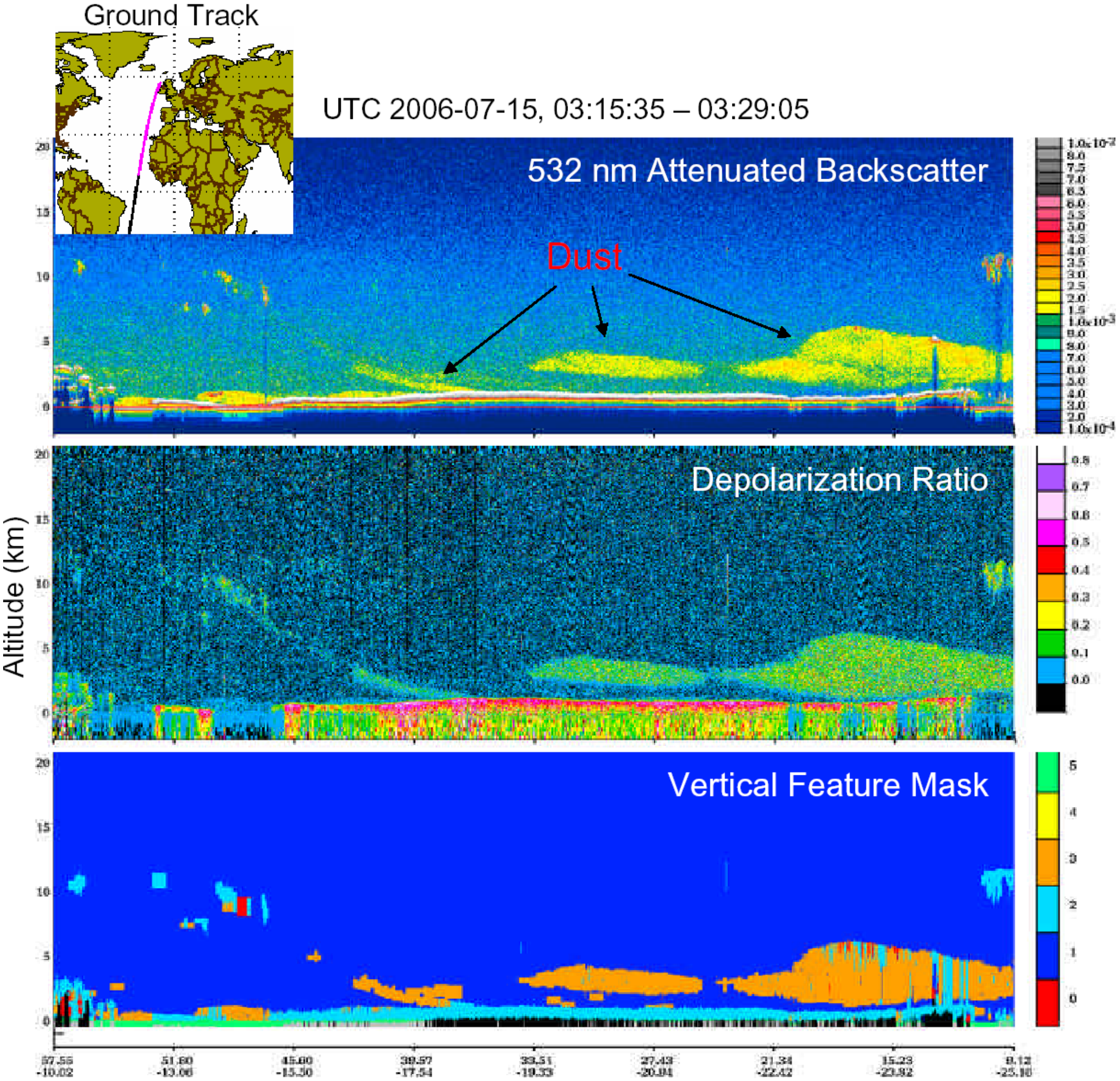
[2] Vaughan, M., S. Young, D. Winker, K. Powell, A. Omar, Z. Liu, Y. Hu, and C. Hostetler, 2004,: Fully automated analysis of space-based lidar data: an overview of the CALIPSO retrieval algorithms and data products, *Proc. SPIE*, **5575**, 16-30.

[3] Liu, Z., M. Vaughan, D. Winker, C. Hostetler, L. Poole, D. Hlavka, W. Hart, and M. McGill, 2004: Use of probability distribution functions for discriminating between cloud and aerosol in lidar backscatter data. *J.*

Geophys. Res., 109, D15202, doi:10.1029/2004JD004732.

[4] Young, S., M. Vaughan, A. Omar, Z. Liu, S. Lee, Y. Hu, and M. Cope, 2008: CALIPSO satellite lidar identification of elevated dust over Australia compared with air quality model PM60 forecasts. Current proceedings.

[5] Uno, I., K. Yumimoto, A. Shimizu, Y. hara, N. Sugimoto, Z. Wang, Z. Liu, and D. Winker, 2008: 3D structure of Asian dust transport revealed by CALIPSO lidar and a 4D VAR dust model. *Gephys. Res. Lett.*, in press.



0 = invalid (bad, missing data or CAD < 20); 1 = clear air; 2 = cloud; 3 = aerosol;
4 = stratospheric feature; 5 = surface

Figure 2. Example of the CALIPSO lidar scene classification.