CALIPSO Quality Statements: Lidar Level 2 Cloud and Aerosol Layer Products Version Releases: 2.01, 2.02

Introduction | Documents and References | Data Products Descriptions | Data Release Versions

---

Introduction

This document provides a high-level quality assessment of the cloud and aerosol layer products derived from the CALIPSO lidar measurements, as described in Section 2.4 of the CALIPSO Data Products Catalog (Version 2.4) (PDF). As such, it represents the minimum information needed by scientists and researchers for appropriate and successful use of these data products. We strongly suggest that all authors, researchers, and reviewers of research papers review this document periodically, and familiarize themselves with the latest status before publishing any scientific papers using these data products.

These data quality summaries are published specifically to inform users of the accuracy of CALIOP data products as determined by the CALIPSO Science Team and Lidar Science Working Group (LSWG). This document is intended to briefly summarize key validation results; provide cautions in those areas where users might easily misinterpret the data; supply links to further information about the data products and the algorithms used to generate them; and offer information about planned algorithm revisions and data improvements.

The primary new parameters included in the version 2.0 release of the Cloud and Aerosol Layer Products are layer optical depth, aerosol type, and cloud ice/water phase. Although extinction and/or optical depths appear in several different products, all extinction retrievals are produced by the same algorithm. **PLEASE NOTE**: users of the CALIOP extinction and backscatter profile data should read and thoroughly understand the information provide in the Profile Products Data Quality Summary. This summary contains an expanded description of the extinction retrieval process from which the layer optical depths are derived, and provides essential guidance in the apppropriate use of all CALIOP extinction-related data products.

Additional Documentation and References

Algorithm Theoretical Basis Documents (ATBDs)

- PC-SCI-202.01 - Mission, Instrument, and Algorithms Overview (PDF)
- PC-SCI-202.02 - Feature Detection and Layer Properties Algorithms (PDF)
- PC-SCI-202.03 - Scene Classification Algorithms (PDF)
- PC-SCI-202.04 - Extinction Retrieval Algorithms (PDF)

General References

- PC-SCI-503 : CALIPSO Data Products Catalog (Version 2.4) (PDF)
- Data analysis overview: Fully automated analysis of space-based lidar data: an overview of the CALIPSO retrieval algorithms and data products (PDF)
- Additional publications (journal articles and conference proceedings about CALIPSO science, algorithms, and data processing)
- CALIPSO Data Read Software

CALIPSO Cloud and Aerosol Layer Products

Each of the CALIPSO layer products contains a sequence of two tightly coupled data types. The first of these is a set of column properties, which describe the temporal and geophysical location of the vertical column (or curtain) of atmosphere being sampled. Column properties include satellite position data and viewing geometry, information about the surface type and lighting conditions, and the number of features (e.g., cloud and/or aerosol layers) identified within the column. For each set of column properties, there is an associated set of layer properties . These layer properties specify the spatial and optical characteristics of each feature found, and include quantities such as layer base and top altitudes, integrated attenuated backscatter, layer-integrated volume depolarization ratio, and optical depth. Below we provide brief descriptions of each of the column properties and the layer properties. Where appropriate, we also provide an assessment of the quality and accuracy of the data in the current release.

The layer products are generated at three different spatial resolutions.

- The *1/3 km layer products* report cloud detection information obtained at the highest spatial resolution of the lidar: 1/3 km horizontally and 30-m vertically. Due to constraints on CALIPSO's downlink bandwidth, this full resolution data is only available from ~8.3 km above mean sea level, down to -0.5 km below sea level.

- The *1 km layer products* report cloud detection information obtained at a horizontal resolution of 1 km, over a vertical range extending from ~20.2 km above mean sea level, down to -0.5 km below sea level.
- The *5 km layer products* report (separately) cloud and aerosol detection information on a 5 km horizontal grid. At present there is no separate stratospheric data product. Stratospheric features are recorded in the 5 km aerosol product.

The fundamental data product provided by the CALIPSO layer products is the vertical location of cloud and aerosol layer boundaries. All other layer properties -- e.g., integrated attenuated backscatters and layer two-way transmittances -- are computed with reference to these boundaries. To make proper use of the CALIPSO layer products, all users must be aware of the uncertainties inherent in the fully automated recognition of layer boundaries. Note too that **clouds and aerosols are reported separately** in the CALIPSO layer products. Therefore, to obtain a complete representation of all features detected within any region, users must use both the cloud and the aerosol layer products.

Column Properties: Data Release 2.01

Profile ID

The lidar profile ID is a 32-bit integer generated sequentially for each single-shot profile record. Each profile ID is unique within each granule. Profile IDs reported in the 1/3 km layer products are for the individual laser pulses from which the layer statistics were derived. Profile IDs reported in the 1 km layer products designate the profile at the temporal midpoint of the three laser pulses averaged to generate the 1 km horizontal resolution. For the 5 km layer products, three values are reported: the profile ID for the first pulse included in the 15 shot average; the profile ID for the final pulse; and the profile ID at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots).

Latitude

Geodetic latitude, in degrees, of the laser footprint. Latitudes reported in the 1/3 km layer products are for the individual laser pulses from which the layer statistics were derived. The latitudes reported in the 1 km layer products represent footprint latitude at the temporal midpoint of the three laser pulses averaged to generate the 1 km horizontal resolution. For the 5 km layer products, three values are reported: the footprint latitude for the first pulse included in the 15 shot average; the footprint latitude for the final pulse; and the footprint latitude at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots).

Longitude

Longitude, in degrees, of the laser footprint. Longitudes reported in the 1/3 km layer products are for the individual laser pulses from which the layer statistics were derived. The longitudes reported in the 1 km layer products represent footprint longitude at the temporal midpoint of the three laser pulses averaged to generate the 1 km horizontal resolution. For the 5 km layer products, three values are reported: the footprint longitude for the first pulse included in the 15 shot average; the footprint longitude for the final pulse; and the footprint longitude at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots).

Profile Time TAI

Time expressed in International Atomic Time (TAI). Units are in seconds, starting from January 1, 1993. Times reported in the 1/3 km layer products are for the individual laser pulses from which the layer statistics were derived. Times reported in the 1 km layer products represent the temporal midpoint of the three laser pulses averaged to generate the 1 km horizontal resolution. For the 5 km layer products, three values are reported: the time for the first pulse included in the 15 shot average; the time for the final pulse; and the time at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots).

Profile Time UTC

Time expressed in Coordinated Universal Time (UTC), and formatted as 'yymmdd.ffffffff', where 'yy' represents the last two digits of year, 'mm' and 'dd' represent month and day, respectively, and 'ffffffff' is the fractional part of the day. Times reported in the 1/3 km layer products are for the individual laser pulses from which the layer statistics were derived. Times reported in the 1 km layer products represent the temporal midpoint of the three laser pulses averaged to generate the 1 km horizontal resolution. For the 5 km layer products, three values are reported: the time for the first pulse included in the 15 shot average; the time for the final pulse; and the time at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots).

Day Night Flag

Indicates the lighting conditions at an altitude of ~24 km above mean sea level; 0 = day, 1 = night.

Off Nadir Angle

The angle, in degrees, between the viewing vector of the lidar and the nadir angle of the spacecraft. Beginning in June 2006, CALIPSO operated with the lidar pointed at 0.3 degrees off-nadir (along track in the forward direction), with the exception of November 6-17, 2006 and August 21 to September 7, 2007. During these periods, CALIPSO operated with the lidar pointed at 3.0 degrees off nadir. Beginning November 28, 2007, the off-nadir angle was permanently changed to 3.0 degrees.

Solar Zenith Angle

The angle, in degrees, between the zenith at the lidar footprint on the surface and the line of sight to the sun.

Solar Azimuth Angle

The azimuth angle, in degrees, from north of the line of sight to the sun.

Scattering Angle

The angle, in degrees, between the lidar viewing vector and the line of sight to the sun.

Parallel Column Reflectance 532

Bi-directional column reflectance derived from the root-mean-square (RMS) variation of the 532 nm parallel channel background measurements. For the 1/3-km layer products, single shot values are reported; for the 1-km and 5-km layer products, mean values are reported.

Parallel Column Reflectance Uncertainty 532

Not calculated for the current release; data products contain fill values in this field.

Parallel Column Reflectance RMS Variation 532 (5 km products only)

The RMS variation of the the parallel channel reflectance values computed using the 15 samples that comprise a nominal 5-km horizontal swath of CALIOP lidar measurements.

Perpendicular Column Reflectance 532

Bi-directional column reflectance derived from the RMS variation of the 532 nm perpendicular channel background measurements. For the 1/3-km layer products, single shot values are reported; for the 1-km and 5-km layer products, mean values are reported.

Perpendicular Column Reflectance Uncertainty 532

Not calculated for the current release; data products contain fill values in this field.

Perpendicular Column Reflectance RMS Variation 532 (5 km products only)

The RMS variation of the the perpendicular channel reflectance values computed using the 15 samples that comprise a nominal 5-km horizontal swath of CALIOP lidar measurements.

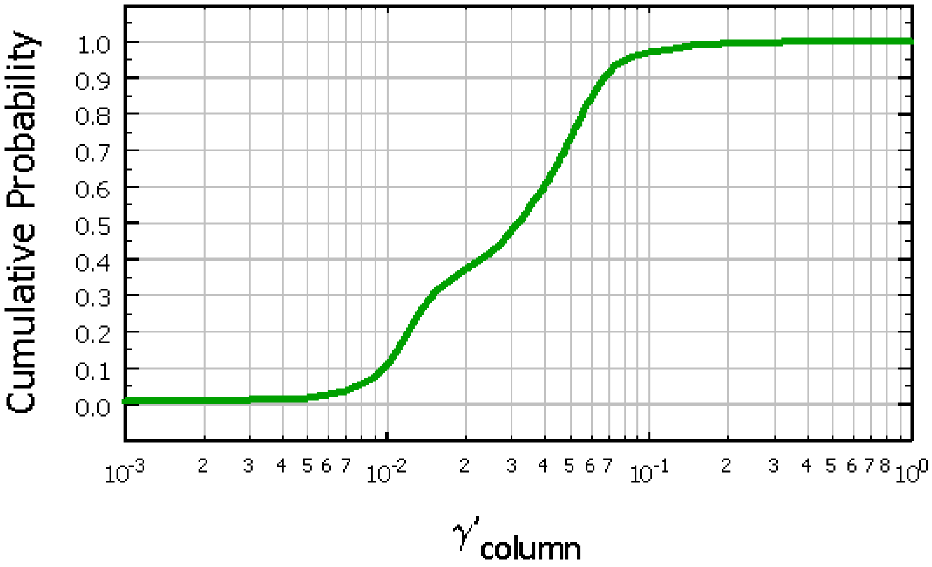
Column integrated attenuated backscatter (IAB) 532

The integral with respect to altitude of the 532 nm total attenuated backscatter coefficients. The limits of integration are from the onset of the backscatter signal at ~40-km, down to the range bin immediately prior to the surface elevation specified by the digital elevation map. This quantity represents the total attenuated backscatter measured within a column. Physically meaningful values of the column integrated attenuated backscatter (hereafter, γ'_{column}) range from ~0.01 sr (completely clear air), to greater than 1.5 sr (e.g., due to anomalous backscatter from horizontally oriented ice crystals; see Hu et al. (Optics Express **15**, 2007)).

Column IAB cumulative probability

The cumulative probability of measuring a total column integrated attenuated backscatter value equal to the value computed for the current profile. Values in this field range between 0 and 1. The cumulative probability distribution function, shown below in Figure 1, was compiled using all CALIOP total column IAB measurements acquired between 15 June, 2006 and 18 October, 2006.

Figure 1: Distribution of γ'_{column} at 532 nm



Tropopause Height

Tropopause height, in kilometers above local mean sea level; derived from the GEOS-5 data product provided to the CALIPSO project by the GMAO Data Assimilation System.

Tropopause Temperature

Tropopause temperature, in degrees C; derived from the GEOS-5 data product provided to the CALIPSO project by the GMAO Data Assimilation System.

IGBP Surface Type

International Geosphere/Biosphere Programme (IGBP) classification of the surface type at the lidar footprint. The IGBP surface types reported by CALIPSO are the same as those used in the CERES/SARB surface map.

NSIDC Surface Type

Snow and ice coverage for the surface at the lidar footprint; data obtained from the National Snow and Ice Data Center (NSIDC).

Lidar Surface Elevation

Surface elevation at the lidar footprint, in kilometers above local mean sea level, determined by analysis of the lidar backscatter signal; see section 7.3 of the CALIPSO CALIPSO Feature Detection ATBD (PDF). The 1/3 km and 1 km layer products report the base and top of the detected surface spike. The 5 km layer products report statistics (minimum, maximum, mean, and standard deviation for both the upper and lower boundaries of the surface echo) derived from an analysis of the 1 km signal. If the surface is detected at the 5 km resolution but not at 1 km, only the maximum and minimum values are reported for each boundary. If no surface is detected, this field will contain fill values.

In the very best case, lidar surface elevations are as reliable as the DEM. GTOPO30 tends to be very reliable over oceans, and considerably less so in rugged terrain such as in the Andes mountains of Peru. However, due to aberrations in the signal caused by a non-ideal transient response in the 532 nm detectors, the geometric thickness associated with the lidar surface elevation (i.e., surface top - surface base) can be extremely misleading. **IMPORTANT**: At present, users should treat **ALL** signal beneath the reported lidar surface elevation top as being pure instrument artifact introduced by the non-ideal transient response of the detectors. No geophysical significance should be attributed to the (apparent!) subsurface portion of the lidar return.

DEM Surface Elevation

Surface elevation at the lidar footprint, in kilometers above local mean sea level, obtained from the GTOPO30 digital elevation map (DEM).

Surface Elevation Detction Frequency (5 km products only)

A bit-mapped 8-bit integer that reports both the horizontal averaging resolution at which the surface was originally detected and, where applicable, the frequency with which the surface was subsequently detected at the 1-km averaging resolution. Bit interpretation is as follows.

Bits 1, 2, and 3 indicate the horizontal resolution at which the surface was detected:

- 0 = not detected
- 1 = detected at 1/3-km averaging
- 2 = detected at 1-km averaging
- 3 = detected at 5-km averaging
- 4 = detected at 20-km averaging
- 5 = detected at 80-km averaging

Bits 4 and 5 are not used and are set to zero. Taken together, bits 6, 7, and 8 report the 5-km detection frequency:

- 0 = detection frequency = 0%
- 1 = detection frequency = 20%
- 2 = detection frequency = 40%
- 3 = detection frequency = 60%
- 4 = detection frequency = 80%
- 5 = detection frequency = 100%

Normalization Constant Uncertainty 532 (5 km products only)

Not calculated for this data release; data products contain fill values in this field.

Calibration Altitude 532 [provisional; 5 km products only]

Top and base altitudes, in kilometers above mean sea level, of the region of the atmosphere used for calibrating the 532 nm parallel channel. The calibration algorithm and procedures are explained in detail in the CALIOP Level 1 ATBD (PDF).

Feature Finder QC Flags [provisional; 5 km products only]

To generate data at a nominal 5 km horizontal resolution requires averaging 15 consecutive laser pulses. For each 5 km average, we report a set of feature finder QC flags. Conceptually, these flags are a set of 15 Boolean values which tell the user whether or not a feature (cloud, aerosol, or surface echo) was detected in each of the 15 laser pulses. The flags are implemented as a 16-bit integer. The most significant bit is unused, and always set to zero. Each of the 15 remaining bits represents the "features found" state for a single full-resolution profile. A bit value of zero indicates that one or more features were found within the profile. A feature finder QC flag value of zero for any 5 km column indicates complete feature finder success.

Number Layers Found [provisional]

The number of layers found in this column; cloud data products report (only) the number of cloud layers found, and aerosol report (only) the number of aerosol layers found.

Surface Wind Speed [provisional; 5 km aerosol products only]

Zonal and meridional surface wind speeds, in meters per second, obtained from the GEOS-5 data product provided to the CALIPSO project by the GMAO Data Assimilation System.

Layer Properties: Data Release 2.01

Layer Top Altitude and **Layer Base Altitude** [provisional]

Layer top and base altitudes are reported in units of kilometers above *mean sea level*. Due to the on-board data averaging scheme, the precision with which CALIPSO can make this measurement is itself a function of altitude. Between -0.5 km and ~8.2 km, the vertical resolution of the lidar is 30-meters. From ~8.2 km to ~20.2 km, the vertical resolution of the lidar is 60-meters. Above ~20.2 km, the vertical resolution is 180-meters.

The uncertainties associated with detection of cloud and aerosol layers in backscatter lidar data are examined in detail in Section 5 of the CALIPSO Feature Detection ATBD (PDF). The ATBD contains quantitative assessments of feature finder performance derived using simulated data sets, for which all layer boundaries were known exactly. In the real world of layer detection, we do not have access to this underlying truth. Therefore in this document we provide the following set of "rules of thumb" that users can apply to the data products to obtain a qualitative understanding of the layer boundaries reported, and of the optical properties associated with these layers.

a. Strongly scattering features are easier to detect than weakly scattering features. The scattering intensity of each layer is reported in the 532 nm and 1064 nm attenuated backscatter statistics and by the integrated attenuated backscatter at 532 nm and 1064 nm.

b. Detection of layers during the nighttime portion of the orbits is more reliable than during the daytime portion of the orbits. Due to solar background signals, the noise levels in the daytime measurements are much larger than those at night, and this additional noise can obscure faint features, and can lead to boundary detection errors even in more strongly scattering layers.

c. Features become increasingly difficult to detect with increasing optical depth above feature top. Put another way, detection of the lower layers in a multi-layer scene is made more difficult by the signal losses that occur as the laser light passes through the upper layers. (In a sense, this is a restatement of (a), since the backscatter intensity of secondary features is reduced from what it otherwise might be by the signal attenuation caused by the overlying features.)

d. In general, our confidence in the location of the top of a layer is somewhat greater than our confidence in the location of the base of the same layer. For transmissive features, one reason for this is that the backscatter signal is attenuated by traversing the feature, thus degrading the potential contrast between feature and "non-feature" at the base. Additionally, in strongly scattering layers, multiple scattering effects and signal perturbations introduced by the non-ideal transient response of the 532 nm detectors can also make base determination less certain.

e. For opaque features that completely attenuate the backscatter signal, the base altitude reported must be considered as an "apparent" base rather than a true base.

f. Stratospheric features reported during daylight -- especially those reported between 60N and 60S -- should be treated with extreme suspicion.

Opacity Flag [provisional; 5 km products only]

In the context of the 5-km CALIOP layer products, a layer is considered opaque if (a) it is the lowest feature detected in a column, and (b) it is not subsequently classified as a surface return. An opacity flag value of 1 indicates an opaque layer; values of 0 indicate transparent layers.

Users should be aware that the opacity flag *does not* indicate that an individual layer is opaque. Instead, it identifies the layer in which the backscatter signal becomes completely attenuated (i.e., indistinguishable from the background signal level), so that for those features having an opacity flag of 1, the reported base altitude must be considered as an apparent base, rather than a true base. Furthermore, as noted in section 7.4 of the CALIPSO Feature Detection ATBD (PDF), the identification of any layer as "opaque" depends on the amount of averaging applied to the signal prior to initiating the layer detection algorithm. Data users intending to employ the opacity flag in their analyses are strongly advised to consult the feature detection ATBD (PDF).

Because all features reported in 1/3-km and 1-km layer products are detected at a single horizontal averaging resolution (i.e., either at 1/3-km or 1-km), the opacity flag is not reported. When using these products, opacity, in the sense described above, can be assessed as follows. If the surface was detected (i.e., the lidar surface altitude field does not contain fill values) then there are no opaque layers in the column. If the surface was not detected, then the lowest layer in the column is considered to be opaque.

Horizontal Averaging [provisional; 5 km products only]

The amount of horizontal averaging required for a feature to be detected. For all data versions up to and including the 2.01 release, the values in this field will be either 0, 5, 20, or 80. 0 is a fill value; the remaining values indicate features detected at 5-km, 20-km, and 80-km averaging intervals, respectively.

Integrated Attenuated Backscatter 532 [provisional]

The 532 nm integrated attenuated backscatter (hereafter, γ'_{532}) for any layer is computed according to equation 3.14 in section 3.2.9.1 of the CALIPSO Feature Detection ATBD (PDF). The values reported for γ'_{532} will always be positive.

For the uppermost layer in any column, the quality of the estimate for γ'_{532} is determined by the accuracy of the top and base identification, the reliability of the 532 nm channel calibrations, and by the signal-to-noise ratio (SNR) of the backscatter data within the layer. For layers beneath the uppermost, the quality of our estimate for γ'_{532} also depends on either obtaining an independent estimate of the two-way transmittance, T^2, for all overlying layers, or by estimating this quantity directly from the lidar backscatter data. Estimating T^2 directly from the data is something of a black art. In tractable situations (i.e., where there exists an extended region of "clear air" between successive layers, and where the uppermost layer has no more than a moderate optical depth of -- say -- 1.0 or less), the calculation can be fairly reliable. In especially awkward situations (e.g., vertically adjacent layers, such as clouds embedded in aerosols), the only way to estimate T^2 is to compute a full extinction retrieval for the profile being examined. Furthermore, the effects of errors caused by misestimating T^2 can increase sharply as the optical thickness above a layer increases. We note that the CALIOP processing scheme ***always*** attempts to correct estimates of γ'_{532} for the attenuation imparted by previously identified overlying features. As a consequence, we will occasionally report unrealistically large values for γ'_{532}.

Integrated Attenuated Backscatter Uncertainty 532 [provisional]

The uncertainties reported for the 532 nm integrated attenuated backscatters provide an estimate of the random error in the backscatter signal. The general procedure used for calculating uncertainties for integrated quantities is described by Liu et al., 2006 (PDF). The specific formula is given by equation 6.7 in the CALIPSO Feature Detection ATBD (PDF).

There are occasions (e.g., in regions of especially low SNR) where the uncertainty calculation can fail. In these cases, the value recorded in the data product will be set to negative one (-1). In all other cases, uncertainty values will be positive.

Attenuated Backscatter Statistics 532 [provisional]

This field reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the 532 nm attenuated backscatter coefficients for each layer. Formulas used for each of the statistical calculations can be found in section 6 of the CALIPSO Feature Detection ATBD (PDF).

Integrated Attenuated Backscatter 1064 [provisional]

The 1064 nm integrated attenuated backscatter (hereafter, γ'_{1064}) for any layer is computed according to equation 6.6 in section 6.5 of the CALIPSO Feature Detection ATBD (PDF).

As is the case for γ'_{532}, in the uppermost layer within any column, the quality of the estimate for γ'_{1064} is determined by the accuracy of the top and base identification, the reliability of the 1064 nm calibration constant, and by the signal-to-noise ratio (SNR) of the backscatter data within the layer. In layers beneath the uppermost, γ'_{1064} will be underestimated by a factor equal to the total particulate two-way transmittance, T^2, above the layer. In contrast to the techniques applied at 532 nm, reliable estimates of T^2 cannot be derived from an analysis of the 1064 nm backscatter signal in the (assumed to be) clear air regions.

The CALIOP layer detection algorithm operates exclusively on the 532 nm backscatter signals. Users should thus be aware that, unlike γ'_{532}, negative (i.e., non-physical) values can occasionally be reported for γ'_{1064}. This situation occurs most often for very weak features and in those layers for which the backscatter signal has been highly attenuated by other, overlying layers.

Integrated Attenuated Backscatter Uncertainty 1064 [provisional]

The uncertainties reported for the 1064 nm integrated attenuated backscatter values provide an estimate of the random error in the backscatter signal. The general procedure used for calculating uncertainties for integrated quantities is described by Liu et al., 2006 (PDF). The specific formula is given by equation 6.7 in the CALIPSO Feature Detection ATBD (PDF).

There are occasions (e.g., in regions of especially low SNR) where the uncertainty calculation can fail. In these cases, the value recorded in the data product will be set to negative one (-1). In all other cases, uncertainty values will be positive.

Attenuated Backscatter Statistics 1064 [provisional]

This field reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the 1064 nm attenuated backscatter coefficients for each layer. Formulas used for each of the statistical calculations can be found in section 6 of the CALIPSO Feature Detection ATBD (PDF).

Integrated Volume Depolarization Ratio [provisional]

The layer integrated 532 nm volume depolarization ratio (hereafter, δ_{layer}) is computed according to equation 6.10 in section 6.7 of the CALIPSO Feature Detection ATBD (PDF).

The quality of the estimate for δ_{layer} is determined by the accuracy of the top and base identification, the reliability of the polarization gain ratio calibration, and by the signal-to-noise ratio (SNR) of the backscatter data within the layer. In general, the CALIOP δ_{layer} estimates are highly reliable. Histograms of δ_{layer} compiled for midlatitude cirrus in the northern hemisphere compare very well with previously reported distributions, e.g., Sassen & Benson, 2001 (PDF).

Integrated Volume Depolarization Ratio Uncertainty [provisional]

The uncertainties reported for the 532 nm layer-integrated volume depolarization ratios provide an estimate of the total random error in the combined backscatter signals (i.e., the 532 nm parallel and perpendicular signals within the feature). The general procedure used for calculating uncertainties for integrated quantities is described by Liu et al., 2006 (PDF). The specific formula is given by equation 6.11 in the CALIPSO Feature Detection ATBD (PDF).

There are occasions (e.g., in regions of especially low SNR) where the uncertainty calculation can fail. In these cases, the value recorded in the data product will be set to negative one (-1). In all other cases, uncertainty values will be positive.

Volume Depolarization Ratio Statistics [provisional]

This field reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the 532 nm volume depolarization ratios for each layer. Formulas used for each of the statistical calculations can be found in section 6 of the CALIPSO Feature Detection ATBD (PDF).

In regions with acceptable SNR, the accuracy with which the range resolved depolarization ratios can be determined will depend almost entirely on the accuracy of the polarization gain ratio calibration.

Users can have high confidence in the *calculation* of all of the values in the depolarization ratio statistics fields. However, the meaning of these numbers can be somewhat obscure. This is because each of the range resolved depolarization ratios within any layer is the ratio of two noisy numbers. Especially where the feature is relatively faint, and in regions of low SNR, data values in both the numerator (the 532 nm perpendicular channel) and the denominator (the 532 nm parallel channel) can randomly and independently approach zero, which in turn can generate extremely large or extremely small (and even non-physical) depolarization ratios. When computing layer means, standard deviations, and centroids, these values can dominate the calculation, and thus return entirely unrealistic estimates. When assessing the depolarization ratio that characterizes a layer, δ_{layer} and the layer median are both more reliable indicators than the mean.

Integrated Attenuated Total Color Ratio [provisional]

The layer integrated attenuated total color ratio (hereafter, χ'_{layer}) is computed according to equation 6.13 in section 6.7 of the CALIPSO Feature Detection ATBD (PDF).

The quality of the estimate for χ'_{layer} is determined by the accuracy of the top and base identification, the reliability of the 532 nm calibration constant and the 1064 nm calibration constant, and by the signal-to-noise ratio (SNR) of the backscatter data within the layer.

Integrated Attenuated Total Color Ratio Uncertainty [provisional]

The uncertainties reported for the layer-integrated attenuated total color ratios provide an estimate of the total random error in the combined backscatter signals (i.e., at 532 nm and 1064 nm). The general procedure used for calculating uncertainties for integrated quantities is described by Liu et al., 2006 (PDF). The specific formula is given by equation 6.14 in the CALIPSO Feature Detection ATBD (PDF).

There are occasions (e.g., in regions of especially low SNR) where the uncertainty calculation can fail. In these cases, the value recorded in the data product will be set to negative one (-1). In all other cases, uncertainty values will be positive.

Attenuated Total Color Ratio Statistics [provisional]

This field reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the attenuated total color ratios for each layer. Formulas used for each of the statistical calculations can be found in section 6 of the CALIPSO Feature Detection ATBD (PDF).

Users can have high confidence in the *calculation* of all of the values in the attenuated total color ratio statistics fields. However, as with the 532 nm depolarization ratio statistics, the meaning of the various numbers can be somewhat misleading. Like the depolarization ratios, the attenuated total color ratios are produced by dividing one noisy number (the 1064 nm attenuated backscatter coefficient) by a second noisy number (the 532 nm attenuated backscatter coefficient). Depending on the noise in any pair of samples, the resulting values can range from large negative values to extremely large positive values. When computing layer means, standard deviations, and centroids, these outliers can dominate the calculation, and thus return entirely unrealistic estimates.

Overlying Integrated Attenuated Backscatter (IAB) 532 [provisional]

Similar to the column integrated attenuated backscatter, the overlying integrated attenuated backscatter (hereafter, γ'_{above}) is the integral with respect to altitude of the 532 nm total attenuated backscatter coefficients. The upper limit of integration is once again the first range bin in the measured signal profile, and the lower limit is now the range bin immediately above the layer top altitude.

γ'_{above} provides a qualitative assessment of the confidence that users should assign to each layer reported. As noted earlier (see the discussion for layer base and top heights), layer detection, and the assessment of the associated layer descriptors, becomes increasingly uncertain as the overlying optical depth increases. This uncertainty cannot be easily quantified, because backscatter lidars such as CALIOP cannot measure optical depth directly, and must instead derive optical depth estimates in subsequent data processing. However, γ'_{above} can easily be obtained directly from the calibrated backscatter signal, and hence can provide a qualitative proxy for the optical depth above each layer detected.

Layer IAB QA Factor [provisional]

The single layer analog of the column IAB cumulative probability; the layer IAB QA factor is defined as 1 - F(γ'_{above}), where F(γ'_{above}) is the cumulative probability of measuring a complete column integrated attenuated backscatter equal to γ'_{above}.

Feature Classification Flags [beta]

For each layer, we report a set of feature classification flags that provide assessments of (a) feature type (e.g., cloud vs. aerosol vs. stratospheric layer); (b) feature subtype; (c) layer ice-water phase (clouds only); and (d) the amount of horizontal averaging required for layer detection. The complete set of flags is stored as a single 16-bit integer. A comprehensive description of the feature finder classification flags, including their derivation and physical significance, quality assessments, and guidelines for interpreting them in computer codes, can be found in the documentation for the vertical feature mask data product.

Correct interpretation of the feature subtype bits depends on the status of the feature type; e.g., the interpretation is different for clouds and aerosols. For aerosols, the feature subtype is one of eight types: desert dust, biomass burning, background, polluted continental, marine, polluted dust, other, and 'not determined'. Desert dust is mostly mineral soil. Biomass burning is an aged smoke aerosol consisting primarily of soot and organic carbon (OC), clean continental (also referred to as background or rural aerosol) is a lightly loaded aerosol consisting of sulfates (SO_4^{2-}), nitrates (NO_3^-), OC, and Ammonium (NH_4^+), polluted continental is background aerosol with a substantial fraction of urban pollution, marine is a hygroscopic aerosol that consists primarily of sea-salt (NaCl), and polluted dust is a mixture of desert dust and smoke or urban pollution. An aerosol will be classified as 'not determined' only in those cases where the classification scheme fails; as yet, no such cases have been identified in the version 2.01 release. The 'other' designation is a place-holder for another, yet to be determined, aerosol type. While this set does not cover all possible aerosol mixing scenarios, it accounts for a majority of mesoscale aerosol layers. In essence the algorithm trades off complex transient multi-component mixtures for relatively stable layers with large horizontal extent (10-1000 km).

CAD score [beta]

The cloud-aerosol discrimination (CAD) score provides the numerical result obtained for each layer by the CALIOP cloud-aerosol discrimination algorithm. The CAD algorithm separates clouds and aerosols based on multi-dimensional, altitude-dependent histograms of scattering properties (e.g., intensity and spectral dependence). In areas where there is no overlap or intersection between these histograms, features can be classified with complete confidence. Detailed descriptions of the CAD algorithm can be found in Sections 4 and 5 of the CALIPSO Scene Classification ATBD (PDF) and in Liu et al., 2004 (PDF).

The standard CAD scores reported in the CALIPSO layer products range between -100 and 100. The sign of the CAD score indicates the feature type: positive values signify clouds, whereas negative values signify aerosols. 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. Users are encouraged to refer to the CAD score when the cloud and aerosol classification results are used and interpreted. Beginning with the version 2.01 release, several additional CAD score values have been added. Each of these new values represents a classification result that is based on additional information beyond that normally considered in the standard CAD algorithm.

CAD score	Interpretation
-101	negative mean attenuated backscatter encountered; layer is most likely an artifact
101	initially classified as aerosol, but layer depolarization mandates classifying layer as cloud
102	layer exhibits very high integrated backscatter and very low depolarization characteristic of oriented ice crystals
103	layer integrated attenuated backscatter at 532 nm is suspiciously high; feature authenticity and classification are both highly uncertain

The accuracy of the CAD score depends on how accurately the PDFs approximate the cloud and aerosol distributions found in the real world, and on how completely clouds and aerosols are separated in the selected attribute space (i.e., by the attributes of layer averaged attenuated backscatter, color ratio, and layer altitude). 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. Using the standard algorithm alone produced an unacceptable level of misclassifications for very dense aerosols at low latitudes and optically thin low clouds in the polar regions. To improve the classifications of features in these regions, an additional depolarization ratio threshold criteria has been incorporated into the current algorithm.

Measured Two Way Transmittance 532 [beta; 5 km products only]

The measured value of the layer two-way transmittance, provided for isolated transparent layers. In this context, an isolated layer is one that is not in contact with another layer or the surface at either its upper or lower boundaries. The two-way transmittance is derived by computing the ratio of the mean attenuated scattering ratios in the "clear air" regions immediately below and above the layer. Details of the calculation are provided in the layer detection ATBD (PDF). This quantity is reported only for the 532 nm data, as the CALIOP 1064 nm channel is essentially insensitive to molecular backscatter. Physically meaningful measurements of two-way transmittance lie between 0 and 1; however, due to noise in the backscatter signal, and perhaps to undetected aerosol contamination of the "clear air" regions, the values reported in the CALIOP data products will sometimes exceed these bounds.

Measured Two Way Transmittance Uncertainty 532 [beta; 5 km products only]

The relative error in the two-way transmittance measurement, calculated using standard techniques for error propagation in ratioed quantities.

Two Way Transmittance Measurement Region [beta; 5 km products only]

Specifies the upper and lower boundaries for the "clear air" altitude region beneath each layer; the mean attenuated scattering ratio that determines the layer two-way transmittance is calculated using this region.

Feature Optical Depth 532 [beta; 5 km products only]
Feature Optical Depth 1064 [beta; 5 km aerosol products only]

Reports estimates of layer optical depth computed according to the procedures outlined in the CALIOP extinction retrieval ATBD (PDF). Estimates for aerosol optical depths are provided at both wavelengths. Because the extinction coefficients for clouds are largely independent of wavelength in the spectral region sampled by CALIOP, cloud optical depth is reported only for the 532 nm measurements. When using any of these values in scientific studies, users are cautioned to take note of several important caveats:

- For the vast majority of cases, CALIOP cannot provide a direct measurement of layer optical depth. In these cases, estimates of optical depth are derived using extinction-to-backscatter ratios (i.e., lidar ratios) that are specified based on an assessment of layer type and subtype. Uncertainties in the value of the lidar ratio, which can arise both from natural variability and from occasional misclassification, propagate non-linearly into subsequent estimates of layer optical depth.

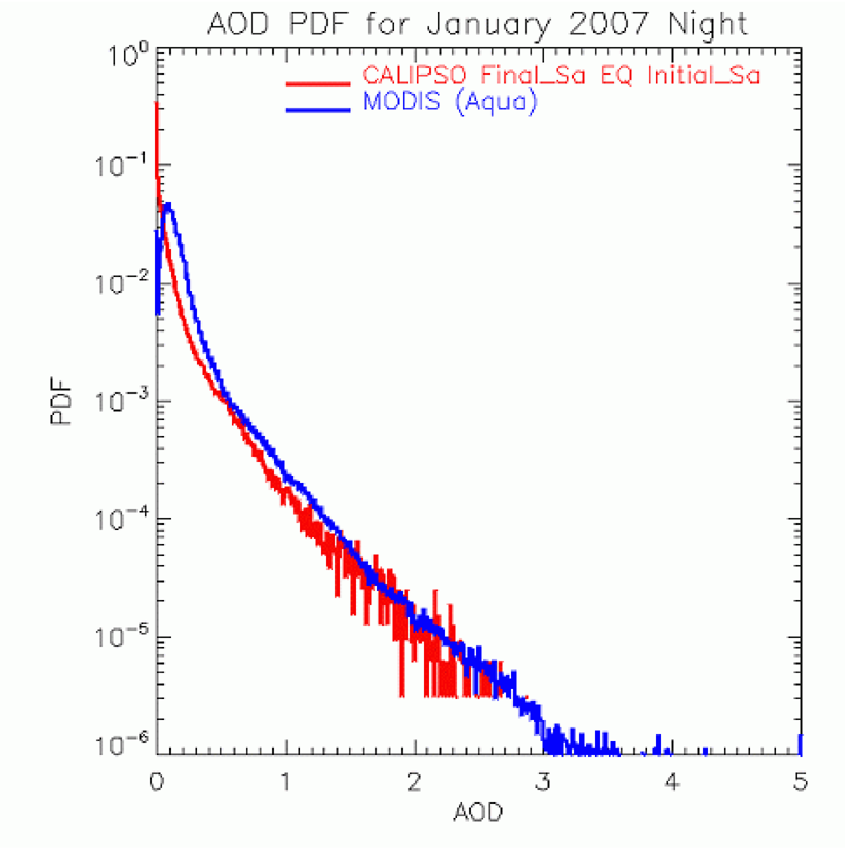
- Retrievals of optical depth from space-based lidar measurements must account for contributions from multiple scattering that are generally considered negligible in ground-based and aircraft based measurements. The theoretical basis for CALIPSO's treatment of multiple scattering is provided in the extinction retrieval ATBD (PDF) and in Winker, 2003 (PDF).

- Similar to the layer detection problem, estimates of layer optical depth become increasingly fraught with error in multiple layer scenes, as errors incurred in overlying layers are propagated into the solutions derived for underlying features.

- **IMPORTANT NOTICE**: before proceeding, all users of the CALIOP optical depth data should read and thoroughly understand the information provided in the Profile Products Data Quality Summary. This summary contains an expanded description of the extinction retrieval process from which the layer optical depths are derived, and provides essential guidance in the apppropriate use of all CALIOP extinction-related data products.

Despite these caveats, users should not be unduly pessimistic about the quality and usability of the CALIPSO optical depth estimates. Figure 2 (below) shows a preliminary comparison of CALIPSO aerosol optical depths with the optical depths derived from MODIS for nighttime measurements acquired during January 2007. The comparison is generally good, with MODIS appearing to slightly over-estimate values at the lower end of the optical depth range.

Figure 2: Comparison of CALIPSO aerosol optical depths to those derived from MODIS (Preliminary - January 2007, nighttime only) final lidar ratio = initial lidar ratio **only)**



Feature Optical Depth Uncertainty 532 [beta; 5 km products only]
Feature Optical Depth Uncertainty 1064 [beta; 5 km aerosol products only]

Not calculated for this data release; data products contain fill values in this field

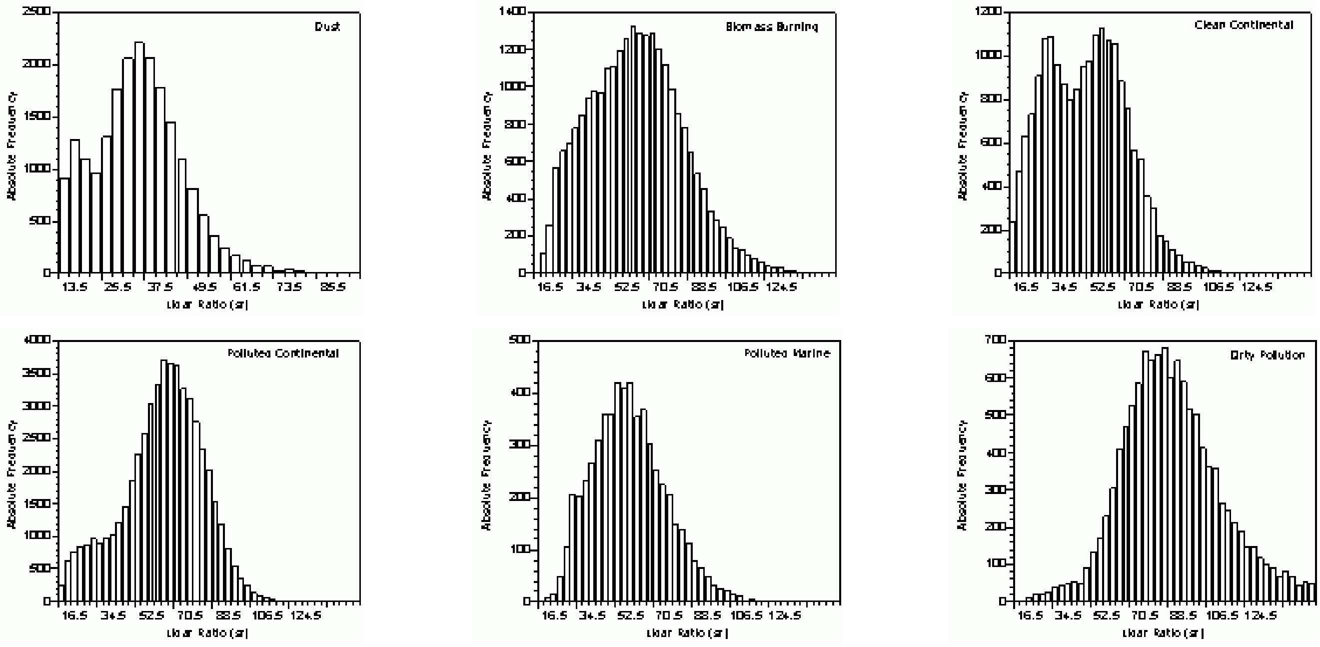
Initial 532 Lidar Ratio [beta; 5 km products only]
Initial 1064 Lidar Ratio [beta; 5 km aerosol products only]

Retrieving optical depth and profiles of extinction and backscatter coefficients from the CALIOP measurements requires an estimate of the particulate extinction-to-backscatter ratio, which in the lidar community is commonly known as the "lidar ratio". These initial estimates are selected based on the type and subtype of the layer being analyzed. The values used in the current release are as follows:

Type	Subtype	Initial 532 nm lidar ratio	Initial 1064 nm lidar ratio
cloud	water	18 sr	N/A
cloud	ice	25 sr	N/A
aerosol	marine	20 sr	45 sr
aerosol	desert dust	40 sr	30 sr
aerosol	polluted continental	70 sr	30 sr
aerosol	clean continental	35 sr	30 sr
aerosol	polluted dust	65 sr	30 sr
aerosol	biomass burning	70 sr	40 sr
stratospheric	all	15 sr	15 sr

The aerosol lidar ratios used for the CALIOP analyses represent well established mean values that are characteristic of the natural variability exhibited for each aerosol species (e.g., see Omar et al., 2005 (PDF); Cattrall et. al., 2005; and Figure 3 below). The clear implication of this natural variability is that even for those cases where the aerosol type is correctly identified, the initial lidar ratio represents an imperfect estimate of the layer-effective lidar ratio of any specific aerosol layer. These same caveats apply equally to the mean values used for the initial cloud lidar ratios. For all layer types, cloud-aerosol discrimination errors can exacerbate the error associated with the specification of the initial lidar ratio. Uncertainty can also be introduced by the cloud ice-water phase classification and the aerosol subtype identification procedures. However, the CALIOP extinction algorithm incorporates some error-correcting mechanisms that in many cases will adjust the initial estimate of lidar ratio so that a more suitable value is ultimately used in the retrieval. Details of the lidar ratio adjustment scheme are provided in the extinction retrieval ATBD (PDF). Algorithm architectural information and generalized error analyses for CALIPSO's cloud-aerosol discrimination algorithms, cloud ice-water phase algorithms, and aerosol subtyping algorithms can be found in the CALIPSO Scene Classification ATBD (PDF).

Figure 3: Distributions for AERONET-derived lidar ratios, computed for aerosol types described in Omar et al., 2005 (PDF);



Final 532 Lidar Ratio [beta; 5 km products only]
Final 1064 Lidar Ratio [beta; 5 km aerosol products only]

This parameter reports the lidar ratio in use at the conclusion of the extinction processing for each layer. The final lidar ratio may be the initial lidar ratio supplied by the Scene Classification Algorithms, it may be the result of modifications to this initial lidar ratio to avoid a non-physical solution, or it may be a lidar ratio determined from measured layer transmittance. In cases where a measured transmittance is available, the lidar ratio derived from that transmittance will be used as an initial condition. The extinction processing terminates when either a successful solution is obtained, or when the required adjustments to the lidar ratio exceed some predetermined bounds. Within those bounds, which range from 0 sr to 250 sr in the version 2.01 release, the extinction retrieval will, if necessary to avoid a non-physical solution, adjust the initial lidar ratio as required to produce a solution consistent with the measured data. Users can determine the status of the final lidar ratio by examining the extinction QC flags.

For weakly scattering features, the lidar ratio is most often left unchanged by the extinction solver, as a physical solution is usually obtained on the first iteration. In these cases, the uncertainties in the final lidar ratio are the same as the uncertainties in the initial lidar ratio. The exception to this statement would be if either the cloud-aerosol discrimination or the layer sub-typing procedures have misclassified the layer. However, for weak layers, the relative error in the lidar ratio is (approximately) linearly related to the resulting error in the derived optical depth estimate.

Retrievals of opaque and strongly scattering layers are very sensitive to the initial lidar ratio selection. Too large a value will cause the retrieval algorithm to, in effect, extinguish all available signal before reaching the measured base of the feature. When that point is reached the retrieval becomes numerically unstable and the calculated extinction coefficients will asymptote toward positive infinity. In these cases, a successful solution can only be obtained by reducing the lidar ratio. The CALIOP extinction routine does this automatically, and will repeat the process until a stable solution is achieved. When this happens, the final lidar ratio reported in the data products is the first one for which a physically meaningful (albeit not necessarily correct) solution was obtained for the entire measured depth of the layer.

The optical depths and extinction profiles derived in those cases where the layer lidar ratio must be reduced are generally not accurate. The current lidar ratio reduction scheme terminates after identifying (a very close estimate of) the largest lidar ratio for which a physically meaningful solution can be generated for the backscatter measured in the layer. However, the optical depths and extinction profiles reported in these situations can only be considered as upper bounds; the true values are somewhat, or perhaps even significantly, lower. Because the associated uncertainties cannot be reasonably estimated, these data should be excluded from statistical analyses of layer optical properties, and even the most sophisticated users are advised to treat theses cases with extreme caution.

When an independent estimate of layer optical depth is available from a measured layer two-way transmittance, the CALIOP extinction algorithm will retrieve the optimal estimate of the layer-effective lidar ratio, irrespective of layer type, and use this retrieved lidar ratio in the extinction retrieval. These so-called 'constrained' retrievals are more accurate than unconstrained retrievals. For constrained retrievals, the uncertainty in the final lidar ratio can be well estimated using equation 7.4 from the CALIPSO Scene Classification ATBD (PDF). An extinction QC value of 1 indicates a successful constrained retrieval.

Lidar Ratio 532 Selection Method [beta; 5 km products only]
Lidar Ratio 1064 Selection Method [beta; 5 km aerosol products only]

Specifies the internal procedure used to select the initial lidar ratio for each layer; valid values in this field are...

Value	Method
0	not determined
1	constrained retrieval (using two-way transmittance)
2	based on cloud phase
3	based on aerosol species
99	fill value

Layer Effective 532 Multiple Scattering Factor [beta; 5 km products only]
Layer Effective 1064 Multiple Scattering Factor [beta; 5 km aerosol products only]

The layer effective multiple scattering factors, η_{532} and η_{1064}, are specified at each wavelength according to layer type and subtype. Values range between 0 and 1; 1 corresponds to the limit of single scattering only, with smaller values indicating increasing contributions to the backscatter signal from multiple scattering. For the CALIOP viewing geometry, multiple scattering effects are nearly independent of range and, for parameterization in terms of extinction adjustments, are nearly independent of extinction. For the version 2.01 release, ice clouds are assigned a constant multiple scattering factor of $\eta_{532} = \eta_{1064} = 0.6$. All other layer types are assigned multiple scattering factors of 1. For weak to moderate aerosol layers, multiple scattering reduces the apparent extinction near the top of the aerosol layer but the multiple scattering effect decreases with penetration into the layer and usually has a negligible effect on the layer optical depth, so that a multiple scattering factor of unity is a good approximation. In dense aerosol layers, multiple scattering effects can become significant at the 10-20% level, but in this case errors due to uncertainties in the lidar ratio are usually more important. In water clouds, multiple scattering becomes a function of both extinction and range and is much more difficult to parameterize. The multiply-scattered component of the signal from water clouds is typically much larger than the single-scattered component of the signal so that the retrieval is very sensitive to errors in the multiple scattering factor used. In Version 2 data products, a multiple scattering factor of unity is assumed for water clouds. This introduces large errors, and extinction and optical depth retrievals in water clouds should be ignored. The rationale for the current CALIOP parameterization of multiple scattering, along with estimates of multiple scattering factors for ice clouds and several aerosol types, is described in detail in Winker, 2003 (PDF).

Extinction QC 532 [beta; 5 km products only]
Extinction QC 1064 [beta; 5 km aerosol products only]

The extinction QC flags are bit-mapped 16-bit integers, reported for each layer and for each wavelength for which an extinction retrieval was attempted. Aerosol extinction is computed for both wavelengths; cloud extinction is only reported at 532 nm. The information content of each bit is as follows:

Bit	Value	Interpretation
1	0	unconstrained retrieval; initial lidar ratio unchanged during solution process
1	1	constrained retrieval
2	2	Initial lidar ratio reduced to prevent divergence of extinction solution
3	4	Initial lidar ratio increased to reduce the number of negative extinction coefficients in the derived solution
4	8	Calculated backscatter coefficient exceeds the maximum allowable value
5	16	Layer being analyzed has been identified by the feature finder as being totally attenuating (i.e., opaque)
6	32	Estimated optical depth error exceeds the maximum allowable value
7	64	Solution converges, but with an unacceptably large number of negative values
8	128	Retrieval terminated at maximum iterations
9	256	No solution possible within allowable lidar ratio bounds
16	32768	Fill value or no solution attempted

The bit assignments are additive, so that (for example) an extinction QC value of 18 represents an unconstrained retrieval (bit 1 is NOT set) for which the lidar ratio was reduced to prevent divergence (+2; bit 2 is set), and for which the feature finder has indicated that the layer is opaque (+16; bit 5 is set). For the version 2.01 release, bits 10-15 are not used. Complete information about the conditions under which each extinction QC bit is toggled can be found in the CALIPSO Extinction Retrieval ATBD (PDF).

Midlayer Temperature [provisional]

Temperature, in degrees C, at the geometric midpoint of the layer in the vertical dimension; derived from the GEOS-5 data product provided to the CALIPSO project by the GMAO Data Assimilation System.

Relative Humidity [provisional; 5 km aerosol products only]

Relative humidity, in percent, at the geometric midpoint of the layer in the vertical dimension; derived from the GEOS-5 data product provided to the CALIPSO project by the GMAO Data Assimilation System.

Layer Properties: Planned for Future Releases

Future releases of the CALIPSO cloud and aerosol layer products will include several additional fields that will report an expanded range of layer optical properties. All currently planned fields are listed (but not described) below.

Integrated Particulate Color Ratio (5 km aerosol products only)

Not calculated for the current release; data products contain fill values in this field.

Integrated Particulate Color Ratio Uncertainty (5 km aerosol products only)

Not calculated for the current release; data products contain fill values in this field.

Particulate Color Ratio Statistics (5 km aerosol products only)

Not calculated for the current release; data products contain fill values in this field.

Cloud Fraction (5 km aerosol products only)

Not calculated for the current release; data products contain fill values in this field.

Integrated Particulate Depolarization Ratio (5 km products only)
Not calculated for the current release; data products contain fill values in this field.

Integrated Particulate Depolarization Ratio Uncertainty (5 km products only)
Not calculated for the current release; data products contain fill values in this field.

Particulate Depolarization Ratio Statistics (5 km products only)
Not calculated for the current release; data products contain fill values in this field.

Cirrus Shape Parameter (5 km cloud products only)
Not calculated for the current release; data products contain fill values in this field.

Cirrus Shape Parameter Uncertainty (5 km cloud products only)
Not calculated for the current release; data products contain fill values in this field.

Cirrus Shape Parameter Invalid Points (5 km cloud products only)
Not calculated for the current release; data products contain fill values in this field.

Ice Water Path (5 km cloud products only)
Not calculated for the current release; data products contain fill values in this field.

Ice Water Path Uncertainty (5 km cloud products only)
Not calculated for the current release; data products contain fill values in this field.

Cloud Fraction (5 km aerosol products only)
Not calculated for the current release; data products contain fill values in this field.

Fixed 532 Lidar Ratio (5 km aerosol products only)
Not calculated for the current release; data products contain fill values in this field.

Fixed 532 Lidar Ratio Optical Depth (5 km aerosol products only)
Not calculated for the current release; data products contain fill values in this field.

Fixed 532 Lidar Ratio Optical Depth Uncertainty (5 km aerosol products only)
Not calculated for the current release; data products contain fill values in this field.

Data Release Versions

<table>
<tr><th colspan="4">Lidar Level 2 Cloud and Aerosol Layer Information
<i>Half orbit (Night and Day) lidar cloud and aerosol layer products describe both column and layer properties</i></th></tr>
<tr><th>Release Date</th><th>Version</th><th>Data Date Range</th><th>Maturity Level</th></tr>
<tr><td>October 2008</td><td>2.02</td><td>09/14/2008 to 02/16/2009,
03/17/2009-03/28/2010</td><td>• Layer Heights - Provisional
• Aerosol/Cloud/ Stratospheric Classifications - Beta</td></tr>
<tr><td>January 25, 2008</td><td>2.01</td><td>June 13, 2006 to September 13, 2008</td><td>• Layer Heights - Provisional
• Aerosol/Cloud/ Stratospheric Classifications - Beta</td></tr>
</table>

Data Quality Statement for the release of the CALIPSO Lidar Level 2 Cloud and Aerosol Layer Products Version 2.02, October 2008

Version 2.02 of the Level 2 data products is a maintenance release that implements the following changes.

- Corrections were made to the code used to interpolate the GMAO meteorological data products to the CALIPSO orbit tracks.
- The Cabannes backscattering cross-sections used to derive the molecular scattering models used for the Level 1 and Level 2 analyses were revised downward by ~0.8%.
- A typographical error was identified in the runtime script that controls the behavior of the aerosol subtyping algorithm in the Level 2 analyses.

The impacts of these changes on the Level 2 data products are as follows:

- Layer detection: As a result of the first two changes, the 532 nm and 1064 nm calibration constants are larger, on average, by ~1%, resulting in corresponding decreases in the magnitudes of the attenuated backscatter coefficients at both wavelengths. These changes in the level 1 data result only small changes to the layer detection statistics. For example, the difference in the total number of layers detected by the two different versions on August 12, 2006 was 4: 9680 layers were detected by the version 2.01 code, versus 9676 layers by the version 2.02 code.

- Cloud-aerosol discrimination: with one exception, there were only minimal changes in cloud-aerosol discrimination results. The exception occurs in the polar regions when PSCs are present. For the August 12, 1006 test case, corrections to the interpolation algorithms applied to the GMAO data result in a slight upward shift in the tropopause heights, and as a consequence, more clouds and fewer stratospheric layers are identified in the version 2.02 results.
- Ice-water phase determination: because this classification is based on depolarization ratio and temperature no substantial changes, there were no substantial changes in the assessments of cloud thermodynamic state.
- Aerosol subtype identification: Correcting the level 2 runtime script error will reduce the number of layers identified as smoke, and increase the number of layers identified as sea salt.

Cloud and aerosol extinction profiles and optical properties: changes in backscatter and extinction coefficients at the tops of layers are small, and proportional to the changes in the calibration coefficients ... however, due to the cumulative nature of error propagation in the extinction retrieval, differences increase with increasing penetration depths, and can grow large when the optical depths of the clouds are large (i.e., > 3).

Data Quality Statement for the release of the CALIPSO Lidar Level 2 Cloud and Aerosol Layer Products Version 2.01, January 25, 2008

The primary new parameters included in this release are aerosol and cloud extinction and backscatter profiles (Aerosol Profile Product and Cloud Profile Product), layer optical depth, aerosol type, and cloud ice/water phase (Aerosol and Cloud Layer Product), and aerosol type and cloud ice/water phase in the Vertical Feature Mask (VFM). Although extinction and/or optical depth appear in several different products, all extinction retrievals are produced by the same algorithm. Therefore, the Data Quality Summaries include a section which discusses general characteristics of the extinction and optical depth data applicable to all products.

Layer Detection

Given the accuracy of the CALIPSO altitude registration, the layer heights reported in the Lidar Level 2 Cloud and Aerosol Layer Products appear to be quite accurate. In optically dense layers, the lowest altitude where signal is reliably observed is reported as the base. In actuality, this reported base may lie well above the true base. In this release, the layers which are reported represent a choice in favor of high reliability over maximum sensitivity. Weakly scattering layers sometimes will go unreported, in the interest of minimizing the number of false positives.

Cloud-Aerosol Discrimination

Based on the initial CALIOP measurements, an improved version of the cloud-aerosol discrimination (CAD) algorithm has been implemented for this release. Overall, the updated 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. Nevertheless, several types of misclassifications still occur with some frequency. Among these, the most prevalent are:

1. Dense aerosol layers (primarily very dense dust and smoke over and close to the source regions) are sometimes labeled as cloud. 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 can 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 probability distribution functions. 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.

2. 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 at lower latitudes and misclassifications of clouds as aerosol are more common. In particular, thin ice clouds which can extend from the surface to several kilometers in altitude, are sometimes misclassified as aerosol.

3. 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. 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.

4. 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).

Aerosol Type Identification

The main objective of the aerosol subtyping scheme is to estimate the appropriate value of the aerosol extinction-to-backscatter ratio (S_a) to within 30% of the true value. S_a is an important parameter used in the determination of the aerosol extinction and subsequently the optical depth from CALIOP backscatter measurements. S_a is an intensive aerosol property, i.e., a property that does not depend on the number density of the aerosol but rather on such physical and chemical properties as size distribution, shape and composition. These properties depend primarily on the source of the aerosol and such factors as mixing, transport, and in the case of hygroscopic aerosols, hydration.

The extinction products are produced by first identifying an aerosol type and then using the appropriate values of S_a and the multiple scattering factor, $\eta(z)$. Note that multiple scattering corrections have not yet been implemented for the current data release, so that $\eta(z) = 1$ for all aerosol types. The accuracy of the S_a value used in the lidar inversions depends on the correct identification of the type of aerosol. In turn, the accuracy of the subsequent optical depth estimate depends on the accuracy of S_a.

The underlying paradigm of the type classification is that a variety of emission sources and atmospheric processes will act to produce air masses with a typical, identifiable aerosol 'type'. This is an idealization, but one that allows us to classify aerosols based on observations and location in a way to gain insight into the geographic distribution of aerosol types and constrain the possible values of S_a for use in aerosol extinction retrievals.

The aerosol subtype product is generated downstream of the cloud-aerosol discrimination (CAD) scheme and, therefore, depends on the cloud-aerosol classification scheme in a very fundamental way. If a cloud feature is misclassified as aerosol, the aerosol subtype algorithm will identify this 'aerosol' as one of the aerosol subtypes. The user must exercise caution where the aerosol subtype looks suspicious or unreasonable. Such situations can occur with some frequency in the southern oceans and the polar regions.

Cloud Ice/Water Phase Discrimination

Cloud phase is determined using a depolarization/backscatter relation, together with temperature and backscatter thresholds. Complete descriptions of the algorithm mechanics and underlying theory are given in Section 6 of the CALIPSO Scene Classification ATBD (PDF). The algorithm implemented for the version 2.01 release identifies obvious water and ice clouds and clear cases of oriented ice crystals. Improvements for recognizing mixed phase clouds are planned for future release.

Cloud and Aerosol Optical Depths

Because comprehensive data assessment and screening procedures have not yet been developed, the CALIOP cloud and aerosol optical depths reported in the version 2.0 data release are not considered appropriate for scientific publication. The data is being released as a beta quality product, for evaluation by the user community, and to provide feedback to the CALIOP algorithm development team.

PLEASE NOTE: users of the CALIOP extinction and backscatter profile data should read and thoroughly understand the information provide in the Profile Products Data Quality Summary. This summary contains an expanded description of the extinction retrieval process from which the layer optical depths are derived, and provides essential guidance in the apppropriate use of all CALIOP extinction-related data products. Validation and improvements to the profile products QA are ongoing efforts, and additional data quality information will be included with future releases.