National Aeronautics and Space Administration

Langley Research Center

Hampton, Virginia 23681-2199

Cloud – Aerosol LIDAR and Infrared Pathfinder Satellite Observations (CALIPSO)

Data Description and Quality Summary

Lidar Level 2 Layer Data

DOI: https://doi.org/10.5067/CALIOP/CALIPSO/CAL_LID_L2_Layer_Data-Description_Document_V5

Cloud – Aerosol LIDAR and Infrared Pathfinder Satellite Observations

Data Description and Quality Summary

Lidar Level 2 Layer Data

Primary Authors

Mark Vaughan, Michael Pitts, Charles Trepte, David Winker, Brian Getzewich, Jason Tackett, Xia Cai
NASA Langley Research Center
Hampton, Virginia 23681-2199

Jacques Pelon
LATMOS/IPSL, Université Pierre et Marie Curie
75252 Paris Cedex 5 (France)

Anne Garnier, Jayanta Kar, Kam-Pui Lee
Analytical Mechanics Associates (AMA)
21 Enterprise Parkway, Hampton, Virginia 23666

Timothy Murray, Sharon Rodier
ADNET Systems, Inc.
6720B Rockledge Drive, Suite #504, Bethesda, Maryland 20817

Robert Ryan
Coherent Applications, Inc. – Psionic LLC
1100 Exploration Way, Hampton, Virginia 23666

January 30, 2026

CALIPSO Lidar Level 2 Layer Data Description Document
Version 5.00

Data Version: 5.00
Data Release Date: October 01, 2025
Data Date Range: June 13, 2006 to June 30, 2023

Introduction

By combining two disparate but tightly coupled data realizations, the CALIPSO cloud, aerosol, and merged layer products deliver comprehensive insights into the large-scale vertical structure of Earth's atmosphere. The first of these is a set of *column properties*, which describe the temporal and spatial location of the vertical column (or, for averaged data, the vertical 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 detected in the column and include quantities such as layer base and top altitudes, integrated attenuated backscatter, layer-integrated volume depolarization ratio, and optical depth.

The CALIPSO project creates "standard" layer products for different subsets of feature types and at several different horizontal averaging resolutions. The cloud layer products are distributed at 1 km (3 shot) and 5 km (15 shot) resolutions and report layer properties for clouds only. Aerosol layer products are distributed at 5 km, with layer properties limited solely to aerosols detected in either the stratosphere or the troposphere. The merged layer products, distributed at 5 km and single shot resolutions, report layer properties for all features detected, irrespective of feature type. File naming conventions shown in Table 1.

Table 1: the left gives the resolution and feature type for each of the CALIOP V5.00 layer products; the right column specifies the file naming convention used to indicate the product type in the left column

Product	File Naming Convention
333 m standard merged layer	CAL_LID_L2_333mMLay-Standard-V5-00*.hdf
1 km standard cloud layer	CAL_LID_L2_01kmCLay-Standard-V5-00*.hdf
5 km standard cloud layer	CAL_LID_L2_05kmCLay-Standard-V5-00*.hdf
5 km standard aerosol layer	CAL_LID_L2_05kmALay-Standard-V5-00*.hdf
5 km standard merged layer	CAL_LID_L2_05kmMLay-Standard-V5-00*.hdf
5 km beta merged layer diagnostic	CAL_LID_L2_MLay_Diagnostic-Beta-V5-00*.hdf

The 5 km layer products report features detected at all altitudes between ~30 km and –0.5 km AMSL. Due to CALIPSO's altitude-dependent on-board data averaging scheme (Hunt et al., 2009), data downlinked at 1 km vertical resolution are only available at altitudes between ~20.2 km and –0.5 km AMSL, thus placing limits on the altitude range of the layers reported in the 1 km cloud layer products. The vertical range of layers reported in the single shot merged layer products is similarly restricted by the downlinked raw data resolution to altitudes between ~8.2 km and –0.5 km AMSL.

As described in Vaughan et al., 2009, the so-called "5 km layer products" actually report properties of features that can be detected at 5 km, 20 km, and 80 km along-track averaging resolutions. Thus, while the column properties are unique for each 5 km segment, the layer properties may not be. Layer properties of features detected at 20 km and 80 km are replicated as necessary along the lidar orbit track to span the full horizontal extent of each layer. Also, the 'Single Shot Detection' Vgroup reported in all 5 km layer products contains a substantial subset of the column and layer properties reported in the single shot merged layer product. This additional information is

included to provide data users the means to readily reconstruct heterogeneous layer composition below 4 km, where CALIOP's single shot boundary layer cloud clearing algorithm operates (Vaughan et al., 2009; Vaughan et al., 2010; Tackett et al., 2023). The scientific data set (SDS) names in the 'Single Shot Detection' Vgroup are close to identical to those used in the 333 m merged layer product. The difference is that the SDS names in these Vgroups all include a two-letter prefix ('ss') indicating that these parameters are retrieved/reported at single shot resolution.

CALIPSO's version 5.00 data release also includes one entirely new layer product: a beta version of the 5 km merged layer product augmented with diagnostic information not included elsewhere (i.e., CAL_LID_L2_MLay_Diagnostic-Beta-V5-00*.hdf). Among the new and enhanced parameters included in these diagnostic files are ocean-derived column optical depth (ODCOD) retrievals at 1064 nm, along with a wealth of ODCOD intermediate calculations, and a full set of cloud optical property retrievals at 1064 nm, including cloud optical depths. In addition to all SDSs and Vgroups contained in the standard 5 km merged layer product, the beta 5 km merged diagnostic file also contains a '01km_Detection' Vgroup which includes a comprehensive set of the layer and column properties reported in the 1 km cloud layer product. As in the 'Single Shot Detection' Vgroup, the SDS names in the '01km_Detection' Vgroup all include a prefix ('01km') indicating that these parameters are retrieved/reported at 1 km horizontal resolution.

Table of Contents

Additional Documentation

Project Documentation

- CALIPSO Data Management Team: CALIPSO Data Products Catalog, PC-SCI-503, Release 5.00.
- Hostetler, C. A., Z. Liu, J. Reagan, M. Vaughan, D. Winker, M. Osborn, W. H. Hunt, K. A. Powell, and C. Trepte, 2006: CALIOP Algorithm Theoretical Basis Document: Calibration and Level 1 Data Products, PC-SCI-201 Release 1.0, https://ntrs.nasa.gov/citations/20250006623.
- Winker, D. M., C. A. Hostetler, M. A. Vaughan, and A. H. Omar, 2006: CALIOP Algorithm Theoretical Basis Document: Part 1 : CALIOP Instrument, and Algorithms Overview, PC-SCI-202 Part 1 Release 2.0, https://ntrs.nasa.gov/citations/20250006626.
- Vaughan, M. A., D. M. Winker, and K. A. Powell, 2005: CALIOP Algorithm Theoretical Basis Document Part 2: Feature Detection and Layer Properties Algorithms, PC-SCI-202 Part 2 Release 1.01, https://ntrs.nasa.gov/citations/20250006627.
- Liu, Z., A. H. Omar, Y. Hu, M. A. Vaughan, and D. M. Winker, 2005: CALIOP Algorithm Theoretical Basis Document: Part 3: Scene Classification Algorithms, PC-SCI-202 Part 3 Release 1.0, https://ntrs.nasa.gov/citations/20250006628.

Peer-Reviewed Algorithm Papers

- Avery, M. A., R. A. Ryan, B. J. Getzewich, M. A. Vaughan, D. M. Winker, Y. Hu, A. Garnier, J. Pelon, and C. A. Verhappen, 2020: CALIOP V4 Cloud Thermodynamic Phase Assignment and the Impact of Near-Nadir Viewing Angles, *Atmos. Meas. Tech.*, **13**, 4539–4563, https://doi.org/10.5194/amt-13-4539-2020.
- Hunt, W. H, D. M. Winker, M. A. Vaughan, K. A. Powell, P. L. Lucker, and C. Weimer, 2009: CALIPSO Lidar Description and Performance Assessment, *J. Atmos. Oceanic Technol.*, **26**, 1214–1228, https://doi.org/10.1175/2009JTECHA1223.1.
- Kim, M.-H., A. H. Omar, J. L. Tackett, M. A. Vaughan, D. M. Winker, C. R. Trepte, Y. Hu, Z. Liu, L. R. Poole, M. C. Pitts, J. Kar, and B. E. Magill, 2018: The CALIPSO Version 4 Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *Atmos. Meas. Tech.*, **11**, 6107-6135, https://doi.org/10.5194/amt-11-6107-2018.
- Liu, Z., J. Kar, S. Zeng, J. Tackett, M. Vaughan, M. Avery, J. Pelon, B. Getzewich, K.-P. Lee, B. Magill, A. Omar, P. Lucker, C. Trepte, and D. Winker, 2019: Discriminating Between Clouds and Aerosols in the CALIOP Version 4.1 Data Products, *Atmos. Meas. Tech.*, **12**, 703–734, https://doi.org/10.5194/amt-12-703-2019.
- Ryan, R. A., M. A. Vaughan, S. D. Rodier, J. L. Tackett, J. A. Reagan, R. A. Ferrare, J. W. Hair, and B. J. Getzewich, 2024: Total Column Optical Depths Retrieved from CALIPSO Lidar Ocean Surface Backscatter, *Atmos. Meas. Tech.*, **17**, 6517–6545, https://doi.org/10.5194/amt-17-6517-2024.
- Tackett, J. L., J. Kar, M. A. Vaughan, B. Getzewich, M.-H. Kim, J.-P. Vernier, A. H. Omar, B. Magill, M. C. Pitts, and D. Winker, 2023: The CALIPSO version 4.5 stratospheric aerosol subtyping algorithm, *Atmos. Meas. Tech.*, **16**, 745–768, https://doi.org/10.5194/amt-16-745-2023.
- Vaughan, M., K. Powell, R. Kuehn, S. Young, D. Winker, C. Hostetler, W. Hunt, Z. Liu, M. McGill, and B. Getzewich, 2009: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, *J. Atmos. Oceanic Technol.*, **26**, 2034–2050, https://doi.org/10.1175/2009JTECHA1228.1.
- Winker, D. M., M. A. Vaughan, A. H. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young, 2009: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, *J. Atmos. Oceanic Technol.*, **26**, 2310–2323, https://doi.org/10.1175/2009JTECHA1281.1.
- Young, S. A. and M. A. Vaughan, 2009: The retrieval of profiles of particulate extinction from Cloud Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) data: Algorithm description, *J. Atmos. Oceanic Technol.*, **26**, 1105–1119, https://doi.org/10.1175/2008JTECHA1221.1.
- Young, S. A., M. A. Vaughan, R. E. Kuehn, and D. M. Winker, 2013: The Retrieval of Profiles of Particulate Extinction from Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) Data: Uncertainty and Error Sensitivity Analyses, *J. Atmos. Oceanic Technol.*, **30**, 395–428, https://doi.org/10.1175/JTECH-D-12-00046.1.
- Young, S. A., M. A. Vaughan, J. L. Tackett, A. Garnier, J. B. Lambeth, and K. A. Powell, 2018: Extinction and Optical Depth Retrievals for CALIPSO's Version 4 Data Release, *Atmos. Meas. Tech.*, **11**, 5701–5727, https://doi.org/10.5194/amt-11-5701-2018.

Glossary and Acronym Dictionary

Term	Meaning
1kmCLay	CALIOP 1 km cloud layer product
333mMLay	CALIOP 333 m merged layer product (reports clouds and aerosols)
5kmALay	CALIOP 5 km aerosol layer product
5kmCLay	CALIOP 5 km cloud layer product
5kmMLay	CALIOP 5 km merged layer product (reports clouds and aerosols)

Term	Meaning
5kmDiag	CALIOP 5 km merged layer diagnostic product (reports clouds and aerosols)
AERONET	AErosol RObotic NETwork
AFWA	U.S. Air Force Weather Agency
ATBD	Algorithm Theoretical Basis Document
AMSL	above mean sea level
AOD	aerosol optical depth
ATBD	algorithm theoretical basis document
BLaCC	boundary layer cloud clearing
CAD	cloud-aerosol discrimination
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CERES	Clouds and the Earth's Radiant Energy System
CRM	CALIOP response model
DCRM	downlinked CALIOP response model
DEM	digital elevation model
EARLINET	European Aerosol Lidar Network
ECR	Earth Centered Rotating
frame	a fundamental data averaging interval used extensively in CALIOP's level 1 and level 2 data processing. A frame consists of 15 consecutive single-shot profiles spanning an along-track distance of ~5 km. 15 shots is the least common multiple of the along-track averaging intervals defined by CALIOP's vertically varying onboard data averaging scheme (Hunt et al., 2009).
GMAO	Global Modeling and Assimilation Office
GOCART	Goddard Chemistry Aerosol Radiation and Transport
granule	continuous data segment in which all measurements were acquired while the lidar was configured for daytime data acquisition only or nighttime data acquisition only; each granule spans approximately one half of a full orbit, with daytime granules being slightly larger/longer than nighttime granules
HERA	Hybrid Extinction Retrieval Algorithm
HDF	Hierarchical Data Format
HOI	horizontal oriented ice
HSRL	high spectral resolution lidar
IGBP	International Geosphere–Biosphere Programme
IAB	integrated attenuated backscatter
IIR	imaging infrared radiometer
IRF	instrument response function
IVDR	integrated volume depolarization ratio
L1B	level 1B
LEM	low energy mitigation
LSWG	CALIPSO lidar science working group
MAC	Maps of Aerosol lidar ratios for CALIPSO
MERRA-2	Modern-Era Retrospective analysis for Research and Applications, Version 2

Term	Meaning
MODIS	Moderate Resolution Imaging Spectroradiometer
MSL	mean sea level
N/A	not applicable
NAMMA	NASA African Monsoon Multidisciplinary Analyses
nm	nanometer
ODCOD	ocean-derived column optical depth
PBL	planetary boundary layer
PDF	probability distribution function
PSC	polar stratospheric cloud
QA	quality assurance/quality control
QC	quality control/quality assurance
RMS	root mean square
SNR	signal-to-noise ratio
SDS	scientific data set
sr	steradian
TAI	International Atomic Time
UTC	Coordinated Universal Time
WRS	Worldwide Reference System

Scientific Data Sets: Measurement Altitudes

Lidar_Data_Altitudes (all layer products)

Units: km

Format: Float_32

Valid Range: -2.0, 40.0

Description: Altitudes (above mean sea level) that specify the vertical midpoints of the 583 range bins in the profile measurements downlinked from the CALIPSO satellite. This scientific data set (SDS) is only available in the V5.00 data release. In all prior versions, the lidar data altitude array was stored in the file metadata.

Scientific Data Sets: Time and Position

Profile_Time (all layer products)

Units: TAI seconds

Format: Float_64

Valid Range: 4.203E8, 9.623E8

Description: International Atomic Time (TAI) in elapsed seconds from January 1, 1993. Times reported in the 333 m 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 data. For all 5 km layer products, three values are reported: the time for the first pulse included in the 15-shot average; the time at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots); and the time for the final pulse.

Profile_UTC_Time (all layer products)

Units: yymmdd.ffffffff

Format: Float_64

Valid Range: 60428.0, 230701.0

Description: Coordinated Universal Time (UTC), formatted as yymmdd.ffffffff, where yy is a two-digit data acquisition year number (06 to 23), mm is a month number (01 to 12), dd is a day number (01 to 31), and ffffffff is the elapsed fraction of the data acquisition day. Times reported in the 333 m 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 all 5 km layer products, three values are reported: the time for the first pulse included in the 15-shot average; the time at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots); and the time for the final pulse.

Day_Night_Flag (all layer products)

Units: NoUnits

Format: Int_8

Valid Range: 0, 1

Description: As CALIPSO approaches the terminator, the lidar is automatically reconfigured to adapt to changing lighting conditions that directly impact signal-to-noise (SNR) levels. These changes occur at Sun-Earth-Satellite (SES) angles of 95° (day to night) and 265° (night to day), corresponding to changes in lighting conditions at an altitude of ~24 km above mean sea level. 0 indicates daytime measurement configuration, 1 indicates nighttime.

Profile_ID (all layer products)

Units: NoUnits

Format: Int_32

Valid Range: 1, 228630

Description: Unique profile identifier generated sequentially for each lidar level 1b profile. Level 1b profile IDs are guaranteed to be unique within each granule but will not be unique over multiple granules. The 333 m (single shot) layer products report the profile ID assigned to each pulse. The 1 km products report the profile ID of the first of the three consecutive laser pulses averaged to create data at 1 km horizontal resolution. The 5 km layer products report the profile IDs for the first and last of the 15 consecutive lasers pulses averaged to create each 5 km profile product data record.

Latitude (all layer products)

Units: degrees north

Format: Float_32

Valid Range: -90.0, 90.0

Description: Geodetic latitude of the laser footprint on the Earth's surface. Latitudes reported in the 333 m 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 at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots); and the footprint latitude for the final pulse.

Longitude (all layer products)

Units: degrees east

Format: Float_32

Valid Range: -180.0, 180.0

Description: Longitude of the laser footprint on the Earth's surface. Longitudes reported in the 333 m 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 at the temporal midpoint (i.e., at the 8th of 15 consecutive laser shots); and the footprint longitude for the final pulse.

Scientific Data Sets: Column Optical Properties

Column_Integrated_Attenuated_Backscatter_532 (all layer products)

Units: 1/sr

Format: Float_32

Valid range: 0.0, 2.0

Dimensions: number of profiles × 1

Fill value: -9999

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The integral with respect to altitude of the 5 km horizontal average (15 shots) of the profiles of 532 nm total attenuated backscatter coefficients reported in the CALIOP level 1b product. 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 model (DEM). The column integrated attenuated backscatter provides a measure of atmospheric turbidity above the surface.

Column_IAB_Cumulative_Probability (all layer products)

Units: NoUnits

Format: Float_32

Valid range: 0.0, 1.0

Dimensions: number of profiles × 1

Fill value: -9999

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The cumulative probability of measuring a 532 nm total column integrated attenuated backscatter (γ'_{532}) value equal to the value computed for the current profile. For any profile, the limits of integration for γ'_{532} are from the first range bin in the profile, at ~40 km, down to one range bin above the surface altitude specific by the CALIPSO DEM. Values in this field range between 0 and 1. Figure 1 shows the cumulative probability distribution compiled using all CALIOP single shot γ'_{532} measurements acquired between 2008-07-01 thru 2008-07-15.

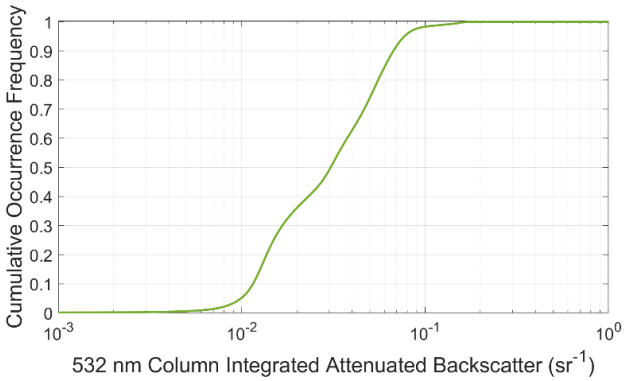


Figure 1: Cumulative distribution of 532 nm single shot total column integrated attenuated backscatter

Column_Optical_Depth_Cloud_532 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid range: 0.0, 25.0

Dimensions: number of profiles × 1

Fill value: -9999

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM
–333.0 failed extinction retrieval within a column
–444.0 suspicious/incomplete single-shot cloud clearing within a column

Description: Optical depth of all clouds detected within each 5 km vertical column, obtained by integrating the 532 nm cloud extinction profile reported in the CALIPSO 5 km Cloud Profile Products from the uppermost range bin to the lowest range bin in which clouds are identified.

Column_Optical_Depth_Cloud_Uncertainty_532 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Dimensions: number of profiles × 1

Fill value: -9999

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM
–333.0 failed extinction retrieval within a column
–444.0 suspicious/incomplete single-shot cloud clearing within a column

Description: Estimated uncertainty in the Column Optical Depth Cloud 532 parameter calculated using

$$\Delta\tau_p = \sqrt{\left(\gamma_p \cdot \Delta\overline{S_p}\right)^2 + \left(\Delta\gamma_p \cdot \overline{S_p}\right)^2}\,. \tag{1}$$

$\Delta\tau_p$ is the uncertainty in the column optical depth and γ_p and $\Delta\gamma_p$ are, respectively, the column integrated particulate backscatter and its associated uncertainty. These quantities are computed using trapezoid rule integration.

$$\gamma_p = \frac{1}{2}\sum_{n=TOA+1}^{surface}\left(z_{n-1} - z_n\right)\left(\beta_p\left(z_{n-1}\right) + \beta_p\left(z_n\right)\right) \tag{2}$$

$$\Delta\gamma_p = \sqrt{\frac{1}{4}\sum_{n=TOA+1}^{surface}\left(z_{n-1} - z_n\right)^2 \times \left(\Delta\beta_p\left(z_n\right)^2 + \Delta\beta_p\left(z_{n-1}\right)^2\right)} \tag{3}$$

In those range bins for which $\beta(z_n)$ is a fill value (i.e., no feature detected), $\beta(z_n)$ is set to 0 prior to computing γ_p and $\Delta\gamma_p$.

$\overline{S_p}$ and $\Delta\overline{S_p}$ are, respectively, the backscatter-weighted column mean lidar ratio and its uncertainty, calculated for non-zero values of $\beta(z_n)$ as follows.

$$\overline{S_p} = \frac{\frac{1}{N}\sum_{n=TOA}^{surface} \beta(z_n) \times S_p(z_n)}{\frac{1}{N}\sum_{n=TOA}^{surface} \beta(z_n)} \quad (4)$$

The quantity N is initialized to 0, then incremented by 1 whenever $\beta(z_n) \neq 0$.

$$\Delta\overline{S_p} = \frac{\frac{1}{N}\sum_{n=TOA}^{surface} \beta(z_n) \times \Delta S_p(z_n)}{\frac{1}{N}\sum_{n=TOA}^{surface} \beta(z_n)} \quad (5)$$

Column_Optical_Depth_Tropospheric_Aerosols_532 (all 5 km layer products)

Column_Optical_Depth_Tropospheric_Aerosols_1064 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid range: 0.0, 3.0

Dimensions: number of profiles × 1

Fill value: -9999

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM
–333.0 failed extinction retrieval within a column
–444.0 suspicious/incomplete single-shot cloud clearing within a column

Description: Optical depth of all tropospheric aerosols detected within each 5 km vertical column, obtained by integrating, respectively, the 532 nm and 1064 nm tropospheric aerosol extinction coefficients reported in the CALIPSO 5 km Aerosol Profile Products from the uppermost range bin to the lowest range bin in which tropospheric aerosols are identified. At both wavelengths, total column aerosol optical depth (AOD) is the sum of the tropospheric and stratospheric column AODs.

Column_Optical_Depth_Tropospheric_Aerosols_Uncertainty_532 (all 5 km layer products)

Column_Optical_Depth_Tropospheric_Aerosols_Uncertainty_1064 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Dimensions: number of profiles × 1

Fill value: -9999

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM
–333.0 failed extinction retrieval within a column
–444.0 suspicious/incomplete single-shot cloud clearing within a column

Description: Estimated uncertainty in, respectively the Column Optical Depth Tropospheric Aerosol 532 and Column Optical Depth Tropospheric Aerosol 1064 parameters, calculated according to equations 1 through 5 above.

Column_Optical_Depth_Stratospheric_Aerosols_532 (all 5 km layer products)

Column_Optical_Depth_Stratospheric_Aerosols_1064 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 25.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM
–333.0 failed extinction retrieval within a column
–444.0 suspicious/incomplete single-shot cloud clearing within a column

Description: Optical depth of all stratospheric aerosols detected within each 5 km vertical column, obtained by integrating, respectively, the 532 nm and 1064 nm stratospheric aerosol extinction coefficients reported in the CALIPSO 5 km Aerosol Profile Products from the uppermost range bin to the lowest range bin in which stratospheric aerosols are identified. At both wavelengths, total column AOD is the sum of the tropospheric and stratospheric column AODs.

Column_Optical_Depth_Stratospheric_Aerosols_Uncertainty_532 (all 5 km layer products)

Column_Optical_Depth_Stratospheric_Aerosols_Uncertainty_1064 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Dimensions: number of profiles × 1

Fill value: -9999

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM
–333.0 failed extinction retrieval within a column
–444.0 suspicious/incomplete single-shot cloud clearing within a column

Description: Estimated uncertainty in, respectively the Column Optical Depth Stratospheric Aerosol 532 and Column Optical Depth Stratospheric Aerosol 1064 parameters, calculated according to equations 1 through 5 above.

Column_Particulate_Optical_Depth_Above_Opaque_Water_Cloud_532 (all layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 3.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM
–33.33 failed retrieval

Description: 532 nm particulate optical depths in the vertical column above an opaque water cloud, derived according to the method detailed in Hu et al., 2007.

Column_Particulate_Optical_Depth_Above_Opaque_Water_Cloud_Uncertainty_532 (all layer products)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM
–33.33 failed retrieval

Description: Estimated uncertainty in the column 532 nm particulate optical depths above opaque water clouds. These uncertainty estimates combine both systematic biases arising from the assumption of a water cloud

lidar ratio and random uncertainties arising from the SNR of the 532 nm parallel and perpendicular channel measurements within the water clouds.

Column_Feature_Fraction (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The fraction of the 5-km horizontally averaged profile, between 30 km and the DEM surface elevation, which has been identified as containing a feature (i.e., either a cloud, an aerosol, or a stratospheric layer) at any of the 5 km, 20 km, or 80 km horizontal averaging resolutions.

Some Notes on Evaluating CALIOP Column Optical Properties

Opacity: The CALIPSO lidar is only capable of penetrating to the surface in columns with total particulate optical depths less than ~5. (Note that this value considers the contribution of multiple scattering.) In opaque columns (i.e., when the column backscatter signal is totally attenuated before the Earth's surface can be detected), the reported "column" optical depths do not characterize the full vertical column, but instead report the integral of the particulate extinction coefficients from ~30 km down to the apparent base of the lowest feature observed. In the layer products, the opacity of any column can be determined by inspecting the opacity flag of the lowest detected layer. Values of 1 indicate that the lowest detected layer in the column is classified as opaque.

Extinction QC: The extinction QC values in the column should be examined to determine if any of the cloud or aerosol extinction retrievals either failed or converged under suspicious circumstances. Low quality extinction retrievals in overlying layers can degrade the reliability of all extinction retrievals in the layers below (Young & Vaughan, 2009). For transparent layers, constrained retrievals (extinction QC = 1) yield the highest quality solutions. Solutions for which the final lidar ratio is unchanged from the initial estimate (extinction QC = 0) also generate physically plausible solutions. On the other hand, retrievals for which the initial lidar ratio must be reduced to achieve a solution can be problematic, as the retrieved optical depths in these cases tend to be biased high.

CAD Scores and Feature Subtypes: Irrespective of the extinction QC flag, extinction solutions for features with low magnitude cloud-aerosol discrimination (CAD) scores (e.g., $|CAD| \leq 20$) or "special" CAD scores should be regarded with caution. These features could not be confidently classified as either cloud or aerosol, and this lack of confidence extends to the choice of initial lidar ratio used in the extinction solution. Similarly, uncertain aerosol type classifications lead to increased uncertainties in the lidar ratios and hence lower confidence in the extinction retrievals.

Cloud phase: The extinction solutions in columns containing clouds predominately composed of horizontal oriented ice (HOI) crystals are likely to be highly unreliable. The anomalously high backscatter generated by HOI clouds, coupled with very low skill in selecting an appropriate lidar ratio for these clouds, frequently yields very large overestimates of cloud extinction coefficients and optical depths (e.g., Mioche et al., 2010).

Scientific Data Sets: Column QA Information

Calibration_Altitude_532 (all 5 km layer products)

Units: km

Format: Float_32

Valid Range: 0.0, 40.0

Fill Value: -9999.0

Description: Top and base altitudes, above mean sea level, of the region of the atmosphere used for calibrating the 532 nm parallel channel. Note that while the nighttime values in this SDS are both constant (top = 39 km, base = 36 km) and correct, daytime values are erroneous, as they simply replicate the nighttime values.

FeatureFinderQC (all 5 km layer products)

Units: NoUnits

Format: UInt_16

Valid Range: 0, 32767

Description: 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.

Normalization_Constant_Uncertainty (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.0

Fill Value: -9999.0

Description: Uncertainty in the 532 nm and 1064 nm calibration constants due solely to random error in the backscatter measurements in the calibration region; reported as a relative error (i.e., dC/C) in a N x 2 array, with the 532 nm uncertainties stored in first column, and the 1064 nm uncertainties in the second column.

Scientific Data Sets: Layer Spatial Properties

Layer_Top_Altitude (all layer products)

Layer_Base_Altitude (all layer products)

Units: km

Format: Float_32

Valid Range: -0.5, 8.2 (333mMLay); -0.5, 20.1 (1kmCLay); -0.5, 30.1 (5kmALay, 5kmCLay, 5kmMLay)

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Fundamentals of the CALIOP layer detection algorithm are explained in detail in Vaughan et al., 2009, with subsequent enhancements described in Vaughan et al., 2010 (aerosol base extension), Tackett et al., 2021 (smoke base refinements), and Tackett et al., 2022 (boundary layer cloud clearing).

The layer top and base altitudes detected by this scheme are reported in units of kilometers above mean sea level. Because CALIOP does onboard data averaging to reduce downlink data volume, and because the amount of averaging varies with altitude (Hunt et al., 2009), the precision of the layer boundaries determination is itself a function of altitude. In the lower troposphere, between -0.5 km and ~8.2 km, the vertical resolution of the lidar is 30 m. In the upper troposphere, between ~8.2 km to ~20.2 km, the vertical resolution is 60 m. In the lower stratosphere, between ~20.2 km to ~30.1 km the vertical resolution is 180-meters. Above ~30.1 km and below –0.5 km, the vertical resolution is 300 m.

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. 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.) In the 5 km layer products, the signal attenuation can be calculated directly using the 'Overlying Particulate Optical Depth 532' SDS; e.g., for an overlying optical depth of $\tau = 1$ in a feature having a multiple scattering factor of $\eta = 1$, the signal is attenuated by a factor of $\exp(-2 \cdot \eta \cdot \tau) = 0.135$. In the 333 m and 1 km products, the 'Overlying Integrated Attenuated Backscatter' and 'Layer Integrated Attenuated Backscatter QA Factor' SDSs serve as optical depth proxies, and thus allow qualitative assessments of the confidence that users should assign to the reported layer properties.

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 transparent 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 (Hunt et al., 2009) can also make base determination less certain.

e. The Opacity Flag is used to indicate features that completely attenuate the backscatter signal. For these features, the base altitude reported must be considered as an "apparent" base rather than a true base.

f. In those cases where the layer base has been extended to 90 m above the local surface (Vaughan et al., 2010), the assumption is that extended region contains aerosol that lies below the detection limits of the standard algorithm. The resulting increase in aerosol optical depths indicates that this procedure is appropriate far more often than not.

g. Stratospheric features reported during daylight – especially those reported above 20 km between 60°N and 60°S – are often noise artifacts and should be treated with suspicion. These false-positive feature detections typically have no confidence CAD scores ($|CAD| < 20$) and are often misclassified as volcanic ash (Tackett et al., 2023).

Horizontal_Averaging (all 5 km layer products)

Units: km

Format: Int_8

Valid Range: 5, 80

Fill Value: 0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The amount of horizontal averaging required for a feature to be detected. The values in this field will be either 0, 5, 20, or 80, with 0 being a fill value. The remaining values indicate features detected at, respectively, 5 km, 20 km, and 80 km horizontal averaging intervals.

Number_Layers_Found (all layer products)

Units: NoUnits

Format: Int_32

Valid Range: 333mMLay = 0–5

5kmALay = 0–8

1kmCLay, 5kmCLay = 0–10

5kmMLay, 5kmDiag = 0–15

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The number of layers found in a column; cloud layer products report (only) the number of cloud layers found, aerosol layer products report (only) the number of aerosol layers found, and the merged layer products report the numbers of cloud and aerosol layers found. Interpretation of the number of layers found parameter is straightforward for the 1 km and 333 m layer products: individual layers are always separated by regions of "clear air", and layer boundaries never overlap in the vertical dimension. However, this simplicity of interpretation does not always carry over into the 5 km cloud and aerosol layer products. CALIPSO uses a nested multi-grid feature finding algorithm (Vaughan et al., 2009), and thus the search for layer boundaries is conducted independently at multiple horizontal averaging resolutions. While the 1 km and 333 m layer products report only those features detected at, respectively, averaging resolutions of 1 km and 1/3 km, the 5 km products report layers detected at multiple averaging resolutions (5 km, 20 km, and 80 km). Because the reporting resolution (5 km) is not always identical to the detection resolution, layers may appear to overlap in the vertical dimension.

Figure 2 shows a wholly fictitious but heuristically useful schematic of layer detection results for a data segment extending 80-km horizontally and 465-m vertically. Yellow/orange/brown colors indicate an aerosol layer detected at horizontal averaging resolutions of, respectively, 80, 20 or 5 km. Shades of blue likewise represent clouds detected at 80, 20, and 5 km resolutions. The white regions are (presumably) clear air, where no features were found. The labeled rows at the bottom indicate the 'number of layers found' that will be reported in the cloud and aerosol layer products for each 5 km column. In column 16, the layer labeled F5 (top altitude = 0.285 km, base altitude = 0.165 km) appears to vertically overlap F6 (top altitude = 0.255 km, base altitude = 0.135 km), which in turn appears overlap F7 (top altitude = 0.225 km). However, F5 was detected at an averaging resolution of 5-km, and hence the backscatter data that comprises F5 is *removed from consideration* before construction the 20-km horizontally averaged profile in which F6 was detected. Similarly, the backscatter data from both F5 and F6 were removed from consideration before constructing the 80-km averaged profile in which F7 was detected. Layers detected at higher spatial resolutions are thus seen to **overwrite, rather than overlap** apparently collocated layers detected at coarser spatial resolutions. One strategy for accommodating these vertically adjacent features is described in Thorsen et al., 2013.

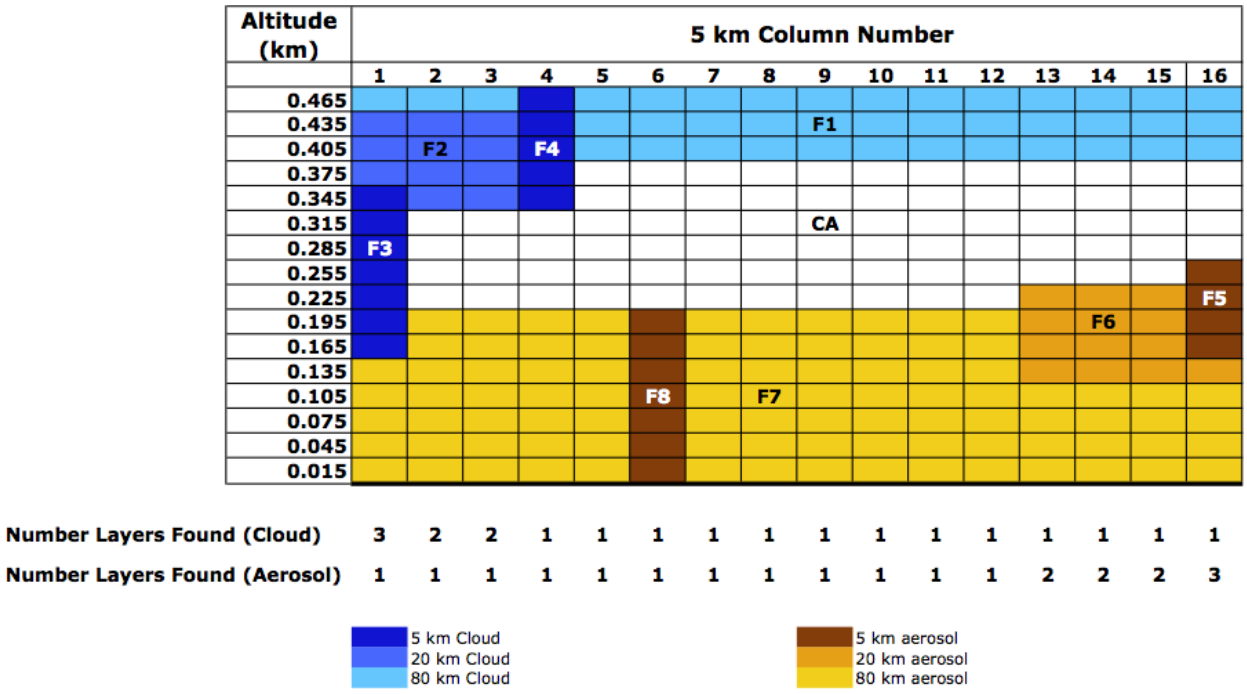


Figure 2: Diagram illustrating the interpretation of the 'number of layers' parameter. Detected features are enumerated (F1 through F8) and color coded according to feature type and averaging required for detection Note that layers detected at different horizontal averaging resolutions can be vertically adjacent to one another (e.g., F5, F6, and F7). If these layers are classified as the same type (e.g., all three are classified as dust), it may be appropriate to consider each feature as a fragment of a larger structure rather than as discrete layers.

High_Resolution_Layers_Cleared (all 5 km layer products)

Units: NoUnits

Format: UInt_16

Valid Range: 0, 32767

Description: Bit-mapped integer identifying which single-shot resolution columns in a 5-km average contain clouds detected and removed by the single-shot boundary layer cloud-clearing algorithm (Vaughan et al., 2009; Tackett et al., 2022). The flags are implemented as 16-bit integers, with the most significant bit unused. For example, if clouds are removed in the 1st, 2nd, and 3rd columns (ordered chronologically) but not in any of the other columns, the high resolution layers cleared would be '0000000000000111' (i.e., 7 in decimal notation).

Single_Shot_Cloud_Cleared_Fraction (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Layers detected in the planetary boundary layer (PBL) are subjected to an additional cloud clearing procedure to separate small-scale boundary layer clouds from any surrounding PBL aerosols at the highest possible spatial resolution (i.e., single shot data; see Vaughan et al., 2009 and Tackett et al., 2022). The single shot cloud cleared fraction reports the fraction of the nominal layer area (i.e., horizontal averaging distance times layer height) that was removed by the cloud clearing process.

Was_Cleared (333mMLay)

Units: NoUnits

Format: Int_8

Valid Range: 0, 1

Fill Value: -127

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Single shot layers that were cleared (i.e., removed) by the boundary layer cloud-clearing algorithm (Vaughan et al., 2009; Tackett et al., 2022) prior to reaveraging the surrounding data for further feature searches; 0 = not cleared, 1 = cleared.

Decoupled_Layer_Type (333mMLay, 5kmALay, 5kmCLay, 5kmMLay, 5kmDiag)

Decoupled_Layer_Status (5kmALay, 5kmCLay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Int_8

Valid Range: -1, 9 (Type); -1, 10 (Status)

Fill Value: -127

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Quality assurance parameters used to document the outcome of the revised boundary layer cloud clearing (BLaCC) algorithm (see Figure 5) first implemented in the V4.51 release (Tackett et al., 2022). The 'Decoupled Layer Type' (see Table 2) identifies the specific path within the BLaCC algorithm that determines the mechanism for partitioning clouds and aerosols within boundary layer features. The 'Decoupled Layer Status' (see Table 3) assigns a BLaCC classification to each layer.

Within the context of CALIOP's layer detection scheme, "Features are composed of a generic substance called particulates and are defined as any extended, vertically contiguous region of enhanced backscatter that rises significantly above the signal magnitude expected from a purely molecular atmosphere." (Vaughan et al., 2009) An unspoken assumption in this definition is that the features being identified are vertically homogeneous, irrespective of whether they are subsequently classified as clouds or aerosols. Recognizing the fallacy of this assumption, the boundary layer cloud clearing algorithm was designed specifically to identify and decompose heterogeneous features with layer tops below a somewhat arbitrary upper bound of 4 km AMSL. The conceptual model used in designing the initial iteration of this algorithm was the presence of small-scale fair weather cumulus embedded within ubiquitous boundary layer aerosols. After making corrections to adapt the prelaunch implementation to on-orbit measurements (Vaughan et al., 2010), the original algorithm performed as expected in the vast majority of cases. But it consistently failed in cases where dense aerosol plumes lay immediately above extended stratus decks. Figure 3 illustrates this phenomenon with an example of dust lofted above opaque stratus measured in the middle of the Atlantic Ocean on 15 June 2013. The left panel shows a time history of 532 nm attenuated backscatter coefficients. The right panel shows a 5 km averaged profile extracted at the vertical line shown in the left panel. From the top of the dust layer, just above 4 km, to the apparent base of the stratus cloud, just below 1 km, the measured backscatter signal never falls below the detection threshold, hence this dust-over-stratus combination is considered to be a single feature. Because the stratus deck is detected in all 15 single shot profiles used to construct the 5 km average, earlier versions of the BLaCC algorithm assumed the feature is a cloud and hence did not attempt to separate the cloud from the aerosol. However, further along in the processing chain, the CAD algorithm will frequently identify these heterogenous features as aerosol. The malign consequence of this miscommunication between the two algorithms is that cloud contamination causes the resulting aerosol extinction retrievals to be biased unacceptably high. While the occurrence frequency of these heterogenous features is low, affecting ~2.7% of all boundary layer features detected at 5 km resolution, the subsequent perturbation to the high end of the aerosol extinction distribution is enormous.

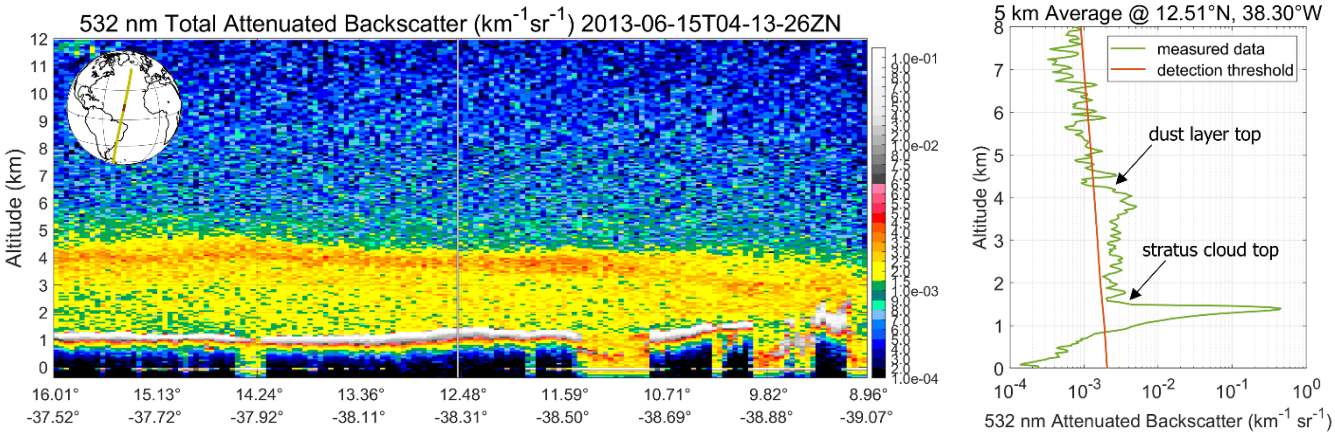


Figure 3: left panel shows a time history of dust lofted above opaque stratus measured 15 June 2013. The vertical line in the left panel marks the location of the 5 km (15 shot) averaged profile shown in the right panel. The right panel shows the averaged data in green and the layer detection threshold in orange. The magnitude of the data first rises continuously above the threshold at the top of the dust layer and remains above the threshold until the signal becomes totally attenuated at the base of the stratus layer. Consequently, two distinct, vertically adjacent layer types are identified as being a single feature.

To remedy the clash of feature classifications between the BLaCC and CAD algorithms, the CALIPSO lidar science working group (LSWG) substantially redesigned the BLaCC for the V4.51 release (Tackett et al., 2022). In doing so, the LSWG introduced the following new descriptive terminology.

- A "Mother Layer" is a boundary layer feature originally detected at 5-km resolution and in which clouds are subsequently detected by the BLaCC in all 15 single shot profiles. The heterogenous feature shown in the right panel of Figure 3 is an example of a mother layer.
- Because multiple clouds can be detected in any single shot profile with a mother layer, a "Fused Layer" is created by first identifying the highest CAD score among all clouds detected in each single shot profile. The top altitude of the fused layer is then set equal to the highest top altitude found within this subset of 15 clouds. The fused layer base altitude is set equal to the base of the mother layer. Finally, the backscatter data in the 15-shot horizontal extent bounded vertically by the fused layer top and base are then "fused" into a new 5 km layer that, by definition, is classified as a cloud. The stratus portion of the heterogenous feature shown in the right panel of Figure 3 is an example of a fused layer.
- A "Severed Layer" is created by separating the fused layer from the mother layer so that the original heterogenous feature detected at 5-km resolution is partitioned into a fused layer and an overlying, vertically adjacent severed layer. In the right panel of Figure 3, the severed layer would extend from the top of the overlying dust layer to one range bin above the top of the stratus cloud below.

Figure 4 shows a more complex conceptual diagram of a heterogenous feature in which annotations identify the mother layer, the fused layer, the severed layer, and all clouds and aerosols detected at single shot resolution. Note that even though a fused layer is created, the severed layer can also contain single shot clouds. While the BLaCC automatically and irrevocably classifies fused layers as clouds, it relies on an invocation of the CAD algorithm to classify the severed layer as either cloud or aerosol.

Figure 5 illustrates the full complexity of this new addition to the BLaCC. The 'Decoupled Layer Type' and 'Decoupled Layer Status' parameters identify the specific algorithm path that was followed and how the classification of the severed layer was determined.

The modified algorithm successfully decouples layers containing single shot resolution clouds in every profile for the vast majority of those cases for which decoupling is appropriate. On rare occasions, layers can be classified as aerosol even though they 15 single-shot resolution clouds. Although the occurrence frequency of these errors is very low (a per profile error rate less than 2.2×10^{-5} %), to further reduce cloud contamination of the aerosol retrievals, the LSWG implemented an extinction post-processor to identify these anomalous cases. When an error is found, the optical depth and all extinction coefficients for the

erroneous layer are assigned a flag value of -444, as are all extinction coefficients in the 5 km column below the erroneous layer.

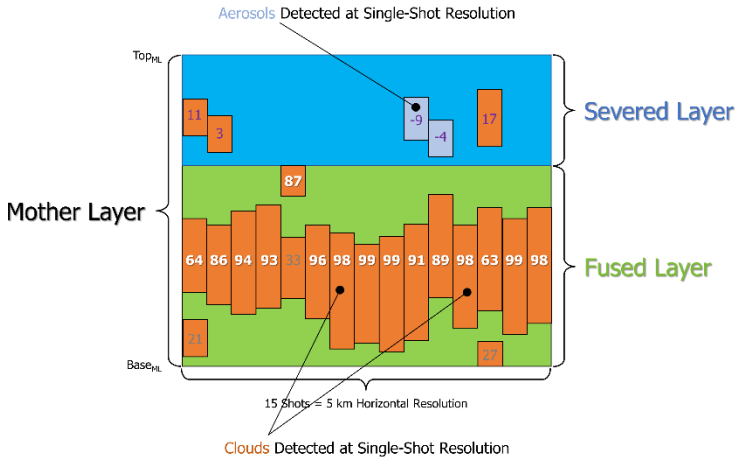


Figure 4: Schematic of a 5 km scene illustrating fused layer determination by identifying the maximum CAD score assigned to layers detected within each single shot profile. CAD scores for the features selected to define the upper boundary of the fused layer are shown in white.

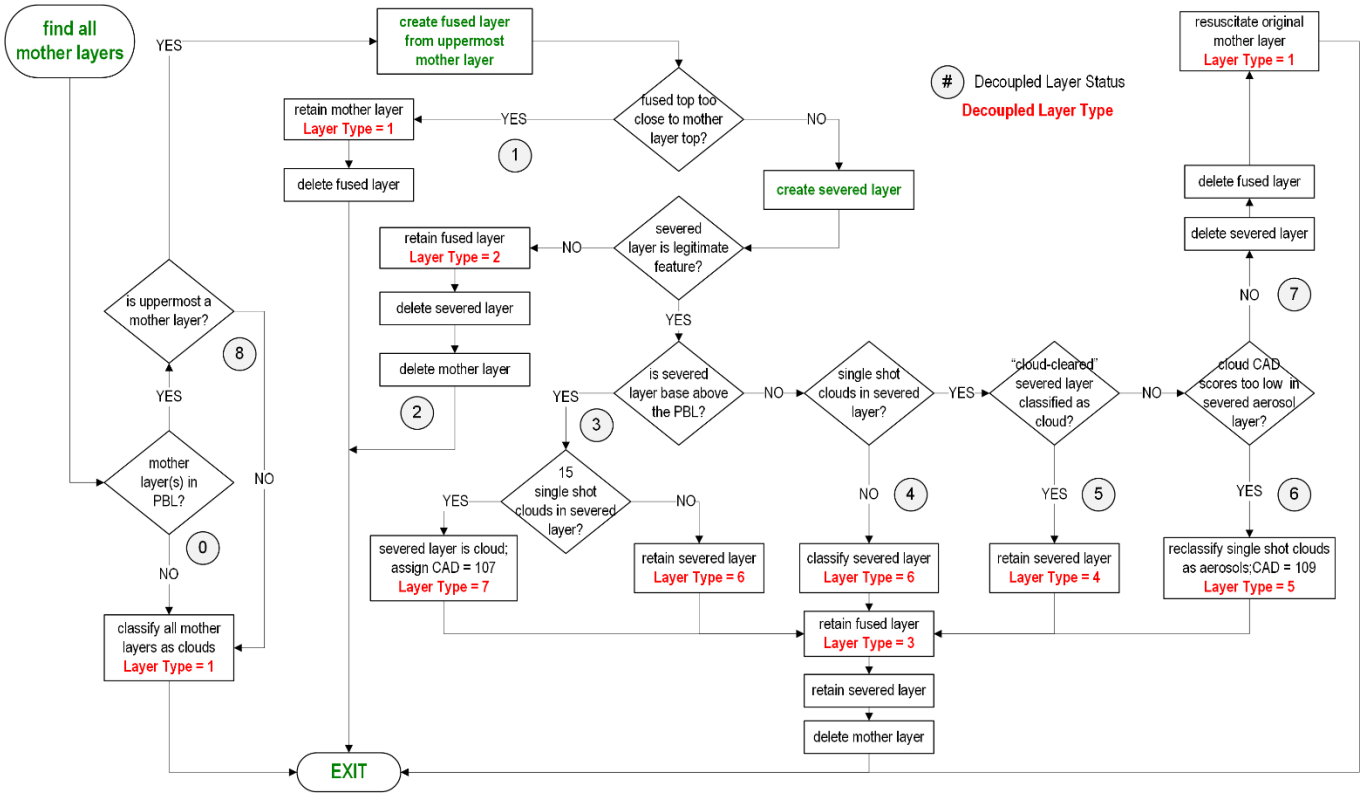


Figure 5: Flow chart detailing the updates to the boundary layer cloud clearing algorithm. Assignments of decoupled layer status are indicated by circled numbers (0–8). Decoupled layer type definitions are given in bold red type.

Table 2: Interpretation for the integers reported in the Decoupled Layer Type SDS

Type	Decoupled Layer Type Definition
-1	Not applicable (fill value)
0	Layer is not in the boundary layer (not used in V5.00)

Type	Decoupled Layer Type Definition
1	Is mother layer; not partitioned
2	Is truncated mother layer; top altitude revised downward (lower)
3	Fused cloud separated from a mother layer; layer type determined by the BLaCC criteria
4	Severed cloud separated from a mother layer; layer type determined by the BLaCC criteria
5	Severed aerosol separated from a mother layer; layer type determined by the BLaCC criteria
6	Severed layer separated from a mother layer; layer type determined by the CALIOP CAD algorithm
7	Severed layer: clouds detected in all single shot profiles
8	Unused
9	BLaCC algorithm error

Table 3: Interpretation for the integers reported in the Decoupled Layer Status SDS

Status	Decoupled Layer Status Definition
-1	Not applicable (fill value)
0	Detected mother layer is not in the boundary layer
1	Fused layer top too close to mother layer top; delete fused layer, retain original mother layer
2	Severed layer is not recognized as a legitimate feature; retain fused layer, delete mother layer and severed layer
3	Severed layer is base is above PBL; retain fused and severed layers, delete mother layer
4	No clouds detected at single shot resolution with a severed layer; retain fused and severed layers, delete mother layer
5	Severed layer classified as cloud after removing clouds detected at single shot resolution; retain fused and severed layers, delete mother layer
6	Severed layer classified as aerosol; single shot detections within severed aerosol layer are not confidently classified as clouds; redefine all single shot detections as aerosols; retain fused and severed layers, delete mother layer
7	Severed layer classified as aerosol; single shot detections within severed aerosol layer are confidently classified as clouds; delete fused and severed layers, retain mother layer
8	Detected mother layer is not the uppermost mother layer in the boundary layer
9	BLaCC algorithm completed with errors
10	BLaCC algorithm failed

Unique_Layer_ID (all 5 km layer products)

Units: NoUnits

Format: Int_32

Valid Range: 1, 83475

Fill Value: -9999

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Unique integer assigned for each layer detected within a granule at 5-km or coarser resolution (i.e., at all resolutions for which layer optical properties are retrieved). Numbering begins at 1 in each granule. Values are replicated as necessary, so that in the 5 km layer products, the unique layer ID for a layer detected at 80 km resolution could appear as many as 16 times. Unique layer IDs are also reported in the CALIOP profile products, thus allowing direct two-way mapping between the altitude-resolved data reported in the profile products and the integrated feature statistics reported in the layer products. The numbering of unique layers restarts anew at 0 for each granule.

Number_Bins_Shift (333mMLay, 01kmCLay)

Units: NoUnits

Format: Int_32

Valid Range: -8, 8

Description: Number of 30-meter altitude bins shifted by the altitude registration algorithm to achieve the best match between apparent MSL in the measured profile and MSL in CALIOP's fixed altitude array. In the altitude registration process, the bin shift required to register the measured data to CALIOP's fixed altitude grid is designated as positive whenever MSL in the measured data is higher than MSL in the fixed altitude grid. Similarly, whenever MSL in the measured data is lower, the bin shift is designated negative. A schematic of the bin shifting geometry is shown in Figure 6. For measured off-nadir angles very near the commanded CALIOP off-nadir angle, the Number of Bins Shifted is approximately equal to round(Surface Altitude Shift / (0.029979 × cos(off-nadir angle))) (units = km).

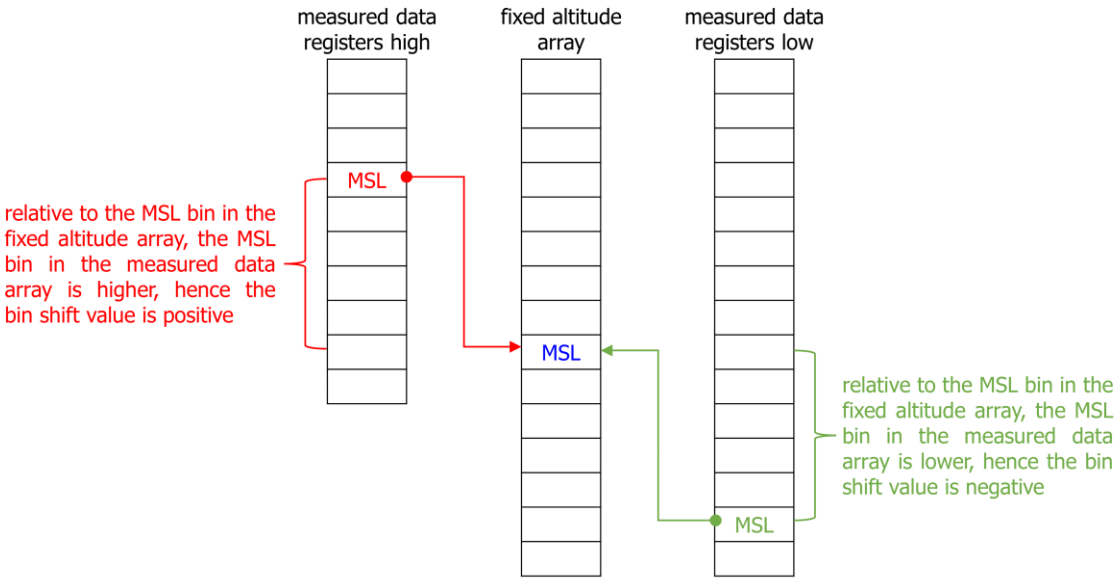


Figure 6: schematic view of the bin shifting that occurs in the CALIOP altitude registration process.

Scene_Flag (all layer products)

Units: NoUnits

Format: Int_32

Fill Value: -9999

Description: Bit mapped 32-bit integer in which individual bits are toggled to indicate the detection, at any averaging resolution, of specific atmospheric feature types within each 5 km column. The interpretations of the scene flag bits are given in Table 4.

Table 4: bit interpretation for the integers reported in the Scene Flag SDS. Bit 0 is the least significant bit.

Bit	Interpretation
0	Tropospheric aerosol, subtype = marine
1	Tropospheric aerosol, subtype = dust
2	Tropospheric aerosol, subtype = polluted continental/smoke
3	Tropospheric aerosol, subtype = clean continental
4	Tropospheric aerosol, subtype = polluted dust
5	Tropospheric aerosol, subtype = elevated smoke
6	Tropospheric aerosol, subtype = dusty marine
7	Stratospheric aerosol, subtype = polar stratospheric aerosol
8	Stratospheric aerosol, subtype = volcanic ash
9	Stratospheric aerosol, subtype = sulfate
10	Stratospheric aerosol, subtype = elevated smoke
11	Stratospheric aerosol, subtype = unclassified
12	Column has ROI phase clouds at any resolution
13	Column has HOI phase clouds at any resolution
14	Column has water phase clouds at any resolution
15	Column has unknown phase clouds at any resolution

Scientific Data Sets: Layer Meteorological Parameters

Layer_Top_Pressure (all layer products)

Midlayer_Pressure (all layer products)

Layer_Base_Pressure (all layer products)

Units: hPa

Format: Float_32

Valid Range: 1.0, 1086.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Atmospheric pressures at, respectively, the top, geometric midpoint, and base of the layer in the vertical dimension; derived from the MERRA-2 meteorological reanalysis products.

Layer_Top_Temperature (all layer products)

Layer_Centroid_Temperature (333mMLay, 1kmCLay, 5kmCLay, 5kmMLay, 5kmDiag)

Midlayer_Temperature (all layer products)

Layer_Base_Temperature (all layer products)

Units: °C

Format: Float_32

Valid Range: -110.0, 60.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Temperature at, respectively, the top, the 532 nm total attenuated backscatter centroid altitude, the geometric midpoint, and the base of the layer in the vertical dimension; derived from the MERRA-2 meteorological reanalysis products. The 532 nm total attenuated backscatter centroid altitude is reported in the Attenuated_Backscatter_Statistics_532 SDS.

Relative_Humidity (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.5

Fill Value: -9999.0

Description: Relative humidity at the geometric midpoint of the layer in the vertical dimension; derived from the MERRA-2 meteorological reanalysis products.

Scientific Data Sets: Layer Measured Optical Properties

Integrated_Attenuated_Backscatter_532 (all layer products)

Integrated_Attenuated_Backscatter_1064 (all layer products)

Units: 1/sr

Format: Float_32

Valid Range: 0.0, 1.8

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The 532 nm and 1064 nm integrated attenuated backscatters (γ'_{532} and γ'_{1064}) are calculated for each layer by integrating the profiles of total attenuated backscatter coefficients from layer top to layer base. Prior to integration, the backscatter data are corrected for wavelength-specific signal attenuation by any overlying layers. At 532 nm, the contribution to the integral made by molecular scattering from within a layer is estimated using the 'clear air trapezoid' technique described in Vaughan et al., 2010. Since molecular backscatter is a factor of ~17 time weaker at 1064 nm than at 532 nm, and because the SNR in the 1064 nm channel is notably lower than at 532 nm (Pauley et al., 2019), clear air trapezoid corrections are not deemed useful or necessary when calculating γ'_{1064}.

At 532 nm, the quality of the γ'_{532} estimate in the uppermost layer in any column is a function of the accuracy of the top and base identification, uncertainties in the backscatter coefficients at the top and the base, the reliability of the 532 nm channel calibrations, and the signal-to-noise ratio (SNR) of the backscatter data within the layer. For layers beneath the uppermost, the signals must be "renormalized" to account for attenuation caused by the overlying layers (Young & Vaughan, 2009). In practice this means that once an extinction solution has been obtained for a layer, all attenuated backscatter coefficients below the layer base are divided by the two-way particulate transmittance, $T^2_{p,532}$, computed for the layer. As a result, the quality of the γ'_{532} calculation for lower layers also depends on the accuracy $T^2_{p,532}$ for all overlying layers. The most reliable $T^2_{p,532}$ estimates come from constrained solutions to the lidar equation (Young & Vaughan, 2009). Uncertainties in estimates of overlying $T^2_{p,532}$ derived from unconstrained solutions will be further amplified by uncertainties in the choice of the lidar ratio(s) and multiple scattering factor(s) used in the extinction solutions for the overlying layer(s) (Young et al., 2013).

The effects of errors caused by misestimating T^2 can increase sharply as the optical thickness above a layer increases: small errors in large optical depths can yield substantial errors in the renormalization of the underlying features. For the 5 km layer products, 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} in the 5 km layer products. These cases are flagged with a special CAD score of 103. However, because extinction solutions are only derived for data averaged to a 5-km (or greater) resolution, the γ'_{532} values reported in the 1 km and 333 m layer products are not corrected for the signal attenuation effects imparted by overlying layers.

The values reported for γ'_{532} should always be positive, and for the results derived directly from the layer detection algorithm (i.e., as reported in the 1 km and 333 m layer products) this is always true. However, in the 5 km products there are two exceptions. First, for layers detected at 5 km resolution only, there are some anomalous cases where negative values occur due to unresolved cloud clearing artifacts. These cases are flagged with a special CAD score of –101 and are extremely rare; there are 238 instances among all 732,601,646 unique layers reported in 5 km merged layer products for the whole mission. Layers with CAD scores of –101 should be excluded from all science investigations.

The second exception is for layers detected after averaging to 20 km or 80 km horizontally, where negative γ'_{532} values can arise when overlying layers detected at finer spatial resolutions have vastly different optical depths. Negatives can occur because this situation leads to very different weightings (i.e., different renormalization factors) of the reaveraged data originally included in the 20 km or 80 km horizontal extent. These layers are assigned a special CAD score of 105 and occur slightly more often than layers with CAD scores –101 (14,477 instances among all 732,601,646 unique layers). While the spatial properties of CAD = 105 layers may be trustworthy, all measured and derived optical properties for these layers are unreliable and should be ignored. In evaluating the reliability of the spatial properties of these layers, users should carefully consider the layer IAB QA factor.

For γ'_{1064}, measurement quality in the uppermost layer is governed by many of the same factors as γ'_{532}; i.e., 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. However, unlike at 532 nm, because the molecular backscatter signal is so faint, reliable estimates of $T^2_{p,1064}$ cannot be derived from an analysis of the 1064 nm backscatter signal in (assumed to be) clear air regions, and thus the $T^2_{p,1064}$ corrections for the attenuation from overlying layers are always obtained from extinction solutions that use prescribed values of lidar ratio and multiple scattering factor. Once again similar to 532 nm, no $T^2_{p,1064}$ corrections are applied to the γ'_{1064} values reported in the 1 km and 333 m layer products. Furthermore, because the CALIOP layer detection algorithm examines only the 532 nm backscatter signals at all resolutions other than single shot (Vaughan et al., 2009), negative (i.e., non-physical) values may occasionally be reported for γ'_{1064} in the 1 km and 5km resolutions of the layer products. Unlike the layers for which γ'_{532} is negative, layers with negative γ'_{1064} are not indicated by a special CAD score. Negative values of γ'_{1064} occur most often for very weakly scattering layers (e.g., faint aerosols) and in those layers for which the backscatter signal has been highly attenuated by other, overlying layers.

Integrated_Attenuated_Backscatter_Uncertainty_532 (all layer products)

Integrated_Attenuated_Backscatter_Uncertainty_1064 (all layer products)

Units: 1/sr

Format: Float_32

Valid Range: 0.0, 0.5

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Estimated uncertainties for the 532 nm and 1064 nm integrated attenuated backscatter values. These uncertainties quantify the random error in the integral of the attenuation-corrected backscatter signal. The general procedure used for calculating uncertainties for integrated quantities is described in section 6 of the CALIPSO Feature Detection ATBD.

Initial_Integrated_Attenuated_Backscatter_532 (5kmDiag)

Initial_Integrated_Attenuated_Backscatter_1064 (5kmDiag)

Units: 1/sr

Format: Float_32

Valid Range: 0.0, 1.8

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The 532 nm and 1064 nm integrated attenuated backscatter (IAB) computed for each layer with no corrections applied for signal attenuation by overlying layers. At 532 nm, these values are required for computing the "Column Particulate Optical Depth Above Opaque Water Cloud 532" parameter (Hu et al., 2007). The 1064 nm value is essential for retrieving estimates of 1064 nm above-cloud optical depths using the method outlined in Chand et al., 2008.

Initial_Integrated_Attenuated_Backscatter_Uncertainty_532 (5kmDiag)

Initial_Integrated_Attenuated_Backscatter_Uncertainty_1064 (5kmDiag)

Units: 1/sr

Format: Float_32

Valid Range: 0.0, 0.5

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Estimated uncertainties of the initial 532 nm and 1064 nm integrated attenuated backscatter values. These uncertainties quantify the random error in the integral of the 532 nm and 1064 nm backscatter signals.

Attenuated_Scattering_Ratio_Statistics_532 (all layer products)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Attenuated scattering ratios are calculated by dividing a profile of measured attenuated backscatter coefficients by a profile of the molecular only attenuated backscatter coefficients derived from collocated atmospheric model data (e.g., MERRA-2). The 532 nm attenuated scattering ratios are used extensively in the CALIOP layer detection algorithm. This SDS reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the 532 nm attenuated scattering ratio coefficients for each layer detected. Formulas used for each of the statistical calculations are given in section 6 of the CALIPSO Feature Detection ATBD.

Attenuated_Backscatter_Statistics_532 (all layer products)

Attenuated_Backscatter_Statistics_1064 (all layer products)

Units: 1/(km · sr)

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the 532 nm and 1064 nm total attenuated backscatter coefficients for each layer.

Integrated_Volume_Depolarization_Ratio (all layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The 532 nm layer integrated volume depolarization ratio (IVDR), calculated as the sum of the perpendicular channel attenuated backscatter coefficients divided by the sum of the parallel channel attenuated backscatter coefficients. The quality of the IVDR estimate 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 ratios (SNR) of the two channels of backscatter data within the layer. In general, IVDR estimates are highly reliable. Validation studies show that the CALIOP depolarization ratios agree well with coincident measurements by ground-based lidars (Tesche et al., 2013; Kim et al., 2014; Dai et al., 2018). The integrated volume depolarization ratio is a key parameter used throughout CALIOP's level 2 scene classification algorithms (e.g., CAD, aerosol subtyping, and cloud thermodynamic phase).

Integrated_Volume_Depolarization_Ratio_Uncertainty (all layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 2.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: 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 specific formula used to generate this parameter is given by equation 6.11 in the CALIPSO Feature Detection ATBD.

Volume_Depolarization_Ratio_Statistics (all layer products)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the altitude-resolved 532 nm volume depolarization ratios for each layer. The mean value reported here is the mean of the ratios, whereas the layer integrated value discussed previously is the ratio of the means.

In regions with sufficient SNR, depolarization ratio accuracy will depend almost entirely on the accuracy of the polarization gain ratio calibration. Unfortunately, in other regions the physical interpretation of the volume depolarization ratios can be somewhat obscure. This is because each of the range resolved depolarization ratios within any layer is the ratio of two noisy measurements. Especially where feature backscattering is relatively weak, and in regions of low SNR, the measured 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. Consequently, when using volume depolarization ratios to characterize cloud and/or aerosol layers, the layer median and (especially) the layer integrated value are both more reliable metrics than the mean.

Integrated_Attenuated_Total_Color_Ratio (all layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 2.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The integrated attenuated total color ratio is the quotient of the integrals of the total attenuated backscatter coefficients at the two wavelengths (1064 nm / 532 nm) computed from layer top to layer base. The 1064 nm term is identical to the 1064 nm integrated attenuated backscatter. However, in the color ratio calculation the 532 nm term is not corrected for in-layer molecular backscatter contributions. The quality of the estimate is wholly determined by the quality of the two integrated backscatter terms. For the 5 km layer products, the attenuated backscatter coefficients used in the color ratio calculation are corrected for the estimated overlying two-way transmittance. No such correction is attempted for the 1 km and 1/3 km values, as no extinction solution is computed at these resolutions. The integrated attenuated total color ratio is a key parameter used in the cloud-aerosol discrimination algorithm.

Integrated_Attenuated_Total_Color_Ratio_Uncertainty (all layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 3.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The uncertainties reported for the layer-integrated attenuated total color ratios, which provides an estimate of the total random error in the combined backscatter signals (i.e., at 532 nm and 1064 nm). The specific formula to derive this parameter is given by equation 6.14 in the CALIPSO Feature Detection ATBD.

Attenuated_Total_Color_Ratio_Statistics (all layer products)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Profiles of total attenuated backscatter color ratios are calculated by dividing the 1064 nm total attenuated backscatter coefficient by the 532 nm total attenuated backscatter coefficient at each CALIOP altitude bin. This SDS reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the total attenuated backscatter color ratios within each layer. As with the 532 nm depolarization ratio statistics, the meaning of the various color ratio quantities can be somewhat misleading. Like the depolarization ratios, the attenuated total color ratios are produced by dividing one noisy measurement by a second noisy measurement that, in turn, is the sum of two noisy measurements. Depending on the noise in any pair of samples, the resulting ratios 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. As with the volume depolarization ratios, if/when total attenuated backscatter color ratios are used to characterize cloud and/or aerosol layers, the layer median and (especially) the layer integrated value are both more reliable metrics than the mean.

Overlying_Integrated_Attenuated_Backscatter_532 (all layer products)

Units: 1/sr

Format: Float_32

Valid Range: 0.0, 2.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: For any layer, overlying integrated attenuated backscatter, γ'_{above}, is the integral with respect to altitude of the 532 nm total attenuated backscatter coefficients between the top of the atmosphere (TOA; i.e., the first range bin in any L1B profile) down to the range bin immediately above the layer top altitude. γ'_{above} provides a qualitative assessment of the layer boundary detection confidence that users can apply to each layer reported. In general, detection confidence diminishes as γ'_{above} increases. γ'_{above} is obtained directly from the calibrated backscatter signal, and hence can also provide a crude proxy for the optical depth above each layer, irrespective of the averaging required for feature detection. 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 or immediately 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 be obtained directly from the calibrated backscatter signal, and hence can provide a qualitative proxy for the optical depth above each layer detected.

Measured_Two_Way_Transmittance_532 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Provides the measured value of the layer two-way transmittance, derived by computing the ratio of the mean attenuated scattering ratios in the "clear air" regions immediately below and above the layer. This calculation can only be accomplished for transparent layers that are not vertically adjacent to other layer types. See the discussion of constrained retrievals in Young & Vaughan, 2009. Two-way transmittances are 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 one or both of the "clear air" regions, the values reported in the CALIOP data products will sometimes exceed these bounds.

Measured_Two_Way_Transmittance_Uncertainty_532 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

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

Two_Way_Transmittance_Measurement_Region (all 5 km layer products)

Units: km

Format: Float_32

Valid Range: 0.0, 30.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Provides the top and base altitudes of the clear air region used to measure the two-way transmittance below a transparent layer.

Opacity_Flag (all layer products)

Units: NoUnits

Format: Int_8

Valid Range: 0, 1

Fill Value: 99

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Flag that identifies a layer in which the backscatter signal becomes completely attenuated. An opacity flag of 0 indicates a transparent layer, while a value of 1 indicates an opaque layer. 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. For those layers identified as opaque, the reported base altitude must be considered as an apparent base, rather than a true base.

Scientific Data Sets: Layer Derived Optical Properties

Feature_Optical_Depth_532 (all 5 km layer products)

Feature_Optical_Depth_1064 (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 5.0

Fill Value: -9999.0

Flag Values: –7.777 invalid feature

–33.333 failed retrieval

–111.0 data quality degraded due to low laser energy and rejected by LEM

–444.0 improper single-shot cloud clearing

Description: Reports estimates of layer optical depths at 532 nm and 1064 nm. 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, the 5 km cloud and merged layer products report cloud optical depths only for the 532 nm retrievals. CALIOP extinction retrievals for clouds and aerosols use exactly the same retrieval framework. These algorithms are documented in detail in Young & Vaughan, 2009 and Young et al., 2018.

When using any of the CALIOP optical depths 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 of layer type, propagate non-linearly into subsequent estimates of layer optical depth (Young et al., 2013).
- 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 (Hu et al., 2007; Garnier et al., 2015).
- Similar to the layer detection problem, estimates of layer optical depth become increasingly fraught with error in multi-layer scenes, as errors incurred in overlying layers are propagated into the solutions derived for underlying features (Young et al., 2013).

- The optical depths reported for layers flagged as opaque are not accurate representations of the true optical depths. Instead, they report the optical depth that can be retrieved before the CALIOP backscatter signal becomes totally extinguished. In general, the extinction profiles in opaque layers are reliable up to optical depths of ~5 in ice clouds and 3 in aerosols. However, users should be especially cautious about retrievals in opaque water clouds, where multiple scattering rapidly becomes so pervasive that "pulse stretching" eliminates the possibility of accurate ranging into the cloud (Miller and Stephens, 1999; Winker, 2003).

Unlike the 5 km cloud and merged layer products, the V5.00 5 km merged diagnostic beta products report cloud optical depths at both 532 nm and 1064 nm. Though yet to be validated, the 1064 nm optical depths for ice clouds are likely to be reliable (Haarig et al., 2016). However, the 1064 nm optical depth retrievals for water clouds are highly suspicious for several reasons. First, the same lidar ratio is used for retrievals at both wavelengths, even though the lidar ratios are known to be slightly lower at 1064 nm than at 532 nm (e.g., $S_{c,1064}$ ~18.2 sr and $S_{c,532}$ ~18.6 sr; see O'Connor et al., 2004). Second, the integrated attenuated backscatters are notably higher at 1064 nm. Consider all water clouds detected at 5 km resolution during daytime operations between December 2007 and December 2016 for which (a) the layers are totally opaque (i.e., the surface is not detected in any of the 15 single shot profiles) and (b) the retrieved multiple scattering factors (Hu et al., 2007) lie between the physically meaningful limits of 0 and 1. Applying these filtering criteria to the global CALIOP data set yields 19,507,473 samples for which the median integrated attenuated backscatter is ~20% higher at 1064 nm than at 532 nm; i.e., γ'_{1064} = 0.0818 sr^{-1} and γ'_{532} = 0.0682 sr^{-1}. If we assume the lidar ratios given by O'Connor et al., 2004, the multiple scattering factor at 532 nm is $\eta_{c,532} = 1 / (2 \times 18.6 \times 0.0682) = 0.394$ (versus a retrieved median value of 0.395) while the multiple scattering factor at 1064 nm is $\eta_{c,1064} = 1 / (2 \times 18.2 \times 0.0818) = 0.336$. These differences notwithstanding, the CALIOP water cloud retrievals use the same lidar ratios and multiple scattering factors at both wavelengths.

Feature_Optical_Depth_Uncertainty_532 (all 5 km layer products)

Feature_Optical_Depth_Uncertainty_1064 (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values:

–111.0	data quality degraded due to low laser energy and rejected by LEM
–33.333	failed retrieval
–29.0	opaque water cloud
–444.0	improper single-shot cloud clearing
99.99	uncertainty calculation failed

Description: Estimated absolute uncertainty in CALIOP's 532 nm and 1064 nm layer optical depths. Young et al., 2013 provides an in-depth assessment of the uncertainties that propagate nonlinearly through the CALIOP extinction retrieval. The numerical recipe used to estimate extinction and optical depth uncertainties in the CALIOP data products is given in the supplementary material accompanying Young et al., 2018.

Errors in layer optical depth calculations typically arise from three main sources: signal-to-noise ratio (SNR) within a layer, calibration accuracy, and the accuracy of the effective lidar ratio specified for use in the solution. The effective lidar ratio is the product of the lidar ratio and multiple scattering factor (e.g., see Young, 1995). Effective lidar ratio uncertainties are almost always the dominant contributor to optical depth uncertainties. The relative error in the layer extinction profile will always be at least as large as the relative error in the effective lidar ratio and will increase nonlinearly as the solution propagates through the layer and the layer two-way particulate transmittance decreases. CALIOP's calculation of the layer optical depth uncertainty is an iterative process. On some occasions when the SNR is poor, or an inappropriate effective lidar ratio is being used, the iteration will attempt to converge asymptotically to positive infinity. When this situation detected, the iteration is terminated, and the layer optical depth uncertainty is assigned a fixed value of 99.99. Any time an uncertainty of 99.99 is reported, the extinction

calculation should be considered to have failed. The associated optical depths are not reliable and should therefore be excluded from all science studies.

Overlying_Particulate_Optical_Depth_532 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 3.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM
–333.0 failed retrieval
–444.0 improper single-shot cloud clearing

Description: For each feature reported in the 5 km layer products, this SDS provides the particulate optical depth above the feature calculated from the CALIOP 532 nm extinction retrievals. If no features are detected above the value will be zero. These data are reported at 5 km along-track resolution, hence multiple values can be associated with a single unique feature. For example, consider a weak boundary layer aerosol lying beneath a continuous cirrus shield. If the cirrus layer is detected at 5 km resolution and the aerosol can only be detected using 80 km horizontal averaging, it is entirely conceivable that there will be 16 different estimates of cirrus optical depth above the single aerosol detection.

Overlying_Particulate_Optical_Depth_Uncertainty_532 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 3.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM
–333.0 failed retrieval
–444.0 improper single-shot cloud clearing

Description: Estimated uncertainty in the overlying 532 nm particulate optical depths, calculated according to the method describe in Young et al., 2018 and reported at 5 km resolution in all of the CALIOP layer products.

Initial_532_Lidar_Ratio (all 5 km layer products)

Initial_1064_Lidar_Ratio (5kmALay, 5kmMLay, 5kmDiag)

Units: sr

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Retrieving optical depth and profiles of particulate extinction and backscatter coefficients from the 532 nm and 1064 nm CALIOP measurements requires an a priori estimate of the particulate extinction-to-backscatter ratio, commonly known as the "lidar ratio". Table 5 lists CALIOP's different feature types and subtypes, the initial estimates used for each subtype, and a literature source that explains the rationale for each choice. The initial lidar ratios and uncertainties reflect the current state of knowledge based on observations by the NASA Langley Airborne High Spectral Resolution Lidar (HSRL), EARLINET, AERONET, CALIPSO and synergistic multi-sensor retrievals.

Table 5: values and selection methods for the Initial Lidar Ratios

Type	Subtype	532 nm	1064 nm	Reference
cloud	water	Night: S_{water} = 17.7 ± 2.7 sr Day: S_{water} = 18.8 ± 2.8 sr	same as at 532 nm	14

Type	Subtype	532 nm	1064 nm	Reference
cloud	ice	S_{ice} = sigmoid function of cloud centroid temperature ± 25%	same as at 532 nm	1, 2
cloud	unknown phase	0.5 x (S_{water} + S_{ice}(T)) ± max(ΔS_{water}, ΔS_{ice}(T))	same as at 532 nm	
tropospheric aerosol	unknown aerosol	35 ± 14 sr	35 ± 14 sr	
tropospheric aerosol	marine	variable	23 ± 5 sr	15, 3
tropospheric aerosol	desert dust	44 ± 9 sr	44 ± 13 sr	4, 5
tropospheric aerosol	polluted continental/smoke	70 ± 25 sr	30 ± 14 sr	6
tropospheric aerosol	clean continental	53 ± 24 sr	45 ± 24 sr	7, 16
tropospheric aerosol	polluted dust	55 ± 22 sr	48 ± 24 sr	8
tropospheric aerosol	elevated smoke	70 ± 16 sr	30 ± 18 sr	4
tropospheric aerosol	dusty marine	variable	37 ± 15 sr	15, 9
stratospheric aerosol	undetermined	50 ± 20 sr	50 ± 20 sr	
stratospheric aerosol	polar stratospheric aerosol	50 ± 20 sr	25 ± 10 sr	10
stratospheric aerosol	volcanic ash	61 ± 17 sr	44 ± 13 sr	11
stratospheric aerosol	sulfate	50 ± 18 sr	30 ± 14 sr	12
stratospheric aerosol	elevated smoke	70 ± 16 sr	30 ± 18 sr	4
stratospheric aerosol	unclassified	50 ± 18 sr	30 ± 14 sr	13

1. Extinction and Optical Depth Retrievals for CALIPSO's Version 4 Data Release (Young et al., 2018).
2. Constrained CALIPSO retrievals of high-confidence randomly oriented ice clouds (Garnier et al., 2015).
3. HSRL measurements in multiple field campaigns, Müller et al., 2007. No wavelength dependence based on Sayer et al., 2012, Josset et al., 2012, and Papagiannopoulos et al., 2016.
4. CALIPSO constrained retrievals by Liu et al., 2015.
5. HRSL measurements of transported Saharan dust. No wavelength dependence based on Tesche et al., 2009.
6. Microphysical measurements made during NAMMA by Omar et al., 2010 and AERONET cluster analysis by Omar et al., 2005.
7. HSRL measurements for layers classified as clean continental by CALIPSO (Rogers et al., 2014).
8. Microphysical measurements made during NAMMA. Lidar ratio values in agreement with Papagiannopoulos et al., 2016 and Müller et al., 2007.
9. Modeled mixture of dust and marine aerosol, 65/35 by surface area (Kim et al., 2018). Uncertainty larger than dust or marine alone.
10. Theoretical model of supercooled ternary solution at stratospheric pressures, Kim et al., 2018.
11. CALIPSO-constrained retrievals of volcanic ash lidar ratio at 532 nm, Tackett et al., 2023.
12. Independent lidar retrievals of multiple sulfate-dominated volcanic plumes and CALIPSO-constrained lidar ratio retrievals, Kim et al., 2018.
13. Identical to the sulfate initial lidar ratio because this is the smallest lidar ratio expected for stratospheric aerosol layers, thereby minimizing propagated extinction retrieval errors, Tackett et al., 2023.
14. Diurnal Differences in Lidar Ratios for Opaque Water Clouds, Vaughan et al., 2022.
15. Mapping CALIPSO Marine and Dusty Marine Aerosol Lidar Ratios using MODIS AOD Constrained Retrievals and GOCART Model Simulations, Toth et al., 2025.
16. Discussion of the spectral slope of the lidar ratio between 355 nm and 1064 nm from multiwavelength Raman lidar observations, Haarig et al., 2025.

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., 2009; Kim et al., 2018; Tackett et al., 2023). 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 for 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 (Avery et al., 2020) and the aerosol subtype identification procedures (Kim et al., 2018 and Tackett et al., 2023). 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 Young et al., 2018.

The initial lidar ratio for ice clouds is a sigmoid approximation function of the temperature at the layer 532 nm attenuated backscatter centroid (Young et al., 2018). Values decrease from ~35 sr to ~20 sr as the centroid temperature decreases. Initial default lidar ratios are derived from the statistical analysis of several years of constrained retrievals, using only those clouds identified as high-confidence randomly oriented ice. The individual ice cloud lidar ratios are retrieved from the layer apparent two-way transmittance, the layer-integrated attenuated backscatter, and the temperature-dependent multiple scattering factor.

For opaque water clouds, 532 nm lidar ratio estimates are retrieved using the opaque layer form of Platt's equation (Platt, 1973) together with the measured integrated attenuated backscatter and a multiple scattering factor derived from layer integrated volume depolarization ratios (Hu et al., 2007).

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.

Final_532_Lidar_Ratio (all 5 km layer products)

Final_1064_Lidar_Ratio (5kmALay, 5kmMLay)

Units: sr

Format: Float_32

Valid Range: 0.0, 250.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Reports the lidar ratio in use at the conclusion of the extinction processing for each layer. The final lidar ratio may be

1) the initial lidar ratio supplied by the scene classification algorithms (i.e., no lidar ratio changes were needed to achieve a successful extinction solution),
2) the result of incremental reductions to the initial lidar ratio to avoid a non-physical solution,
3) iteratively adjusted to yield a calculated optical depth consistent with a measured layer two-way transmittance (i.e., a constrained retrieval), or
4) for opaque layers only, derived from the measured integrated attenuated backscatter and a measured or modeled multiple scattering factor (Young et al., 2018).

Solution paths (3) and (4) yield the most accurate and precise lidar ratios. 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.

The optical depths and extinction profiles derived in those cases where the layer lidar ratio must be reduced (option 2 above) are, in all likelihood, biased high. The lidar ratio reduction scheme terminates after identifying the largest lidar ratio for which a physically meaningful solution can be generated from the backscatter measured in the layer. Consequently, the optical depths and extinction profiles reported in these cases can only be considered as upper bounds; the true values are somewhat, or perhaps even significantly, lower. Similarly, the uncertainties reported for these layers captures only the random component; the unknown bias error cannot be estimated and hence is not included. Using these data in statistical analyses of layer optical properties can be problematic and users are advised to treat these cases with all due caution.

Final_532_Lidar_Ratio_Uncertainty (all 5 km layer products)

Final_1064_Lidar_Ratio_Uncertainty (5kmALay, 5kmMLay)

Units: sr

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Reports the lidar ratio uncertainty at the conclusion of the extinction processing for each layer. The final lidar ratio uncertainty is different from the lidar ratio uncertainty initially assigned based on feature type (see Table 5) only in those cases where the lidar ratio is optimized based on the value of some external constraint (i.e., constrained solutions and solutions in opaque layers).

Lidar_Ratio_532_Selection_Method (all 5 km layer products)

Lidar_Ratio_1064_Selection_Method (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Int_8

Valid Range: 0, 7

Fill Value: 99

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Specifies the internal procedure used to select the initial lidar ratio for each layer. These selection procedures are described briefly in Table 6.

Table 6: Lidar Ratio Selection Method

Value	Selection Method
0	Not Determined
1	Transmittance Method, constrained retrieval; selected only for those layers with a measurable two-way transmittance lying between 0 and 1; also known as a constrained retrieval
2	Opaque Layer Method; selected for all opaque layers
3	Cirrus Temperature Method; selected for transparent ice clouds that are not suitable for solution via constrained retrieval; i.e., clouds for which the two-way transmittance either cannot be measured or falls outside the bounds of 0 to 1
4	Unknown Cloud Phase; selected for unknown phase transparent clouds having centroid temperatures between -37.0 and 0.0 °C that are not suitable for solution via constrained retrieval
5	Fixed Layer Method; selected for transparent water clouds and aerosols that are not suitable for solution via constrained retrieval (e.g., surface-attached aerosols)

Value	Selection Method
6	Solution at 532 nm Method; selected for 1064 nm solutions in transparent clouds. Because spectral independence is assumed for clouds, the 1064 nm solution is initiated using the 532 nm multiple scattering factor and the 532 nm final lidar ratio.
7	MAC Table Interpolation; lidar ratio assigned via interpolation of the 'Maps of Aerosol lidar ratios for CALIPSO'(MAC) lookup table; applies only to transparent marine and dusty marine aerosol types that are not suitable for solution via constrained retrieval

Layer_Effective_532_Multiple_Scattering_Factor (all 5 km layer products)

Layer_Effective_1064_Multiple_Scattering_Factor (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Float_32

Valid Range: 1.0 (aerosols); 0.01, 1.0 (clouds)

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: The CALIPSO 532 nm and 1064 nm multiple scattering factors range between ~0.2 (for opaque water clouds) and 1.0 (for aerosols), where 1.0 corresponds to the single scattering only limit (i.e., no multiple scattering present) and smaller values indicate increasing contributions to the backscatter signal from multiple scattering.

Multiple scattering effects are different in aerosols, ice clouds, and water clouds. An overview of multiple scattering factors for ice clouds and several aerosol types can be found in Winker, 2003. Garnier et al., 2015 derive CALIOP's temperature-dependent multiple scattering factor model for ice clouds using collocated measurements of transparent layers made by CALIOP and CALIPSO's 3-channel Infrared Imaging Radiometer (IIR; Garnier et al., 2021). Hu et al., 2007 develops the theoretical relationship that enables multiple scattering factors to be directly calculated from the integrated volume depolarization ratios measured in opaque liquid water clouds. While the multiple scattering factor for opaque water clouds will vary, transparent water clouds are assigned a constant multiple scattering factor of 0.6. Based on Monte Carlo simulations of multiple scattering, this value appears to be appropriate for semitransparent water clouds with optical depths less than ~1.

Multiple scattering factors for all aerosol subtypes at both wavelengths are set to unity. Simulations of multiple scattering effects on retrievals of aerosol layer optical depths indicate the effects are small in most cases. There is uncertainty in these estimates, however, due to poor knowledge of aerosol scattering phase functions. For example, based on Monte Carlo studies, Kahnert and Scheirer, 2019 suggest that, for low optical depths or when the aerosol effective radius is not too large, a constant 532 nm multiple scattering factor of ~0.93 could be appropriate. Note, however, that the 7% bias in this value relative to CALIOP's default specification is dwarfed by the relative uncertainties in the CALIOP's aerosol lidar ratio models, which range from 20% to 45% at 532 nm and from 22% to 60% at 1064 nm. Analyses of CALIOP particulate depolarization ratios suggest that, even in dust plumes with a significant coarse mode fraction, multiple scattering effects are largely negligible for aerosol extinction retrievals in layers with optical depths less than ~1 (Liu et al., 2011). At larger optical depths and layer penetration depths, and especially for opaque dust layers, neglecting to account for multiple scattering effects can result in significant underestimates of AOD (Liu et al., 2011; Shcherbakov et al., 2022). However, in these high AOD cases uncertainties in the calibration coefficients and lidar ratios can also produce large retrieval errors.

Integrated_Particulate_Depolarization_Ratio (all 5 km layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Layer-integrated particulate depolarization ratio, which represents the contribution to the volume depolarization ratio that is due only to the cloud and/or aerosol particles within the layer. For non-spherical particles (e.g., ice and dust), the particulate depolarization ratio will normally be higher than the volume depolarization ratio, though how much higher depends on the particulate concentration within the volume. For layers consisting of spherical particles, the particulate depolarization ratio would normally be equal to or even slightly lower than the volume depolarization ratios. However, if the layer optical depth is high (e.g., opaque water clouds), multiple scattering can cause the integrated particulate depolarization ratio to be substantially higher than would otherwise be expected. While the integrated volume depolarization ratio, δ_v, is a direct measurement, the integrated 532 nm particulate depolarization ratio, δ_p, is a post-extinction quantity, calculated from ratio of the layer integrated perpendicular and parallel polarization components of the particulate backscatter coefficients within the layer; i.e.,

$$\delta_p = \sum_{k=top}^{base} \beta_{p,\perp}(z_n) \Big/ \sum_{k=top}^{base} \beta_{p,\parallel}(z_n) \quad (6)$$

where $\beta_{P,\perp}$ and $\beta_{P,\parallel}$ are, respectively, the perpendicular and parallel components of the 532 nm particulate backscatter coefficient.

The quality of the estimate for δ_p is determined not only by the SNR of the backscatter measurements in parallel and perpendicular channels, but also the accuracy of the range-resolved two-way transmittance estimates within the layer. The two-way transmittances due to molecules and ozone can be well characterized via the model data obtained from the GMAO. The two-way transmittances due particulates, however, are only as accurate as the CALIOP extinction retrieval. Opaque layers can be particularly prone to errors in the layer-integrated particulate depolarization ratio, as very large attenuation corrections are applied to the weak signals at the base of the layers, and on those occasions where one channel or the other becomes totally attenuated, very large, negative particulate depolarization ratio estimates can result. For layers that are not opaque, δ_p is generally reliable. However, in transparent layers, relative uncertainties in the daytime estimates typically increase by a factor of ~2 due to the larger background noise compared with the nighttime estimates.

Integrated_Particulate_Depolarization_Ratio_Uncertainty (all 5 km layer products)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Absolute uncertainties associated with the estimates of the layer-integrated particulate depolarization ratio, calculated by integrating the uncertainties for the particulate backscatter coefficients measured in the 532 nm parallel and perpendicular channels, and then applying standard techniques for error propagation in ratioed quantities. The uncertainties for the parallel and perpendicular channel particulate backscatter coefficients are derived from previously computed estimates of the extinction uncertainty. There are occasions when the extinction uncertainty calculation can become unstable and excessively large. Whenever this situation is detected, a default uncertainty of 99.99 is assigned to all remaining extinction and backscatter coefficients within a layer. If the extinction calculation anywhere within a layer found to be totally unreliable – i.e., if the extinction uncertainty is 99.99 – then the integrated particulate depolarization ratio is equally unreliable and hence the integrated particulate depolarization ratio uncertainty will also be set to 99.99.

Figure 7 shows the relative particulate depolarization uncertainty, $\Delta\delta_p / \delta_p$, as a function of cirrus optical depth for all transparent cirrus measured between December 2007 and December 2016. When compiling the statistics, cirrus were limited to the uppermost layer detected in a 5 km column and required to have centroid temperatures less than –40 °C. Each point in the figure shows the median $\Delta\delta_p / \delta_p$ for an optical depth increment of 0.03. Error bars are the median absolute distance (MAD) relative to the median. Daytime data are shown in yellow, nighttime data are shown in blue. $\Delta\delta_p / \delta_p$ statistics for opaque layers are shown in Table 7. As explained earlier, the $\Delta\delta_p / \delta_p$ relative uncertainties are higher for opaque layers than for transparent layers, but do not show an appreciable day–night difference.

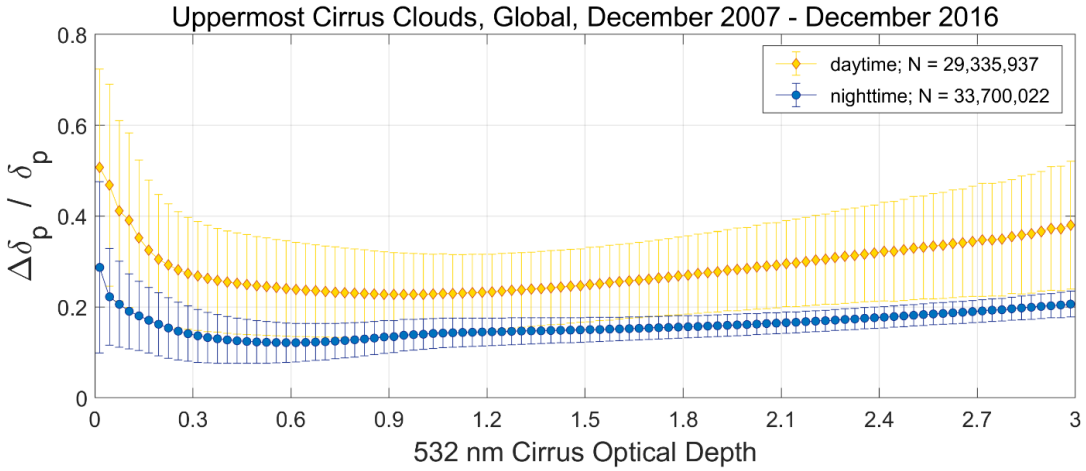


Figure 7: Relative particulate depolarization ratio uncertainty, $\Delta\delta_p / \delta_p$, for all transparent cirrus that are the uppermost layer detected in a 5 km column and have centroid temperatures less than –40 °C. The data are filtered to exclude data points with relative uncertainties greater than 2 in either the volume depolarization ratios or the particulate depolarization ratios.

Table 7: Descriptive statistics for the relative particulate depolarization ratio uncertainties for opaque cirrus that are the only layer detected in a 5 km column and have centroid temperatures less than –40 °C. The data are filtered to exclude data points with relative uncertainties greater than 2 in either the volume depolarization ratios or the particulate depolarization ratios. The mean optical depth is substantially smaller during the day because daytime solar background noise obscures faint, highly attenuated signals approaching the apparent cloud base.

	τ_{cirrus} day	$\Delta\delta_p/\delta_p$ day	τ_{cirrus} night	$\Delta\delta_p/\delta_p$ night
min	0.0082	0.0455	0.0059	0.0404
max	9.8029	2.0000	11.2573	2.0000
median	4.1081	0.5283	6.6021	0.5188
MAD	0.8028	0.2418	0.6996	0.2013
mean	4.1831	0.6133	6.4232	0.5310
stdev	1.0133	0.3170	0.9595	0.2718
samples	9,587,077	9,587,077	7,772,986	7,772,986

Particulate_Depolarization_Ratio_Statistics (all 5 km layer products)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the array of particulate depolarization ratios computed for each layer. Formulas used for each of the statistical calculations can be found in section 6 of the CALIPSO Feature Detection ATBD.

As for the volume depolarization ratios, each of the range resolved particulate depolarization ratios within any layer is the ratio of two noisy retrievals. Especially where feature backscattering is relatively weak, and in regions of low SNR, the retrieved 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. Consequently, when using particulate depolarization ratios to characterize cloud and/or aerosol layers, the layer median and (especially) the layer integrated value are both more reliable metrics than the mean.

Integrated_Particulate_Color_Ratio (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 4.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: The integrated particulate color ratio is derived from ratio of the particulate backscatter coefficients at 532 nm and 1064 nm, each summed over the vertical extent of the layer:

$$\chi_p = \sum_{k=top}^{base} \beta_{p,1064}(Z_k) \Big/ \sum_{k=top}^{base} \beta_{p,532}(Z_k) \quad . \tag{7}$$

Much like the integrated attenuated total color ratio, the quality of χ_p is governed by the accuracy to which layer top and base altitudes are determined and by the signal-to-noise ratios of the backscatter data within the layer. Additionally, since all of the $\beta_{P,\lambda}(z)$ values are derived by CALIOP's Hybrid Extinction Retrieval Algorithm (HERA; Young & Vaughan, 2009), the quality of χ_p also depends on the success of the HERA profile solver in deriving accurate solutions for $\beta_{P,\lambda}(z)$. As such, the quality of χ_p can be partially assessed via the extinction QC flags which report the final state of the HERA solution attempt. In general, solutions where the final lidar ratio is unchanged (extinction QC = 0) or the extinction solution is constrained (extinction QC = 1) yield physically plausible solutions more often. Conversely, solutions tend to be more uncertain in those cases where the lidar ratio for either wavelength must be reduced.

Integrated_Particulate_Color_Ratio_Uncertainty (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Absolute uncertainty associated with the estimate of the layer-integrated particulate backscatter color ratio, calculated by first integrating the uncertainties for the particulate backscatter coefficients measured at each wavelength, and then applying standard techniques for error propagation in ratioed quantities. The error estimates in the particulate backscatter coefficients at both wavelengths are computed by the extinction retrieval algorithm. There are occasions when the extinction uncertainty calculation can become unstable and excessively large. Whenever this situation is detected, a default

uncertainty of 99.99 is assigned to all remaining extinction and backscatter coefficients within a layer. If the backscatter uncertainty calculation fails anywhere within a layer – that is, if the uncertainty in any of the backscatter coefficients used to compute the particulate color ratio uncertainty is set to 99.99 – then the particulate color ratio uncertainty for the entire layer will also be reported as 99.99.

Particulate_Color_Ratio_Statistics (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Reports the minimum, maximum, mean, standard deviation, centroid, and skewness coefficient of the array of particulate backscatter color ratios computed for each layer. Formulas used for each of the statistical calculations can be found in section 6 of the CALIPSO Feature Detection ATBD. Because particulate color ratios are created by dividing one noisy number (the 1064 nm particulate backscatter coefficient) by a second noisy number (the 532 nm particulate backscatter coefficient), the resulting values can range from large negative values to extremely large positive values, depending on the noise in any pair of samples. When computing layer means, standard deviations, and centroids, these outliers can dominate the calculation and thus return entirely unrealistic estimates. Consequently, when using particulate color ratios to characterize cloud and/or aerosol layers, the layer median and (especially) the layer integrated value are both more reliable metrics than the mean.

Ice_Water_Path (5kmCLay, 5kmMLay, 5kmDiag)

Units: g/(m^2)

Format: Float_32

Valid Range: 0.0, 200.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: The integral, from layer top to layer base, of the ice-water content profile within any ice cloud layer. Ice water content is derived from the cloud particulate 532 nm extinction coefficient using a temperature-dependent parameterization derived from in-situ measurements (Heymsfield et al., 2014; Winker et al., 2024).

Ice_Water_Path_Uncertainty (5kmCLay, 5kmMLay, 5kmDiag)

Units: g/(m^2)

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Uncertainty associated with the estimate of ice water path, computed by integrating the uncertainties estimated for the ice water content within the layer.

Scientific Data Sets: Layer QA Information

Feature_Classification_Flags (all layer products)

Units: NoUnits

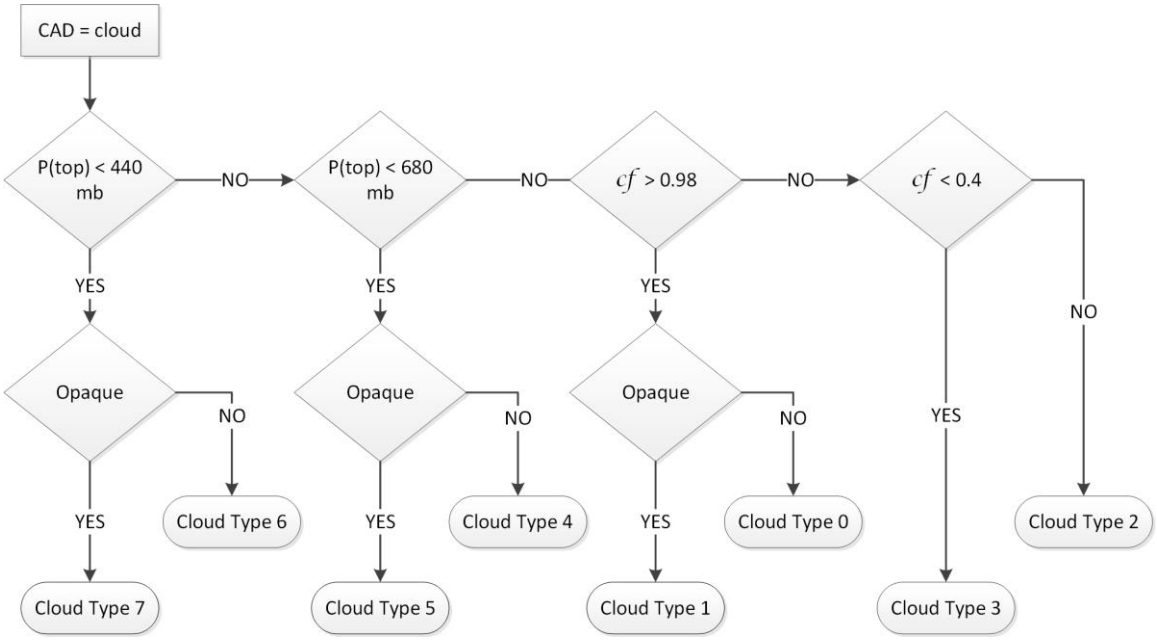
Format: UInt_16

Valid Range: 1, 49146

Description: A set of feature classification flags is derived for each layer detected in the CALIPSO backscatter data. These flags are 16-bit integers with bit interpretations given in Table 8. The first three bits of the feature classification flags (bits 0–2) identify the feature type. Bits 9–11 identify the feature subtype for clouds,

tropospheric aerosols, and stratospheric aerosols. For regions identified as "clear air", in which no atmospheric features were detected, bits 9–11 report additional LEM analyses of the signal within the range bin (Tackett et al., 2025). Note that clear air range bins are only reported in the vertical feature mask product and will not appear in the layer products. Bits 13–15 identify the amount of horizontal averaging required for feature detection.

Correct interpretation of the feature subtype bits depends entirely on the feature type identified in bits 0–2, and the interpretations are different for clouds and aerosols. Distinguishing between features and clear air is accomplished by the CALIOP layer detection algorithms (Vaughan et al., 2009). Cloud-aerosol discrimination (CAD) is accomplished using 5-dimensional probability distribution functions (PDFs) that determine layer type and assign a classification confidence score nominally between 0 (no confidence) and 100 (highest confidence) (Liu et al., 2019). The separation between tropospheric aerosols and stratospheric aerosols is based on the height of the layer attenuated backscatter centroid relative to the MERRA-2 specified tropopause latitude. Tropospheric aerosols are further partitioned into eight subtypes: undetermined, marine, dust, polluted continental/ smoke, clean continental, polluted dust, elevated smoke, and dusty marine. Six subtypes of stratospheric aerosols are identified: invalid, polar stratospheric aerosol, volcanic ash, sulfate, elevated smoke, and unclassified. Comprehensive descriptions of the tropospheric and stratospheric aerosol subtyping algorithms are given in, respectively, Kim et al., 2018 and Tackett et al., 2023. Cloud subtypes are assigned by the CALIOP cloud subtyping algorithm illustrated in Figure 8. High, middle, and low clouds are defined according to the International Satellite Cloud Climatology Project (ISCCP) cloud top pressure (P(top)) boundaries at 680 mb and 440 mb. Further discrimination is done using along track cloud fraction (cf) and cloud opacity. Additional information about the Feature_Classification_Flags is provided in the CALIPSO Vertical Feature Mask Data Description Document.



Cloud Type 0 = Low, overcast, thin (transparent stratus, stratocumulus, and fog)
Cloud Type 1 = Low, overcast, thick (opaque stratus, stratocumulus, and fog)
Cloud Type 2 = Transition stratocumulus
Cloud Type 3 = Low, broken (trade cumulus and shallow cumulus)
Cloud Type 4 = Altocumulus (transparent)
Cloud Type 5 = Altostratus (opaque, altostratus, nimbostratus, altocumulus)
Cloud Type 6 = Cirrus (transparent)
Cloud Type 7 = Deep convective (opaque altostratus, cumulonimbus, nimbostratus)

Figure 8: Diagram of the CALIOP cloud subtyping algorithm, adapted from the CALIOP scene classification ATBD. P(top) represents cloud top pressure and cf represents along track cloud fraction.

Table 8: interpretation of the bits in the Feature Classification Flags. Bit 0 is the least significant bit.

Bits	Field Description	Bit Interpretation
0–2	Feature Type	0 = rejected by LEM * 1 = "clear air" 2 = cloud 3 = tropospheric aerosol 4 = stratospheric aerosol 5 = surface 6 = subsurface 7 = no signal (totally attenuated)
3–4	Feature Type QA	0 = none (0 ≤ |CAD| < 20) 1 = low (20 ≤ |CAD| < 50) 2 = medium (50 ≤ |CAD| < 70) 3 = high (70 ≤ |CAD| ≤ 100)
5–6	Ice/Water Phase	0 = unknown / not determined 1 = ice 2 = water 3 = oriented ice crystals
7–8	Ice/Water Phase QA	0 = none 1 = low 2 = medium 3 = high
9–11	Feature Sub-type	
	If feature type = tropospheric aerosol, bits 10-12 will specify the tropospheric aerosol type	0 = not determined 1 = marine 2 = dust 3 = polluted continental/smoke 4 = clean continental 5 = polluted dust 6 = elevated smoke 7 = dusty marine
	If feature type = cloud, bits 10-12 will specify the cloud type.	0 = low overcast, transparent 1 = low overcast, opaque 2 = transition stratocumulus 3 = low, broken cumulus 4 = altocumulus (transparent) 5 = altostratus (opaque) 6 = cirrus (transparent) 7 = deep convective (opaque)

Bits	Field Description	Bit Interpretation
	If feature type = stratospheric aerosol, bits 10-12 will specify stratospheric aerosol type.	0 = invalid 1 = polar stratospheric aerosol 2 = volcanic ash 3 = sulfate 4 = elevated smoke 5 = unclassified 6 = spare 7 = spare
	If feature type = clear-air, bits 10-12 will specify horizontal averages not searched for features due to low energy mitigation	0 = n/a 1 = not searched at 80 km 2 = not searched at 20 and 80 km
12	Feature Subtype QA	0 = not confident 1 = confident
13-15	Horizontal Averaging Required for Detection	0 = not applicable 1 = 1/3 km 2 = 1 km 3 = 5 km 4 = 20 km 5 = 80 km

* in previous releases, a zero in this field indicated bad or missing data

CAD_Score (all layer products)

Units: NoUnits

Format: Int_8

Valid Range: -101, 110

Fill Value: -127

Flag Values:
–111 column data quality degraded due to low laser energy and rejected by LEM
–101 negative mean attenuated backscatter encountered
103 layer integrated attenuated backscatter at 532 nm is suspiciously high
104 scattered boundary layer clouds that were found to be opaque at the initial 5 km horizontal averaging resolution used by the layer detection algorithm
105 layer detected at 20 km or 80 km for which the initial estimates of measured properties have been negatively impacted by correction for attenuations by overlying layers
106 suspected “cirrus fringe” (Liu et al., 2019); layer initially classified as aerosol by the CAD algorithm and subsequently reclassified as a no-confidence randomly oriented ice cloud
107 correction to single-shot cloud clearing; fused/mother 5-km layer is a cloud by definition
108 correction to single-shot cloud clearing; severed 5-km layer is an aerosol by definition
109 correction to single-shot cloud clearing; single shot cloud is reclassified as an aerosol
110 correction to single-shot cloud clearing; layer is a cloud based on the scene classification algorithm

Description: The cloud-aerosol discrimination (CAD) score provides a numerical confidence level for the classification of layers by the CALIOP cloud-aerosol discrimination algorithm. 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. A value of 0 indicates that a feature has an equal likelihood of being a cloud and an aerosol. Additional special CAD scores are also defined based on additional information beyond that normally considered in the standard CAD algorithm. Full explanations of special scores 103 to 106 and the circumstances under which they are assigned are given in Liu et al., 2019. Special scores 107 to 110 are related to the revised boundary layer cloud clearing (BLaCC) algorithm first implemented in the V4.51 release (Tackett et al., 2022)

The standard CAD algorithm uses five-dimensional (5D) PDFs consisting of layer midpoint altitude, footprint center latitude, 532 nm layer mean attenuated backscatter, layer mean total attenuated backscatter color ratio, and layer mean volume depolarization ratio. Detailed descriptions of the original CAD algorithm can be found in Sections 4 and 5 of the CALIPSO Scene Classification ATBD. As detailed in the V4 data quality summary, major changes to the CAD algorithm were made in the transition from V3 to V4. Further information on the V5 CAD algorithm architecture can be found in Liu et al., 2019 and references cited therein.

The same CAD algorithm is applied to all features detected, including those the stratosphere and those detected at single shot resolution (333 m). However, the users should note that the optical properties of these single shot layers were not used for building the CAD PDFs, which may affect the overall CAD performance including the CAD scores for these single shot layers. For scientific analysis of Polar Stratospheric Clouds (PSCs), users are urged to use CALIPSO's dedicated PSC product.

Initial_CAD_Score (333MLay, 01kmCLay, 5kmCLay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: Int_8

Valid Range: -101, 119

Fill Value: -127

Flag Values: –111 data quality degraded due to low laser energy and rejected by LEM
119 feature initially identified as cloud has been reclassified

Description: The initial CAD score field contains the CAD score assigned in the initial classification of layer type, prior to attenuation corrections made for overlying layers. Because dense smoke and pollution plumes can attenuate the backscattered signal much more strongly at 532 nm than at 1064 nm, the color ratios in underlying clouds can grow to magnitudes far beyond the characteristic values recognized by the CAD algorithm. Consequently, where excessively large color ratios are detected, a CAD postprocessor will attempt to reclassify suspicious layers by using a default color ratio (see section 3.2.2 in Liu et al., 2019). When this occurs, the original layer CAD score is stored in the Initial_CAD_score field.

Extinction_QC_Flag_532 (all 5km layer products)

Extinction_QC_Flag_1064 (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: UInt_16

Valid Range: 0, 32768

Fill Value: 32768

Description: The 532 and 1064 nm extinction quality control flags are bit-mapped 16-bit integers, reported for each layer for which an extinction retrieval was attempted. Aerosol extinction is computed and reported for both wavelengths; cloud extinction is only reported at 532 nm. 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). Complete information about the conditions under which each extinction QC bit is toggled can be found in Young et al., 2018. Bit 13 was introduced after V4.51.

Table 9: interpretation of 532nm and 1064nm Extinction QC flags. Bit 0 is the least significant bit.

Bit	Value	Interpretation
0	0	unconstrained retrieval; initial lidar ratio unchanged during solution process
	1	constrained retrieval
1	2	Initial lidar ratio reduced to achieve successful full-column solutions
2	4	Suspicious retrieval due to layer or overlying integrated attenuated backscatter being too high or excessive lidar ratio reductions
3	8	Initial lidar ratio has been reduced and has converged, but backscatter uncertainty solution does not exist
4	16	Layer being analyzed has been identified by the feature finder as being totally attenuating (i.e., opaque)
5	32	Estimated optical depth error exceeds the maximum allowable value
6	64	Negative signal anomaly detected
7	128	Retrieval terminated at maximum iterations for a constrained retrieval
8	256	No solution possible within allowable lidar ratio bounds
9	512	Two-way particulate transmittance has converged but constrained retrieval still not achieved
10	1024	Backscatter coefficients not converging and maximum lidar ratio correction iterations reached
11	2048	Uncertainties not converging and maximum lidar ratio correction iterations achieved
12	4096	Lidar ratio converged but retrieval still not converging
13	8192	Range bin contains two different feature types that may have different lidar ratios
14	16384	Complex retrieval failure
15	32768	Fill value or no solution attempted

Layer_Type (5kmMLay, 5kmDiag)

Units: NoUnits

Format: Int_8

Valid Range: 2, 4

Fill Value: -127

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Feature type deciphered from bits 1-3 of the Feature Classification Flag; 2 = Cloud, 3 = Tropospheric Aerosol, 4 = Stratospheric Aerosol.

Layer_Base_Extended (all 5 km layer products)

Units: NoUnits

Format: UInt_16

Valid Range: 0, 49146

Description: A non-zero value indicates that the base of an aerosol layer has been extended from the initial altitude assigned by the layer detection algorithm to a new, lower altitude lying three range bins (90 m) above the Earth's surface (Vaughan et al., 2010). Non-zero values represent the layer's feature classification flags prior to the base being extended. Layer base extension is not applied to clouds.

Smoke_Layer_Base_Extended_QC_Flag (5kmALay, 5kmMLay, 5kmDiag)

Units: NoUnits

Format: UInt_16

Valid Range: 0, 127

Fill Value: 9999

Description: Flag used to identify aerosol layers which are candidates for smoke layer base extension, and, if selected, when the extension was successful. The CALIOP feature detection algorithm can at times have difficulty determining the base of optically thick smoke layers. Rajapakshe et al., 2017 documented one such case where a dense smoke layer was in contact with an underlying marine stratocumulus cloud deck, as evidenced by the CALIOP 1064nm attenuated backscatter signal. However, the smoke layer base altitude was detected well above the cloud top, leaving an apparent region of 'clear air' between the smoke and the cloud. This occurred because CALIOP feature detection operates on the 532 nm attenuated backscatter signal, and smoke layers are strongly attenuating at that wavelength. The attenuation of the 532nm signal through the dense smoke layer lowered the attenuated backscatter magnitude rapidly enough to fall below the feature detection threshold prematurely. The resulting base altitude for the smoke layer was set artificially high, leaving the region of clear air between the two layers. In order to improve the ability to detect the full vertical extent of smoke layers overlying clouds, a technique was developed to close this apparent clear air 'gap'. Relying on a similar approach to the CALIOP layer base extension algorithm, the base of the dense smoke layers above clouds are extended downward to 60 m above the cloud top if the following criteria are met, based on empirical analysis of CALIOP smoke observations:

- The subtype of the layer is smoke, and it is the lowest aerosol layer in the column.
- The gap between the smoke layer and the top of the underlying cloud is less than 4.5 km for day and less than 3 km night.
- The 1064 nm scattering ratio of the apparent clear air 'gap' between the smoke and cloud layer is greater than 2.0, indicating aerosol exists within the gap.
- The underlying cloud layer must have an attenuated color ratio greater than 1.4, indicating the presence of smoke overhead.

Tackett et al., 2021 provide additional details.

Table 10: interpretation of the Smoke Layer Base Extended QC Flag. Bit 0 is the least significant bit.

Bit	Interpretation
0	Candidate for smoke layer extension
1	Success, smoke layer extended
2	Failed, smoke layer base already in contact with a cloud layer
3	Failed, no cloud layer is found below the smoke layer
4	Failed, gap distance between smoke ad cloud layer exceeds limit
5	Failed, 1064 nm scattering ratio test is below threshold
6	Failed, cloud color ratio test is below threshold

Layer_IAB_QA_Factor (all layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 1.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Single layer analog of the Column Integrated Attenuated Backscatter Cumulative Probability SDS; the layer IAB QA factor is defined as 1 - $F(\gamma'_{aboveLayer})$, where $\gamma'_{aboveLayer}$ is calculated between the uppermost range bin and the range bin immediately above layer top and $F(\gamma'_{aboveLayer})$ is the cumulative probability of measuring a total column integrated attenuated backscatter equal to $\gamma'_{aboveLayer}$. Values range between 0 and 1, with lower values indicating more turbid skies above and hence reduced confidence in layer

detection boundaries and optical property retrievals. For layers detected at single shot and 1 km resolutions, the layer IAB QA factor offers qualitative insight into the overlying particulate optical depth.

Low_Energy_Mitigation_Column_QC_Flag (all layer products)

Units: NoUnits

Format: UInt_16

Valid Range: single shot = 0–1983

1 km and 5 km average = 0–63

Description: In mid-2016, the CALIOP laser began intermittently emitting low energy laser pulses, and the occurrence frequency of these pulses slowly increased as the mission progressed. For the version 5.00 data release, CALIOP's low energy mitigation algorithm (LEM; Tackett et al., 2025) was enhanced for use in the level 2 data processing. The LEM algorithm identifies level 1 backscatter data resulting from low energy pulses and, when appropriate, excludes it from the level 2 analyses. The low energy mitigation column quality control (QC) flags are N × 1 arrays of bit-mapped 16-bit integers, where N is the number of columns reported in the layer product file. Bit interpretation for the LEM column QC flags is given in Table 11. Bits 0–5 summarize the operation of the LEM algorithm within each 5 km vertical column (i.e., each frame) in the CALIOP profile products and report potential data averaging biases over an 80 km chunk. Bits 7–10 report information specific to single shot retrievals and thus will be toggled on only in the 333 m layer products and Vgroups.

Table 11: interpretation of the bits in the LEM column QC flags reported in the 1 km and 5 km layer products. Bit 0 is the least significant bit. These flags are intended to characterize the influence of individual low energy pulses on profile data that has been averaged horizontally aboard the spacecraft prior to being transmitted to Earth.

bit	Interpretation
0	Column contains LEM affected data (data has been rejected or low energy shots that LEM accepts)
1	Column belongs to a 5 km resolution frame that has been rejected due to too many unusable profiles
2	Column belongs to a 5 km resolution frame that has been rejected due to too many rejected subregions in altitude region 3 (8.2 – 20.2 km)
3	Column belongs to a 5 km resolution frame that has been rejected due to too many rejected subregions in altitude region 4 (20.2 – 30.1 km)
4	Feature detection at 20 km resolution not performed in this column due to insufficient fraction of usable 5 km resolution frames
5	Feature detection at 80 km resolution not performed in this column due to insufficient fraction of usable 5 km resolution frames
6	unused
7	Column has data rejected in altitude regions 1 & 2 (-0.2 – 8.2 km) (333 m only)
8	Column has data rejected in altitude region 3 (8.2 – 20.2 km) (333 m only)
9	Column has data rejected in altitude region 4 (20.2 – 30.1 km) (333 m only)
10	Column does not have low energy, but data is rejected in altitude regions 1 & 2 due to rejected data in altitude region 3 (333 m only)
11-15	Unused

Low_Energy_Mitigation_Feature_QC_Flag (01kmCLay, 5kmALay, 5kmCLay, 5kmMLay)

Units: NoUnits

Format: Int_8

Valid Range: 0, 1

Fill Value: -127

Description: Low energy mitigation (LEM) quality control flag reported for each layer detected.

Table 12: interpretation of the Low Energy Mitigation Feature QC flag

value	interpretation
0	No low energy shots affecting the layer
1	Some portion of the layer contains low energy shots and/or data rejected by the LEM algorithm

Low_Energy_Mitigation_Internal_Column_QC_Flag (5kmDiag)

Units: NoUnits

Format: UInt_16

Valid Range: 0, 768

Fill Value: -127

Description: Internal low energy mitigation quality control flag set during LEM algorithm operations.

Table 13: interpretation of the Low Energy Mitigation Internal Column QC flag

bit	Interpretation
0	Single shot columns have data rejected in regions 1 & 2 but the 5 km resolution column was not rejected by LEM
1	Column has data rejected in region 3 but the 5 km resolution column was not rejected by LEM
2	Column has data rejected in region 4 but the 5 km resolution column was not rejected by LEM
3	Full column is rejected by LEM
4	Column rejected due to region 3
5	Frame rejected due to too many low energy shots
6	Frame rejected due to too many rejected subregions in region 3
7	Frame rejected due to too many rejected subregions in region 4
8	Feature detection was not attempted at 20 km resolution due to too many LEM-rejected columns
9	Feature detection was not attempted at 80 km resolution due to too many LEM-rejected columns
10	Feature detection was not attempted at resolution 6 due to too many LEM-rejected columns
11	Column contains LEM affected data in region 3
12	Column contains LEM affected data in region 4
13-15	unused

Minimum_Laser_Energy_532 (1kmCLay, all 5 km layer products)

Units: J

Format: Float_32

Typical Range: 0.003…0.135

Fill Value: -9999.0

Description: A quality assurance parameter introduced in V4.20 that reports the minimum 532 nm laser energy measured within each consecutive 80 km along-track data segment (80 km = 240 single-shot laser pulses). CALIOP's standard level 2 data analyses sequentially process level 1b data in 80 km "chunks". Within these chunks, layers can be detected at horizontal resolutions as large as 80 km, and anomalously low laser energies contaminating coarse resolution upper layers can introduce biases in the spatial and optical property retrievals of underlying layers detected at finer spatial resolutions. In V5.00, the 'LEM Column QC Flag' SDS largely obviates the need for the minimum laser energy parameter.

Scientific Data Sets: Reflectance

Parallel_Column_Reflectance_532 (all layer products)

Perpendicular_Column_Reflectance_532 (all layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 2.0

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Bi-directional column reflectances derived from the depolarization gain ratio and, respectively, the parallel and perpendicular root mean square (RMS) baseline measurements at 532 nm. As explained in the data quality summary, the calculation of column reflectances was substantially revised in the V5.00 release.

Parallel_Column_Reflectance_Uncertainty_532 (all layer products)

Perpendicular_Column_Reflectance_Uncertainty_532 (all layer products)

Units: NoUnits

Format: Float_32

Valid Range: 0.0, 0.45

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

Description: Column reflectance uncertainties derived for, respectively, the 532 nm parallel channel and the 532 nm perpendicular channel. As explained in the data quality summary, the calculation of column reflectance uncertainties was substantially revised in the V5.00 release.

Parallel_Column_Reflectance_RMS_Variation_532 (all 5 km layer products)

Perpendicular_Column_Reflectance_RMS_Variation_532 (all 5 km layer products)

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 data quality degraded due to low laser energy and rejected by LEM

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

Scientific Data Sets: Ancillary Meteorological Data

Tropopause_Height (all layer products)

Units: km

Format: Float_32

Valid Range: 4.0, 22.0

Fill Value: -9999.0

Description: Tropopause heights AMSL obtained from the MERRA-2 ancillary meteorological data.

Tropopause_Temperature (all layer products)

Units: °C

Format: Float_32

Valid Range: -95.0, -20.0

Description: Tropopause temperatures, obtained from the MERRA-2 ancillary meteorological data.

Surface_Wind_Speeds_02m (5kmALay, 5kmMLay, 5kmDiag)

Units: m/s

Format: Float_32

Valid Range: -80.0, 80.0

Fill Value: -9999.0

Description: Surface wind speeds measured 2 m above the Earth's surface obtained from MERRA-2 ancillary meteorological data. First Dimension = eastward, zonal, u; Second Dimension = northward, meridional, v.

Scientific Data Sets: Surface Information

IGBP_Surface_Type (all layer products)

Units: NoUnits

Format: Int_8

Valid Range: 1, 18

Fill Value: -9

Description: International Geosphere–Biosphere Programme (IGBP) surface type at the laser footprint, provided by the CERES Surface Type ID map.

Table 14: interpretation of the IGBP surface type values.

Value	Surface Type		Value	Surface Type
1	Evergreen-Needleleaf-Forest		10	Grassland
2	Evergreen-Broadleaf-Forest		11	Wetland
3	Deciduous-Needleleaf-Forest		12	Cropland
4	Deciduous-Broadleaf-Forest		13	Urban
5	Mixed-Forest		14	Crop-Mosaic
6	Closed-Shrubland		15	Permanent-Snow
7	Open-Shrubland (Desert)		16	Barren/Desert
8	Woody-Savanna		17	Water
9	Savanna		18	Tundra

Snow_Ice_Surface_Type (all layer products)

Units: NoUnits

Format: UInt_8

Valid Range: 0, 255

Fill Value: -9

Description: Snow and ice coverage for the surface at the laser footprint provided by the U.S. Air Force Weather Agency (AFWA). Values between 0 and 100 represent the percentage of ice within a pixel having a nominal resolution of 25 km^2. Pixels with permanent ice are assigned a value of 101. A value of 103 indicates pixels containing snow. 255 indicates mixed type/uncertain type pixels at coastlines.

DEM_Surface_Elevation (all layer products)

Units: km

Format: Float_32

Valid Range: -1.0, 9.0

Fill Value: -9999.0

Description: Reports the minimum, maximum, mean, and standard deviation of the surface elevation at the laser footprint, computed for all 15 laser pulses in a 5 km along-track averages. Surface elevations are expressed as "above local mean sea level" (AMSL) and are obtained from the CloudSat science team digital elevation model (DEM).

Scientific Data Sets: Lidar Operating Mode

Laser_Energy_532 (333mMLay, all 5 km layer products)

Units: J

Format: Float_32

Typical Range: 0.003...0.135

Description: 532 nm laser energy measured for each laser pulse by a dedicated onboard laser energy monitor. In the 5 km layer products, per pulse laser energies are reported in the 'Single Shot Detection' Vgroups.

Off_Nadir_Angle (all layer products)

Units: degrees

Format: Float_32

Valid Range: 0.0, 10.0

Fill Value: -9999.0

Description: Angle of the lidar viewing vector relative to nadir pointing; the off-nadir angle was fixed at 0.3 degrees at launch/first-light and permanently changed to 3.0 degrees after November 28, 2007. The 333 m layer products report the off-nadir angles at single shot resolution. The 1 km and 5 km products report values averaged over, respectively, 3 and 15 consecutive laser pulses.

There is some jitter in the off-nadir angle. The mean value for the 51,719,250 good laser profiles acquired during October 2014 is 2.9827° ± 0.0128°, with minimum and maximum values of, respectively, 2.9348° and 3.0274°. There is also a very small number of granules in which pointing angle experiments are conducted, and during these events the off-nadir angles can be as large as 30° or more (see Appendix 1 of the CALIPSO lidar level 1 data description document). These granules with large off-nadir angles are excluded from all level 2 analyses. On the other hand, data with off-nadir angles of 1.0°, 1.5°, and 2.0°, acquired during February and June of 2017, are included in the level 2 analyses. The time periods for these small off-nadir angle experiments are reported in Table 1 of Avery et al., 2020.

Scientific Data Sets: Orbital Environment

Solar_Zenith_Angle (all layer products)

Units: degrees

Format: Float_32

Valid Range: 0.0, 180.0

Fill Value: -9999.0

Description: Angle between the zenith vector at the lidar footprint on the surface and the line of sight to the sun.

Solar_Azimuth_Angle (all layer products)

Units: degrees

Format: Float_32

Valid Range: -180.0, 180.0

Fill Value: -9999.0

Description: Azimuth angle from north of the line of sight to the sun.

Scattering_Angle (all layer products)

Units: degrees

Format: degrees

Valid Range: 0.0, 180.0

Fill Value: -9999.0

Description: Angle between the lidar viewing angle and the line of sight of the sun.

Scientific Data Sets: Spacecraft Position

Spacecraft_Position (all layer products)

Units: km

Format: Float_64

Valid Range: -8000.0, 8000.0

Fill Value: -9999.0

Description: Position of the CALIPSO satellite expressed in Earth Centered Rotating (ECR) coordinates (also known as Earth-centered, Earth-fixed (ECEF) coordinates).

Vgroups : Single Shot Lidar Surface Detection (333mLay, all 5 km layer products)

The Lidar_Surface_Detection Vgroups report the results of CALIPSO's dedicated search for backscatter signals from the Earth's surface. These "surface returns" have both a top altitude, indicating the onset of the surface signal, and a base altitude, indicating that the signal has been completely attenuated. Both altitudes are reported in kilometers above local mean sea level (AMSL). The apparent depth of the surface return is due to low-pass Bessel filters in the CALIOP receiver and, at 532 nm, the non-ideal response of the photomultipliers. The surface top and base altitudes can be different at the two wavelengths, since the 532 nm signals are downlinked at 30 m vertical resolution whereas the 1064 nm signals are downlinked at 60 m resolution. When no surface is detected for a given wavelength, all surface detection data fields report fill values.

Search results for lidar surfaces detected at single shot resolution are reported in the following layer products.

- In the Lidar_Surface_Detection Vgroup in the 333 m merged layer product
- In the Lidar_Surface_Detection Vgroup contained within the Single_Shot_Detection Vgroup in

- the 5 km aerosol layer products;
- the 5 km cloud layer products;
- the 5 km standard merged layer products; and
- the 5 km beta merged layer products.

Names and descriptions for all SDSs in the Lidar_Surface_Detection Vgroup reported in the 333 m merged layer product are listed below. In the Single_Shot_Detection Vgroups reported in the 5 km layer products, these same SDS names are prefixed with 'ss' to signify retrievals at single shot resolution. As an example, the Surface_Top_Altitude_532 SDS reported in the 333 m merged layer product is reported as the ssSurface_Top_Altitude_532 SDS in the Single_Shot_Detection Vgroups contained in all 5 km layer products.

Surface_Top_Altitude_532

Surface_Top_Altitude_1064

Units: km

Format: Float_32

Valid range: -0.5, 8.2

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Top altitude of the surface return at the lidar footprint detected at, respectively, 532 nm and 1064 nm.

Surface_Base_Altitude_532

Surface_Base_Altitude_1064

Units: km

Format: Float_32

Valid range: -0.5, 8.2

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Apparent "base" altitude of the surface return at the lidar footprint detected at, respectively, 532 nm and 1064 nm.

Surface_Integrated_Attenuated_Backscatter_532

Surface_Integrated_Attenuated_Backscatter_1064

Units: 1/sr

Format: Float_32

Valid range: 0.0, 2.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Vertically integrated total attenuated backscatter of the surface return detected at either 532 nm or 1064 nm. Limits of integration are the wavelength-specific surface top and base altitudes. The integrated attenuated backscatter measurements quantify the relative strengths of the surface returns.

Surface_532_Integrated_Depolarization_Ratio

Surface_1064_Integrated_Depolarization_Ratio

Units: NoUnits

Format: Float_32

Valid range: 0.0, 1.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: The depolarization ratio of surface return at 532 nm, computed as the ratio of the vertically integrated 532 nm perpendicular channel attenuated backscatter to the vertically integrated 532 nm parallel channel attenuated backscatter. Integration limits for the 532 nm integrated depolarization ratio are from the 532 nm surface top altitude to the 532 nm surface base altitude. For the 1064 nm integrated depolarization ratio, integration limits are from the 1064 nm surface top altitude to the 1064 nm surface base altitude. Surface depolarization ratios provide some insight into surface type, with liquid water surfaces having very low depolarization ratios, snow and ice having very high depolarization ratios, and land surfaces typically having midrange values that depend on vegetation and soil conditions (see Lu et al., 2017).

Surface_532_Integrated_Attenuated_Color_Ratio

Surface_1064_Integrated_Attenuated_Color_Ratio

Units: NoUnits

Format: Float_32

Valid range: 0.0, 2.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Total attenuated backscatter color ratio of the surface return, computed by dividing the vertically integrated 1064 nm attenuated backscatter coefficients by the vertically integrated 532 nm total attenuated backscatter coefficients. Integration limits for the 532 nm integrated attenuated color ratio are from the 532 nm surface top altitude to the 532 nm surface base altitude. For the 1064 nm integrated attenuated color ratio, integration limits are from the 1064 nm surface top altitude to the 1064 nm surface base altitude.

Surface_Detection_Flags_532

Surface_Detection_Flags_1064

Units: NoUnits

Format: UInt_16

Valid range: 0, 8192

Description: Surface detection quality assurance flags that describe the wavelength-specific outcomes of the respective surface searches. Surface detection flags are stored as 16-bit integers with bit 0 being the least significant bit. Bit interpretations are given in Table 15.

Table 15: Bit interpretations for the surface detection flags reported for each wavelength for all 532 nm and 1064 nm surface detection attempts. Bit 0 is the least significant bit.

Bit(s)	Interpretation
0	Surface detected? 0 = no 1 = yes
1-2	Surface detection method 0 = derivative test 1 = multi-shot averaged data test 2 = single shot surface detection fraction test 3 = unused
3	532 nm parallel channel saturated
4	532 nm perpendicular channel saturated
5	1064 nm channel saturated
6	532 nm parallel channel negative signal anomaly

Bit(s)	Interpretation
7	532 nm perpendicular channel negative signal anomaly
8	1064 nm channel negative signal anomaly
9	Derivative method failure: Z(min gradient) < Z(max gradient)
10	Derivative method failure: vertical extent exceeds limit
11	Derivative method failure: peak signal below threshold
12	Failure: vertical separation between 532 and 1064 surface top altitudes exceeds allowable limit
13	Failure: surface detected at 1064 nm only, but overlying atmospheric total attenuated backscatter color ratio is below a threshold value. This test is disabled in the V5.00 release.
14-15	Unused

Surface_Overlying_Integrated_Attenuated_Backscatter_532

Surface_Overlying_Integrated_Attenuated_Backscatter_1064

Units: 1/sr

Format: Float_32

Valid range: 0.0, 2.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Wavelength-specific vertically integrated total attenuated *atmospheric* backscatter from the top of the profile measurement (40 km at 532 nm, 30 km at 1064 nm) to one range bin above the detected surface top altitude.

Surface_Scaled_RMS_Background_532

Surface_Scaled_RMS_Background_1064

Units: 1/(km * sr)

Format: Float_32

Valid range: 0.0, 0.05

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Background noise estimate computed from RMS baseline noise measurements between 65 and 80 km above mean sea level, rescaled to create pseudo-attenuated backscatter coefficients at, respectively, 532 nm and 1064 nm. These pseudo-attenuated backscatter coefficients are used to determine surface detection thresholds (Vaughan et al., 2016).

Surface_Peak_Signal_532

Surface_Peak_Signal_1064

Units: 1/(km * sr)

Format: Float_32

Valid range: 0.0, 3.5

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Maximum attenuated backscatter coefficient in the surface signals detected at, respectively, 532 nm and 1064 nm.

Vgroups : 1 km Lidar Surface Detection (1kmLay, 5kmDiag)

In the 1 km cloud layer product, the Lidar_Surface_Detection Vgroup reports surfaces detected in attenuated backscatter profiles averaged to 1 km (3 shot) along-track resolution. The same layer detection algorithm is used for both single shot and 1 km resolutions. The SDS names and descriptions in the Lidar_Surface_Detection Vgroup reported in the 1 km cloud product are identical to those given previously for the single shot lidar detection Vgroup. However, the 1 km resolution Lidar_Surface_Detection Vgroup contains two SDSs, listed below, that are not reported in the single shot product.

Surface_Detections_1km_532

Surface_Detections_1km_1064

Units: NoUnits

Format: Int_16

Valid range: 0, 5

Flag values: -111 = column data quality degraded due to low laser energy and rejected by LEM

Description: These fields report the number of 1 km resolution surface detections within each 5 km resolution profile. Detection results are reported separately for the 532 nm and 1064 nm channels.

Surface detection results retrieved at 1 km along-track resolution are also reported in the Lidar_Surface_Detection Vgroup nested in the 01km_Detection Vgroup in the 5 km beta merged layer product. The parameters reported in the 01km_Detection Vgroup are identical to those reported in the stand-alone 1 km cloud layer product. To distinguish 1 km results from the single shot and 5 km results reported in the 5 km beta merged layer product, the names of SDSs reporting parameters retrieved at 1 km are prefixed with '01km'. As an example, the Surface_Top_Altitude_532 SDS reported in the 1 km cloud layer product is reported as the 01kmSurface_Top_Altitude_532 SDS in the 01km_Detection Vgroup contained in all 5 km layer products.

Vgroups : 5 km Lidar Surface Detection (all 5 km layer products)

In the 5 km standard layer products, the Lidar_Surface_Detection Vgroup reports surfaces detected in attenuated backscatter profiles averaged to 5 km (15 shot) along-track resolution. The same layer detection algorithm is used for both single shot, 1 km, and 5 km resolutions. The SDS names and descriptions in the Lidar_Surface_Detection Vgroup reported in the 5 km cloud product are identical to those given previously for the single shot lidar detection Vgroup. However, the 5 km resolution Lidar_Surface_Detection Vgroup contains two SDSs, listed below, that are not reported in the single shot product.

Surface_Detections_333m_532

Surface_Detections_333m_1064

Units: NoUnits

Format: Int_16

Valid range: 0, 15

Flag values: -111 = column data quality degraded due to low laser energy and rejected by LEM

Description: These fields report the number of single shot surface detections within each 5 km resolution profile. Detection results are reported separately for the 532 nm and 1064 nm channels.

Vgroups : Ocean-Derived Column Optical Depths, Standard (all layer products)

ODCOD_Effective_Optical_Depth_532

Units: NoUnits

Format: Float_32

Valid range: 0.0, 25.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Ocean Derived Column Optical Depth (ODCOD) estimates of total column *effective* particulate optical depth derived from the magnitude of the integrated ocean surface backscatter measured at 532 nm (Ryan et al., 2024). These are effective (i.e., apparent) optical depths rather than true optical depths because no corrections are made for multiple scattering effects arising from cloud and aerosol layers that may be present within any column.

ODCOD_Effective_Optical_Depth_532_Uncertainty

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Estimated absolute uncertainties in the ODCOD effective optical depths at 532 nm; see equation 12 in Ryan et al., 2024.

ODCOD_QC_Flag_532

Units: NoUnits

Format: UInt_32

Valid range: 0, 4294967295

Description: The ODCOD quality control flags are bit-mapped 32-bit integers describing outcome of ODCOD retrievals at 532 nm. Bit interpretation is given in Table 16.

Table 16: Interpretation of the bits in the ODCOD quality control flags. The least significant bit is bit 0.

Bit(s)	Interpretation
0	The fundamental time delay was shifted from the first point
1	Lidar surface data too many data points for a valid surface spike
2	Range bin above recorded lidar surface data added
3	Range bin below recorded lidar surface data added
4	The surface point ratio indicated that the first point of the surface spike is missing
5	To find the time delay, the surface spike index range had to be shifted down
6	The retrieval inputs fall outside of confidence ranges indicating a confident retrieval
7-9	Unused
10	No Lidar surface was found
11	The surface is not ocean
12	The surface is ice
13	The wind speed is outside the valid range (0.025 to 43 m/s)
14	The fundamental time delay found was too low; insufficient data samples to conduct a second search
15	Too few data points to solve
16	The CALIOP response model curve area was too large
17	The CALIOP response model curve scale factor could not be found
18	Lidar surface saturated
19	Lidar surface has negative signal anomaly
20	Lidar surface data is not valid
21	Flags bad input/ancillary data
22	No Lidar surface was detected via derivative method at greater than single shot resolutions

Bit(s)	Interpretation
23-31	Unused

ODCOD_Surface_Wind_Speeds_10m

Units: m/s

Format: Float_32

Valid range: -43.0, 43.0

Description: Mean zonal and meridional components of the wind speeds 10 meters above the Earth's surface reported in the lidar level 1b file. Lidar L1b wind speeds are interpolated from the MERRA-2 ancillary meteorological data and reported for each laser pulse as N × 2 arrays, where the first column gives the magnitude of the eastward, zonal, u component and the second column reports the northward, meridional, v component.

ODCOD_Surface_Wind_Speed_Correction

Units: m/s

Format: Float_32

Valid range: -43.0, 43.0

Description: An additive correction derived from global analysis of AMSR wind speeds and applied during the ODCOD retrievals to the MERRA-2 wind speed magnitudes.

Vgroups : Ocean-Derived Column Optical Depths, 5 km Merged Diagnostic Beta Only

The Ocean-Derived Column Optical Depths Vgroup reported in the 5 km Merged Layer Diagnostic Beta products provides diagnostic information about the 532 nm ODCOD retrievals. It also offers preliminary estimates of the 1064 nm effective optical depths derived using exactly the same technique that is implemented at 532 nm. Because the extensive laboratory characterization of the 532 nm detector response was not replicated for the 1064 nm channel, the 1064 nm instrument response function (IRF) is approximated using the 532 nm IRF. This choice, together with the different vertical averaging done for the two channels, inevitably introduces some increased uncertainties and the possibility of unquantified biases in the 1064 nm results. Studies conducted thus far suggest that the 1064 nm effective optical depth retrievals could be biased low by ~0.07, though the underlying cause for this offset has not been determined. Furthermore, the ODCOD 1064 nm effective optical depths have not been validated and hence should be used with utmost caution.

Understanding and effective use of the diagnostic parameters requires a close reading of the ODCOD algorithm details given in Ryan et al., 2024.

ODCOD_Effective_Optical_Depth_1064

Units: NoUnits

Format: Float_32

Valid range: 0.0, 25.0

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Estimates of total column *effective* particulate optical depth derived from the magnitude of the integrated ocean surface backscatter measured at 1064 nm. These are effective (i.e., apparent) optical depths rather than true optical depths because no corrections are made for multiple scattering effects arising from cloud and aerosol layers that may be present within any column.

ODCOD_Effective_Optical_Depth_1064_Uncertainty

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Flag Values: –111.0 column data quality degraded due to low laser energy and rejected by LEM

Description: Estimated absolute uncertainties in the ODCOD effective optical depths at 1064 nm.

ODCOD_QC_Flag_1064

Units: NoUnits

Format: UInt_32

Valid range: 0, 4294967295

Description: The ODCOD quality control flags are bit-mapped 32-bit integers describing outcome of ODCOD retrievals at 1064 nm. Bit interpretation is given in Table 16.

ODCOD_Two_Way_Transmittance_532

ODCOD_Two_Way_Transmittance_1064

Units: NoUnits

Format: Float_32

Valid range: 0, 1

Fill Value: -9999.0

Description: see equation 6 in Ryan et al., 2024. The ODCOD algorithm first estimates the effective particulate two-way transmittances in each column before converting them to the corresponding optical depths. 532 nm optical depths are reported in the standard layer products. 1064 nm optical depths and the two-way transmittances at both wavelengths are only reported in the 5 km merge beta product.

ODCOD_Two_Way_Transmittance_532_Unc

ODCOD_Two_Way_Transmittance_1064_Unc

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Description: total estimated absolute uncertainty in the effective particulate two-way transmittance calculations at both wavelengths; as seen equation 11 in Ryan et al., 2024, there are multiple components that combine to yield the total uncertainty estimate.

ODCOD_Two_Way_Transmittance_532_Unc_Wind

ODCOD_Two_Way_Transmittance_1064_Unc_Wind

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Description: the wind speed component of the total two-way transmittance uncertainty calculations; see equation 17 in Ryan et al., 2024.

ODCOD_Two_Way_Transmittance_532_Unc_Area

ODCOD_Two_Way_Transmittance_1064_Unc_Area

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Description: portion of the total two-way transmittance uncertainty calculations ascribed to uncertainties in the CALIOP response model (CRM), that approximates the underlying CALIOP instrument response function (IRF); see equation 18 in Ryan et al., 2024.

ODCOD_T2_Product_Molecular_O3_532

ODCOD_T2_Product_Molecular_O3_1064

Units: NoUnits

Format: Float_32

Valid range: 0, 1

Fill Value: -9999.0

Description: magnitude of the combined molecular and ozone two-way transmittance between the lidar and the surface at each wavelength; obtained from MERRA-2 model data.

ODCOD_Two_Way_Transmittance_532_Unc_T2MolO3

ODCOD_Two_Way_Transmittance_1064_Unc_T2MolO3

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Description: portion of the total two-way transmittance uncertainty calculations that would be ascribed to uncertainties in the molecular and ozone two-way transmittances between the lidar and the surface. Although calculated values are given in the respective SDSs, Ryan et al., 2024 state that the contributions made by this term "to the overall uncertainty should be less than 0.02%. As these uncertainties are small compared to that of wind speed and area fit, they are not included in the ODCOD (total) uncertainty estimates".

ODCOD_Ocean_Surface_Retro_Reflectance_532

ODCOD_Ocean_Surface_Retro_Reflectance_1064

Units: NoUnits

Format: Float_32

Valid range: 0, 1

Fill Value: -9999.0

Description: The modeled reflectance of the ocean surface as a function of wind speed at the lidar off-nadir angle at both wavelengths; see equation 7 in Ryan et al., 2024.

ODCOD_Fundamental_Time_Delay_532

ODCOD_Fundamental_Time_Delay_1064

Units: µs

Format: Float_32

Valid range: -0.05, 0.15

Description: Elapsed time, at both 532 nm and 1064 nm, from the surface range bin midpoint to the onset of the CRM within the bin; for details, see Section 2.1 and Figure 2 in Ryan et al., 2024.

ODCOD_Scale_Factor_532

ODCOD_Scale_Factor_1064

Units: NoUnits

Format: Float_32

Valid range: 0, 5

Fill Value: -9999.0

Description: Scale factor to achieve the best fit between the measurements and the CRM; see the α parameter in Table 1 in Ryan et al., 2024.

ODCOD_First_Point_Index_532

ODCOD_First_Point_Index_1064

Units: NoUnits

Format: Float_32

Valid range: 0, 582

Fill Value: -1

Description: The index into the 583-element CALIOP altitude array designating the location of the onset of the ocean surface signal. Due to differences in CALIOP's vertical averaging of the 532 nm and 1064 nm backscatter profiles, these indexes may not be identical for the two wavelengths.

ODCOD_Surface_Points_532

ODCOD_Surface_Points_1064

Units: $km^{-1}\,sr^{-1}$

Format: Float_32

Typical range: -0.1...3.3

Fill Value: -9999.0

Description: The attenuated backscatter coefficients used for the ODCOD retrieval at each wavelength. Level 1 data array indexes for these values extend from the ODCOD_First_Point_Index_532 downward for 1 to 3 range bins (i.e., the ODCOD ocean surface return will contain at least 2 points and no more than 4 points).

ODCOD_IRF_Area_532

ODCOD_IRF_Area_1064

Units: NoUnits

Format: Float_32

Valid range: 0.0, 0.962

Fill Value: -9999.0

Description: The ODCOD algorithm approximates the CALIOP impulse response function (IRF) using the mathematically tractable CRM. The IRF area parameters are the areas calculated under the CRM and subsequently used to determine the ODCOD scale factors at both wavelengths; see equation 5 in Ryan et al., 2024.

ODCOD_IRF_Area_532_Uncertainty

ODCOD_IRF_Area_1064_Uncertainty

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Description: Estimated uncertainty in the integral of the shifted and scaled CRM (i.e., after the time delay and scale factor have been found and applied). These parameters are intermediate values used in calculating the 'ODCOD Two Way Transmittance [532 or 1064] Unc Area" and are reported for diagnostic purposes only.

ODCOD_Curve_Fit_Error_532

ODCOD_Curve_Fit_Error_1064

Units: NoUnits

Format: Float_32

Fill Value: -9999.0

Description: Estimated uncertainty of the ocean return model fit, calculated as the root mean squared difference between the surface measurements and the best fit to the downlinked CALIOP response model (DCRM); see equation 4 in Ryan et al., 2024. The curve fit error quantifies the contributions that uncertainties in the integrated surface backscatter signals make to the total effective optical uncertainty.

Vgroups : Single Shot Detection (all 5 km layer products)

All CALIOP 5 km layer products include a Single Shot Detection (SSD) Vgroup that reports a comprehensive subset of the parameters reported in the 333 m Merged Layer Product. These results are included specifically to give users the ability to identify and characterize boundary layer clouds detected at single shot resolution. The attenuated backscatter coefficients in these boundary layer clouds are excluded when creating the increasingly coarse horizontal averages used for layer detections at 5 km, 20 km, and 80 km (Vaughan et al., 2009; Tackett et al., 2022). By design, the SDS names in the SSD Vgroup are very nearly identical to the SDS names in the root HDF directory. The difference is that the SSD Vgroup names are, in all cases, prefixed with 'ss', where 'ss' represents 'single shot'. Doing this eliminates any potential ambiguity in specifying parameter names when using the 5 km layer products. For example, in the 5 km Cloud Layer Product, cloud top altitudes for layers detected at 5 km, 20 km, and 80 km are reported in the Layer_Top_Altitude SDS stored in the HDF root directory. Similarly, cloud top altitudes for layers detected at single shot resolution are reported in the ssLayer_Top_Altitude SDS stored in the SSD Vgroup. Because the spatial resolution of the ancillary meteorological data present is not better than 5 km, these parameters are not copied from the 333 m Merged Layer Product into the SSD Vgroup. Sun angles and spacecraft position are also excluded from the SSD Vgroup.

Table 17 lists all SDSs contained in the SSD Vgroup and includes links to the parameter descriptions in the left column. Because the spatial resolution of the ancillary meteorological data present is not better than 5 km, these parameters are not copied from the 333 m Merged Layer Product into the SSD Vgroup. Sun angles and spacecraft position are also excluded from the SSD Vgroup.

Table 17: correspondence between SDS names in the 333 m Merged layer product and the Single Shot Detection Vgroups reported in all 5 km layer products.

333 m Merged Layer Product	Single Shot Detection Vgroup
Lidar_Data_Altitudes	
Profile_ID (all layer products)	ssProfile_ID
Latitude (all layer products)Latitude (all layer products)	ssLatitude
Longitude (all layer products)	ssLongitude
Profile_Time (all layer products)Profile_Time (all layer products)	ssProfile_Time
Profile_UTC_Time (all layer products)	ssProfile_UTC_Time
Day_Night_Flag (all layer products)	ssDay_Night_Flag
Off_Nadir_Angle (all layer products)	ssOff_Nadir_Angle
Solar_Zenith_Angle	
Solar_Azimuth_Angle	
Scattering_Angle	
Spacecraft_Position	
Laser_Energy_532 (333mMLay, all 5 km layer products)	ssLaser_Energy_532

333 m Merged Layer Product	Single Shot Detection Vgroup
Low_Energy_Mitigation_Column_QC_Flag (all layer products)	ssLow_Energy_Mitigation_Column_QC_Flag
Parallel_Column_Reflectance_532 (all layer products)	ssParallel_Column_Reflectance_532
Parallel_Column_Reflectance_Uncertainty_532 (all layer products)	ssParallel_Column_Reflectance_Uncertainty_532
Perpendicular_Column_Reflectance_532 (all layer products)	ssPerpendicular_Column_Reflectance_532
Perpendicular_Column_Reflectance_Uncertainty_532 (all layer products)	ssPerpendicular_Column_Reflectance_Uncertainty_532
Column_Integrated_Attenuated_Backscatter_532 (all layer products)	ssColumn_Integrated_Attenuated_Backscatter_532
Column_IAB_Cumulative_Probability (all layer products)	ssDEM_Surface_Elevation
Column_Particulate_Optical_Depth_Above_Opaque_Water_Cloud_532 (all layer products)	ssColumn_Particulate_Optical_Depth_Above_Opaque_Water_Cloud_532
Column_Particulate_Optical_Depth_Above_Opaque_Water_Cloud_Uncertainty_532 (all layer products)	ssColumn_Particulate_Optical_Depth_Above_Opaque_Water_Cloud_Uncertainty_532
Tropopause_Height	
Tropopause_Temperature	
IGBP_Surface_Type	
Snow_Ice_Surface_Type	
DEM_Surface_Elevation	
Number_Layers_Found (all layer products)	ssNumber_Layers_Found
Scene_Flag (all layer products)	ssScene_Flag
Number_Bins_Shift (333mMLay, 01kmCLay)	ssNumber_Bins_Shift
Layer_Top_Altitude (all layer products)	ssLayer_Top_Altitude
Layer_Base_Altitude (all layer products)	ssLayer_Base_Altitude
Layer_Top_Pressure (all layer products)	ssLayer_Top_Pressure
Midlayer_Pressure (all layer products)	ssMidlayer_Pressure
Layer_Base_Pressure (all layer products)	ssLayer_Base_Pressure
Layer_Top_Temperature (all layer products)	ssLayer_Top_Temperature
Layer_Centroid_Temperature (333mMLay, 1kmCLay, 5kmCLay, 5kmMLay, 5kmDiag)	ssLayer_Centroid_Temperature
Midlayer_Temperature (all layer products)	ssMidlayer_Temperature
Layer_Base_Temperature (all layer products)	ssLayer_Base_Temperature
Opacity_Flag (all layer products)	ssOpacity_Flag
Attenuated_Scattering_Ratio_Statistics_532 (all layer products)	ssAttenuated_Scattering_Ratio_Statistics_532
Attenuated_Backscatter_Statistics_532 (all layer products)	ssAttenuated_Backscatter_Statistics_532
Integrated_Attenuated_Backscatter_532 (all layer products)	ssIntegrated_Attenuated_Backscatter_532

333 m Merged Layer Product	Single Shot Detection Vgroup
Integrated_Attenuated_Backscatter_Uncertainty_532 (all layer products)	ssIntegrated_Attenuated_Backscatter_Uncertainty_532
Attenuated_Backscatter_Statistics_1064 (all layer products)	ssAttenuated_Backscatter_Statistics_1064
Integrated_Attenuated_Backscatter_1064 (all layer products)	ssIntegrated_Attenuated_Backscatter_1064
Integrated_Attenuated_Backscatter_Uncertainty_1064 (all layer products)	ssIntegrated_Attenuated_Backscatter_Uncertainty_1064
Volume_Depolarization_Ratio_Statistics (all layer products)	ssVolume_Depolarization_Ratio_Statistics
Integrated_Volume_Depolarization_Ratio (all layer products)	ssIntegrated_Volume_Depolarization_Ratio
Integrated_Volume_Depolarization_Ratio_Uncertainty (all layer products)	ssIntegrated_Volume_Depolarization_Ratio_Uncertainty
Attenuated_Total_Color_Ratio_Statistics (all layer products)	ssAttenuated_Total_Color_Ratio_Statistics
Integrated_Attenuated_Total_Color_Ratio (all layer products)	ssIntegrated_Attenuated_Total_Color_Ratio
Integrated_Attenuated_Total_Color_Ratio_Uncertainty (all layer products)	ssIntegrated_Attenuated_Total_Color_Ratio_Uncertainty
Overlying_Integrated_Attenuated_Backscatter_532 (all layer products)	ssOverlying_Integrated_Attenuated_Backscatter_532
Layer_IAB_QA_Factor (all layer products)	ssLayer_IAB_QA_Factor
Feature_Classification_Flags (all layer products)	ssFeature_Classification_Flags
* in previous releases, a zero in this field indicated bad or missing data CAD_Score (all layer products)	ssCAD_Score
Initial_CAD_Score (333MLay, 01kmCLay, 5kmCLay, 5kmMLay, 5kmDiag)	ssInitial_CAD_Score
Was_Cleared (333mMLay)	ssWas_Cleared
Decoupled_Layer_Type (333mMLay, 5kmALay, 5kmCLay, 5kmMLay, 5kmDiag)	ssDecoupled_Layer_Type

Metadata Parameters

Product_ID

A character string (80-byte maximum) specifying the data product name. For all CALIPSO Level 2 lidar data products, the value of this string is “L2_LIDAR”.

Date_Time_at_Granule_Start

A 27-byte character string that specifies the granule UTC start date and time. The format is yyyy-mm-ddThh:nn:ss.ffffffZ, where yyyy is the year, mm is the month, dd is the day, hh is the hour, nn is the minute, ss is the second, and ffffff is the fractional second. Date and time are separated by the character ‘T’. The ‘Z’ indicates that time is given in UTC.

Date_Time_at_Granule_End

A 27-byte character string that specifies the granule UTC end date and time. The format is yyyy-mm-ddThh:nn:ss.ffffffZ, where yyyy is the year, mm is the month, dd is the day, hh is the hour, nn is the minute, ss is the second, and ffffff is the fractional second. Date and time are separated by the character 'T'. The 'Z' indicates that time is given in UTC.

Date_Time_of_Production

A 27-byte character string that specifies the UTC date and time at which the data was generated. The format is yyyy-mm-ddThh:nn:ss.ffffffZ, where yyyy is the year, mm is the month, dd is the day, hh is the hour, nn is the minute, ss is the second, and ffffff is the fractional second. Date and time are separated by the character 'T'. The 'Z' indicates that time is given in UTC.

Number_of_Bad_Profiles

A 32-bit integer specifying the number of bad attenuated backscatter profiles contained the level 1b file used to generate the level 2 analyses. Profiles are considered bad if (a) any of the three measurement channels are missing (see bits 0, 1, and 2 in the QC Flags SDS); (b) the measurement data could not be geolocated (see bit 3 in the QC flags SDS); or (c) the energy in either the 532 nm or 1064 nm channel falls below the low energy threshold (see bits 5 and 6 in the QC Flags SDS).

Number_of_Good_Profiles

A 32-bit integer specifying the number of good (i.e., not identified as being bad) attenuated backscatter profiles contained in the level 1b file used to generate the level 2 analyses.

Initial_Subsatellite_Latitude

This field reports the first subsatellite latitude contained in the granule.

Initial_Subsatellite_Longitude

This field reports the first subsatellite longitude contained in the granule.

Final_Subsatellite_Latitude

This field reports the last subsatellite latitude contained in the granule.

Final_Subsatellite_Longitude

This field reports the last subsatellite longitude contained in the granule.

Orbit_Number_at_Granule_Start

Orbit_Number_at_Granule_Stop

Orbit_Number_Change_Time

Orbit Number consists of three fields that define the number of revolutions by the CALIPSO spacecraft around the Earth. This number is incremented each time the spacecraft passes the equator on the ascending node. To maintain consistency between the CALIPSO and CloudSat orbit parameters, the Orbit Number is keyed to the CloudSat orbit 2121 at 23:00:47 on 2006/09/20. Because the CALIPSO data granules are organized according to satellite lighting conditions, based on fixed Sun-Earth-Satellite angles, day/night boundaries do not coincide with transition points for defining orbit number. As such, three parameters are needed to describe the orbit number for each granule as:

- **Orbit Number at Granule Start:** orbit number at the granule start time.
- **Orbit Number at Granule End:** orbit number at the granule stop time.
- **Orbit Number Change Time:** time at which the orbit number changes in the granule.

Path_Number_at_Granule_Start

Path_Number_at_Granule_Stop

Path_Number_Change_Time

Path Number consists of three fields that define an index ranging from 1-233 that references orbits to the Worldwide Reference System (WRS). This global grid system was developed to support scene identification for LandSat imagery. Since the A-Train is maintained to the WRS grid within ±10 km, the Path Number provides a convenient index to support data searches, instead of having to define complex latitude and longitude regions along the orbit track. The Path Number is incremented after the maximum latitude in the orbit is realized and changes by a value of 16 between successive orbits. Because the CALIPSO data granules are organized according to satellite lighting conditions, day/night boundaries do not coincide with transition points for defining path number. As such, three parameters are needed to describe the path number for each granule as:

- **Path Number at Granule Start:** path number at the granule start time.
- **Path Number at Granule End:** path number at the granule stop time.
- **Path Number Change Time:** time at which the path number changes in the granule.

While CALIPSO was formation flying in the A-Train all path numbers in the metadata are exact. Beginning in September 2018, when CALIPSO lowered its orbit into the C-Train, path numbers are no longer exact, but instead indicate the closest WRS reference orbit.

Lidar_L1_Production_Date_Time

For each CALIOP Lidar Level 2 data product, the Lidar Level 1 Production Date Time field reports the file creation time and date for the CALIOP Level 1 Lidar data file that provided the source data used in the Level 2 analyses.

Number_of_Single_Shot_Records_in_Files

QA parameter for internal use only

Number_of_Average_Records_in_File

QA parameter for internal use only

Number_of_Features_Found

QA parameter for internal use only

Number_of_Cloud_Features_Found

QA parameter for internal use only

Number_of_Aerosol_Features_Found

QA parameter for internal use only

Number_of_Indeterminate_Features_Found

QA parameter for internal use only

Ocean_Fresnel_Reflection_Coefficient_532

Estimated reflectance coefficient of the 532 nm signal by seawater. The value of this parameter is 0.0213, based on work documented in Vaughan et. al., 2019.

Ocean_Fresnel_Reflection_Coefficient_1064 (5kmDiag)

Estimated reflectance coefficient of the 1064 nm signal by seawater. The value of this parameter is 0.0202, based on work documented in Vaughan et. al., 2019.

MERRA2_Wind_Uncertainty

Estimated global relative uncertainty of the 10 m wind speed reported in the MERRA2 meteorological model data product. The value of this parameter is 0.1506, based on work documented in Ryan et al., 2024

AMSR_Wind_Correction_Uncertainty

Estimated global correction uncertainty between the AMSR and MERRA-2 winds. As documented in Ryan et al., 2024, the value of this parameter is 0.2537.

Lidar_Data_Altitudes

Altitudes (above mean sea level) that specify the vertical midpoints of the 583 range bins in the profile measurements downlinked from the CALIPSO satellite. The values in this field are identical to those in the Lidar_Data_Altitudes SDS and are retained in V5.00 for backward compatibility with existing software.

GEOS_Version

Specifies the version of the meteorological data used in the level 2 analyses. For the version 5.0 data release, this field is always "MERRA2".

GMAO_Files_Used

Lists the four MERRA-2 meteorological data files used to create the level 2 data files for this granule.

Classifier_Coefficients_Version_Number

Version number of the classifier coefficients file that stores the five-dimensional probability distribution functions used by the cloud-aerosol discrimination (CAD) algorithm.

Classifier_Coefficients_Version_Date

Creation date of the classifier coefficients file that stores the five-dimensional probability distribution functions used by the cloud-aerosol discrimination (CAD) algorithm.

Production_Script

Provides the configuration information and command sequences that were executed during the processing of the CALIOP Lidar Level 2 data products. Documentation for many of the control constants found within this field is contained in the CALIPSO Lidar Level 2 Algorithm Theoretical Basis Documents.

CALIPSO Data Quality Information

Relevant External Documentation

This section lists CALIPSO Algorithm Theoretical Basis Documents (ATBDs) and peer-reviewed journal articles that provide detailed descriptions of the algorithms used to calibrate the lidar and retrieve the CALIOP level 2 science data products.

ATBDs and Project Documentation

- CALIPSO Data Management Team: CALIPSO Data Products Catalog, PC-SCI-503, Release 5.00.
- Vaughan, M. A., D. M. Winker, and K. A. Powell, 2005: CALIOP Algorithm Theoretical Basis Document, Part 2: Feature Detection and Layer Properties Algorithm, PC-SCI-202.02, https://ntrs.nasa.gov/citations/20250006627.

- Liu, Z., A. H. Omar, Y. Hu, M. A. Vaughan, and D.M. Winker, 2005: CALIOP Algorithm Theoretical Basis Document, Part 3: Scene Classification Algorithms, PC-SCI-202.03, https://ntrs.nasa.gov/citations/20250006628.
- Winker, D. M., C. A. Hostetler, M. A. Vaughan, and A. H. Omar, 2006: CALIOP Algorithm Theoretical Basis Document, Part 1: CALIOP Instrument and Algorithms Overview, PC-SCI-202.01, https://ntrs.nasa.gov/citations/20250006626.

Peer-Reviewed Journal Papers

- Liu, Z., M. J. McGill, Y. Hu, C. A. Hostetler, M. Vaughan, and D. Winker, 2004: "Validating lidar depolarization calibration using solar radiation scattered by ice clouds", *IEEE Geosci. Remote Sens. Lett.*, **1**, 157–161, https://doi.org/10.1109/LGRS.2004.829613.
- Liu, Z., M. Vaughan, D. Winker, C. A. Hostetler, L. R. Poole, D. L. Hlavka, W. D. Hart, and M. J. McGill, 2004: Use of probability distribution functions for discriminating between cloud and aerosol in lidar backscatter data, *J. Geophys. Res.*, **109**, D15202, https://doi.org/10.1029/2004JD004732.
- Liu, Z., W. Hunt, M. Vaughan, C. Hostetler, M. McGill, K. Powell, D. Winker, and Y. Hu, 2006: Estimating Random Errors Due to Shot Noise in Backscatter Lidar Observations, *Appl. Opt.*, **45**, 4437–4447, https://doi.org/10.1364/AO.45.004437.
- Hu, Y., M. Vaughan, Z. Liu, K. Powell, and S. Rodier, 2007: Retrieving Optical Depths and Lidar Ratios for Transparent Layers Above Opaque Water Clouds From CALIPSO Lidar Measurements, *IEEE Geosci. Remote Sens. Lett.*, **4**, 523–526, https://doi.org/10.1109/LGRS.2007.901085.
- Hu, Y., D. Winker, M. Vaughan, B. Lin, A. Omar, C. Trepte, D. Flittner, P. Yang, W. Sun, Z. Liu, Z. Wang, S. Young, K. Stamnes, J. Huang, R. Kuehn, B. Baum, and R. Holz, 2009: CALIPSO/CALIOP Cloud Phase Discrimination Algorithm, *J. Atmos. Oceanic Technol.*, **26**, 2293–2309, https://doi.org/10.1175/2009JTECHA1280.1.
- Hunt, W. H, D. M. Winker, M. A. Vaughan, K. A. Powell, P. L. Lucker, and C. Weimer, 2009: CALIPSO Lidar Description and Performance Assessment, *J. Atmos. Oceanic Technol.*, **26**, 1214–1228, https://doi.org/10.1175/2009JTECHA1223.1.
- Liu, Z., M. A. Vaughan, D. M. Winker, C. Kittaka, R. E. Kuehn, B. J. Getzewich, C. R. Trepte, and C. A. Hostetler, 2009: The CALIPSO Lidar Cloud and Aerosol Discrimination: Version 2 Algorithm and Initial Assessment of Performance, *J. Atmos. Oceanic Technol.*, **26**, 1198–1213, https://doi.org/10.1175/2009JTECHA1229.1.
- Omar, A., D. Winker, C. Kittaka, M. Vaughan, Z. Liu, Y. Hu, C. Trepte, R. Rogers, R. Ferrare, R. Kuehn, and C. Hostetler, 2009: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *J. Atmos. Oceanic Technol.*, **26**, 1994–2014, https://doi.org/10.1175/2009JTECHA1231.1.
- Powell, K. A., C. A. Hostetler, Z. Liu, M. A. Vaughan, R. E. Kuehn, W. H. Hunt, K. Lee, C. R. Trepte, R. R. Rogers, S. A. Young, and D. M. Winker, 2009: CALIPSO Lidar Calibration Algorithms: Part I - Nighttime 532 nm Parallel Channel and 532 nm Perpendicular Channel, *J. Atmos. Oceanic Technol.*, **26**, 2015–2033, https://doi.org/10.1175/2009JTECHA1242.1.
- Vaughan, M., K. Powell, R. Kuehn, S. Young, D. Winker, C. Hostetler, W. Hunt, Z. Liu, M. McGill, and B. Getzewich, 2009: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, *J. Atmos. Oceanic Technol.*, **26**, 2034–2050, https://doi.org/10.1175/2009JTECHA1228.1.
- Winker, D. M., M. A. Vaughan, A. H. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young, 2009: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, *J. Atmos. Oceanic Technol.*, **26**, 2310–2323, https://doi.org/10.1175/2009JTECHA1281.1.
- Young, S. A. and M. A. Vaughan, 2009: The retrieval of profiles of particulate extinction from Cloud Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) data: Algorithm description, *J. Atmos. Oceanic Technol.*, **26**, 1105–1119, https://doi.org/10.1175/2008JTECHA1221.1.
- Garnier, A., J. Pelon, M. A. Vaughan, D. M. Winker, C. R. Trepte, and P. Dubuisson, 2015: Lidar multiple scattering factors inferred from CALIPSO lidar and IIR retrievals of semi-transparent cirrus cloud optical depths over oceans, *Atmos. Meas. Tech.*, **8**, 2759–2774, https://doi.org/10.5194/amt-8-2759-2015.

- Getzewich, B. J., M. A. Vaughan, W. H. Hunt, M. A. Avery, K. A. Powell, J. L. Tackett, D. M. Winker, J. Kar, K.-P. Lee, and T. Toth, 2018: CALIPSO Lidar Calibration at 532-nm: Version 4 Daytime Algorithm, *Atmos. Meas. Tech.*, **11**, 6309–6326, https://doi.org/10.5194/amt-11-6309-2018.
- Kar, J., M. A. Vaughan, K. P. Lee, J. Tackett, M. Avery, A. Garnier, B. Getzewich, W. Hunt, D. Josset, Z. Liu, P. Lucker, B. Magill, A. Omar, J. Pelon, R. Rogers, T. D. Toth, C. Trepte, J-P. Vernier, D. Winker, and S. Young, 2018: CALIPSO Lidar Calibration at 532 nm: Version 4 Nighttime Algorithm, *Atmos. Meas. Tech.*, **11**, 1459–1479, https://doi.org/10.5194/amt-11-1459-2018.
- Kim, M.-H., A. H. Omar, J. L. Tackett, M. A. Vaughan, D. M. Winker, C. R. Trepte, Y. Hu, Z. Liu, L. R. Poole, M. C. Pitts, J. Kar, and B. E. Magill, 2018: The CALIPSO Version 4 Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *Atmos. Meas. Tech.*, **11**, 6107-6135, https://doi.org/10.5194/amt-11-6107-2018.
- Lu, X., Y. Hu, Y. Yang, M. Vaughan, Z. Liu, S. Rodier, W. Hunt, K. Powell, P. Lucker, and C. Trepte, 2018: Laser pulse bidirectional reflectance from CALIPSO mission, *Atmos. Meas. Tech.*, **11**, 3281–3296, https://doi.org/10.5194/amt-11-3281-2018.
- Young, S. A., M. A. Vaughan, J. L. Tackett, A. Garnier, J. B. Lambeth, and K. A. Powell, 2018: Extinction and Optical Depth Retrievals for CALIPSO's Version 4 Data Release, *Atmos. Meas. Tech.*, **11**, 5701–5727, https://doi.org/10.5194/amt-11-5701-2018.

  Supplement: https://doi.org/10.5194/amt-11-5701-2018-supplement
- Liu, Z., J. Kar, S. Zeng, J. Tackett, M. Vaughan, M. Avery, J. Pelon, B. Getzewich, K.-P. Lee, B. Magill, A. Omar, P. Lucker, C. Trepte, and D. Winker, 2019: Discriminating Between Clouds and Aerosols in the CALIOP Version 4.1 Data Products, *Atmos. Meas. Tech.*, **12**, 703–734, https://doi.org/10.5194/amt-12-703-2019.
- Vaughan, M., A. Garnier, D. Josset, M. Avery, K.-P. Lee, Z. Liu, W. Hunt, J. Pelon, Y. Hu, S. Burton, J. Hair, J. Tackett, B. Getzewich, J. Kar, and S. Rodier, 2019: CALIPSO Lidar Calibration at 1064 nm: Version 4 Algorithm, *Atmos. Meas. Tech.*, **12**, 51–82, https://doi.org/10.5194/amt-12-51-2019.
- Avery, M. A., R. A. Ryan, B. J. Getzewich, M. A. Vaughan, D. M. Winker, Y. Hu, A. Garnier, J. Pelon, and C. A. Verhappen, 2020: "CALIOP V4 Cloud Thermodynamic Phase Assignment and the Impact of Near-Nadir Viewing Angles", *Atmos. Meas. Tech.*, **13**, 4539–4563, https://doi.org/10.5194/amt-13-4539-2020.
- Ryan, R. A., M. A. Vaughan, S. D. Rodier, J. L. Tackett, J. A. Reagan, R. A. Ferrare, J. W. Hair, and B. J. Getzewich, 2024: "Total Column Optical Depths Retrieved from CALIPSO Lidar Ocean Surface Backscatter", *Atmos. Meas. Tech.*, **17**, 6517–6545, https://doi.org/10.5194/amt-17-6517-2024.
- Tackett, J. L., J. Kar, M. A. Vaughan, B. Getzewich, M.-H. Kim, J.-P. Vernier, A. H. Omar, B. Magill, M. C. Pitts, and D. Winker, 2023: "The CALIPSO version 4.5 stratospheric aerosol subtyping algorithm", *Atmos. Meas. Tech.*, **16**, 745–768, https://doi.org/10.5194/amt-16-745-2023.
- Tackett, J. L., R. A. Ryan, A. E. Garnier, J. Kar, B. Getzewich, X. Cai, M. A. Vaughan, C. R. Trepte, R. Verhappen, D. M. Winker and K.-P. A. Lee, 2025: Mitigating Impacts of Low Energy Laser Pulses on CALIOP Data Products, *EGUsphere* [preprint], https://doi.org/10.5194/egusphere-2025-2376.
- Toth, T. D., G. Schuster, M. Clayton, Z. Li, D. Painemal, S. Rodier, J. Kar, T. Thorsen, R. Ferrare, M. Vaughan, J. Tackett, H. Bian, M. Chin, A. Garnier, E. Welton, R. Ryan, C. Trepte and D. Winker, 2025: "Mapping CALIPSO Marine and Dusty Marine Aerosol Lidar Ratios using MODIS AOD Constrained Retrievals and GOCART Model Simulations" *EGUsphere* [AMTD], https://doi.org/10.5194/egusphere-2025-2832.

Data Release Information

At the conclusion of the mission, the CALIPSO project had released five major versions of the lidar level 2 data products, as well as several minor version updates. Table 18 lists all major and minor releases.

Table 18: release date, version number, data date range, and production strategy for all CALIPSO lidar level 2 data products

Lidar Level 2: Half orbits (Night and Day)			
Release Date	**Version**	**Data Date Range**	**Production Strategy**
October 2025	5.00	June 13, 2006 to June 30, 2023	Standard
June 2023	4.51	June 13, 2006 to June 30, 2023	Standard
October 2020	4.21	July 01, 2020 to January 19, 2022	Standard
October 2018	4.20	June 13, 2006 to June 20, 2020	Standard
November 2016	4.10	June 13, 2006 to September 30, 2020	Standard
October 2020	3.41	October 1, 2020 to June 30, 2023	Validated Stage 1
December 2016	3.40	December 1, 2016 to September 31, 2020	Validated Stage 1
April 2013	3.30	March 1, 2013 to November 30, 2016	Validated Stage 1
December 2011	3.02	November 1, 2011 to February 28, 2013	Validated Stage 1
May 2010	3.01	June 13, 2006 to October 31, 2011	Validated Stage 1
October 2008	2.02	September 14, 2008 to October 29, 2009	Provisional
January 2008	2.01	June 13, 2006 to September 13, 2008	Provisional
December 2006	1.10	June 13, 2006 to November 11, 2007	Provisional / Beta

Data Quality Summaries

Data Quality Statement for CALIPSO's Version 5.00 Lidar Level 2 Data Product Release

Data Version: **5.00**
Data Release Date: **1 October 2025**
Data Date Range: **June 13, 2006 to June 30, 2023**

The sections below highlight the most significant changes made in the CALIOP version 5.00 (V5.00) level 2 data products. The magnitude and effects of these changes are frequently illustrated with comparisons to the previous release of the version 4.51 (V4.51) data products.

Modifications for Low Energy Mitigation (LEM)

During final seven years of the CALIPSO mission, a slow leak in the laser canister (Hunt et al., 2009) reduced the internal pressure below the voltage breakdown limit described by Pachen's law. Once this occurred, intermittent coronal arcing across the Q switch caused the CALIOP laser to begin emitting an increasing number of low-to-no energy laser pulses. As seen in Figure 9, early on this behavior was confined almost exclusively to the South Atlantic Anomaly (SAA; Rodriguez et al., 2022) but toward the end of the mission the phenomenon occurred worldwide.

Because CALIOP profiles are time-averaged onboard the satellite prior to downlink (Hunt et al., 2009), the deleterious effects of individual low energy pulses are not localized, but instead cause a degradation in signal quality across multiple consecutive shots. To minimize these effects, the CALIPSO lidar science working group (LSWG) developed a family of low energy mitigation (LEM) algorithms (Tackett et al., 2025) that identify compromised level 1b data segments on small, targeted scales. Using a highly optimized data filtering scheme, these segments are subsequently excluded from all level 2 data analyses. The level 1 LEM algorithm corrects a pervasive low daytime calibration bias (3% to 4%) and reduces daytime calibration uncertainties by 20% to 40% in the SAA latitude band. In addition, the V5.00 level 1 energy normalization process now correctly compensates for the inclusion of no energy laser pulses in data averaged onboard the satellite. As demonstrated by Figure 10, this

level 1 modification, in combination with the level 2 LEM algorithms, yields a substantial reduction in the number of false positive feature detections.

Note that averaged profiles containing one or more low energy pulses are not automatically excluded in the level 2 analyses. Horizontally averaged data segments that are "LEM-affected" (i.e., some low energy shots are present, but the filtered profile is still deemed acceptable for science retrievals) experience a signal-to-noise ratio reduction of 6–9 % which slightly increases the probability of false detections relative to unaffected data. However, as illustrated by Figure 11, the median measured optical properties of the LEM-affected layers typically differ from the properties of unaffected layers by ~1.0% or less.

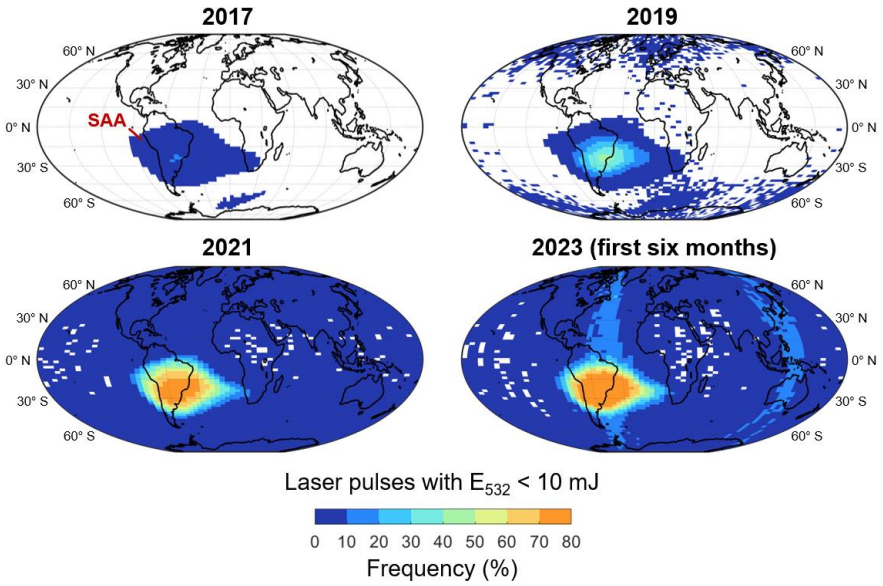


Figure 9: Global frequencies of laser pulses having 532 nm energies less than 10 mJ during 2017, 2019, 2021, and the first six months of 2023. The vertical bands of high frequencies seen in 2023 are caused by daytime granules having nearly continuous low energy pulses during May and June; e.g., see 2023-06-12T05-54-23ZD (Tackett et al., 2025 © Authors 2025, CC BY 4.0).

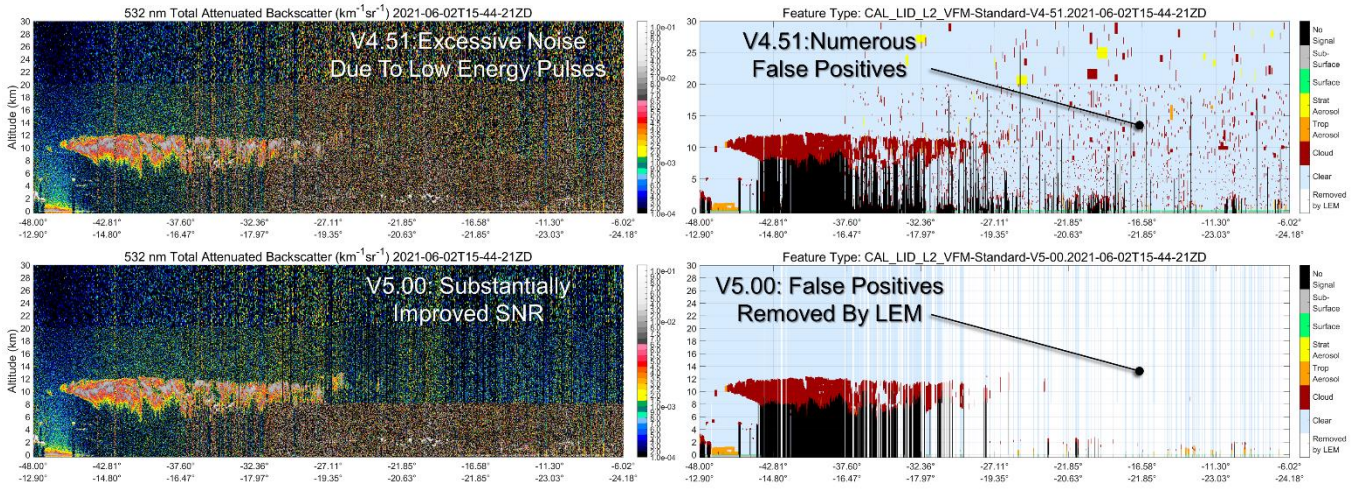


Figure 10: Top row shows the level 1 532 nm total attenuated backscatter (β'(z), left) and the corresponding level 2 vertical feature mask (VFM, right) from the V4.51 release for a granule strongly affected by low energy pulses on 02 June 2021 at 16Z. Similarly, the bottom row shows the same quantities for the V5.00 release. Relative to V4.51, the SNR above 8.2 km is noticeably higher in the LEM-filtered V5.00 532 nm β'(z) example. Similarly, the V5.00 VFM arguably eliminates all of the false positive feature detections seen in V4.51 (Figure adapted from Tackett et al., 2025 © Authors 2025, CC BY 4.0).

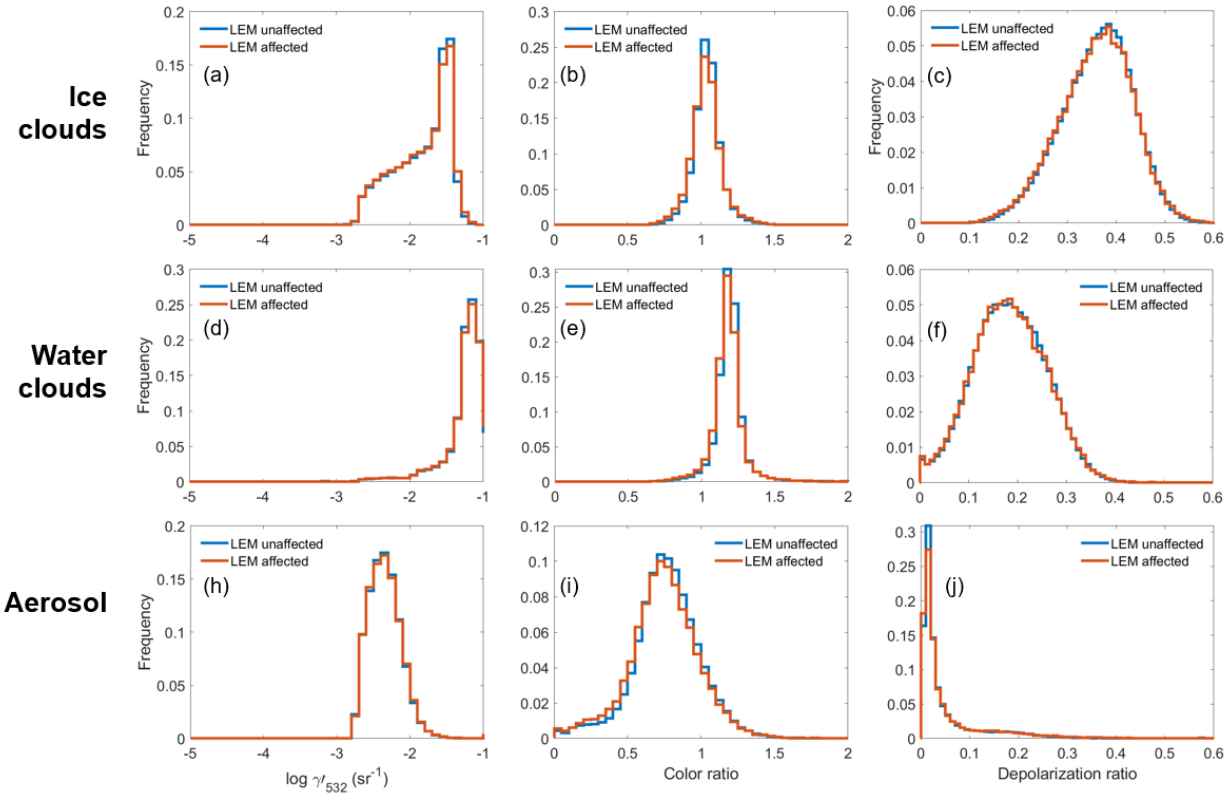


Figure 11: Frequency distributions of layer-integrated 532 nm attenuated backscatter (γ'_{532}), layer-mean attenuated backscatter color ratio ($\bar{\beta}'_{1064}/\bar{\beta}'_{532}$), and layer-integrated volume depolarization ratio for ice clouds (panels a-c), water clouds (panels d-f), and aerosol layers (panels h-j). Distributions from layers wholly unaffected by low energy pulses are shown in aqua. Distributions for LEM-affected layers are shown in orange. Statistics are computed from layers detected in nighttime measurements at 5 km horizontal resolution between 50°S-50°N, excluding the SAA, during all of 2021 (Tackett et al., 2025 © Authors 2025, CC BY 4.0).

To enable users to easily identify LEM-affected features and profiles that have been wholly excluded by the LEM algorithm, the CALIOP level 2 layer products and profile products now report a 'Low Energy Column QC Flag' for each atmospheric column in the respective data sets. These flags are bit-mapped 16-bit integers that indicate which columns contain low energy laser pulses, which columns have been wholly rejected as unusable, why rejected columns have been excluded, and which of CALIOP's altitude-dependent averaging regions contain LEM-rejected data. The interpretations of individual bits are given in Table 19. With the exception of the 333 m merged layer products, all layer products and profile products also report a Boolean 'Low Energy Mitigation Feature QC Flag'. When this flag is true, the corresponding feature (layer products) or range bin (profile products) contains either LEM-affected data, in which some low energy pulses were included in the data averaged onboard the satellite prior to downlink, or LEM-rejected data that have been wholly excluded from the level 2 analyses. The layer descriptions (e.g., top and base altitudes, integrated attenuated backscatter, etc.) in LEM-rejected columns are identified with a LEM flag value of -111.

LEM information is reported in the vertical feature mask (VFM) product via the new 'VFM Feature Detection Quality Flag'. Like the 'Feature Classification Flags', which have always been reported in the VFM files, the 'VFM Feature Detection Quality Flag' is an array of bit-mapped 16-bit integers, with one value reported for each element in the downlinked data stream. As shown in Table 20, these bits inform data users which range bins contain low energy shots, which have been rejected by LEM, and which feature finder averaging resolution was required for the feature to be detected.

Table 19: Interpretation of the individual bits in the Low Energy Column QC Flag; note that "a frame" consists of 15 consecutive laser pulses that have been averaged onboard the satellite to the vertical and horizontal spatial resolutions given in (Hunt et al., 2009). Bit 0 is the least significant bit.

Bit	Interpretation
0	Column contains LEM-affected data (data has been rejected or contains low energy pulses that LEM accepts)
1	Column belongs to a 5 km frame that has been rejected due to too many unusable profiles
2	Column belongs to a 5 km frame that has been rejected due to too many rejected subregions in altitude region 3
3	Column belongs to a 5 km frame that has been rejected due to too many rejected subregions in altitude region 4
4	Feature detection not performed at 20 km resolution in this column due to too many rejected frames
5	Feature detection not performed at 80 km resolution in this column due to too many rejected frames
6	Unused
7	Column has data rejected in altitude regions 1 and 2 (only reported at single-shot resolution)
8	Column has data rejected in altitude region 3 (only reported at single-shot resolution)
9	Column has data rejected in altitude region 4 (only reported at single-shot resolution)
10	Column does not have low energy, but data is rejected in regions 1 & 2 due to rejected data in altitude region 3 (only reported at single-shot resolution)
11–15	Unused

Table 20: bit interpretations for the 'VFM Feature Detection Quality Flag'; bit 0 is the least significant bit.

Bit	Interpretation
0	First single-shot profile has low energy
1	Second single-shot profile has low energy
2	Third single-shot profile has low energy
3	Fourth single-shot profile has low energy
4	Fifth single-shot profile has low energy
5	Bin in first single-shot profile is rejected by LEM
6	Bin in second single-shot profile is rejected by LEM
7	Bin in third single-shot profile is rejected by LEM
8	Bin in fourth single-shot profile is rejected by LEM
9	Bin in fifth single-shot profile is rejected by LEM
10	Contributed to feature detection at 1/3 km resolution
11	Contributed to feature detection at 1 km resolution
12	Contributed to feature detection at 5 km resolution
13	Contributed to feature detection at 20 km resolution
14	Contributed to feature detection at 80 km resolution
15	Unused

The effects of the low energy mitigation filtering are pervasive throughout the CALIPSO level 2 data products, particularly in the final years of the mission. To optimize the functioning of the LEM algorithms required several LEM-related changes to the lidar level 2 analyses.

- The feature classification flags (in the layer products and VFM) and the atmospheric volume descriptions (in the profile products) no longer include an "invalid" feature category. Previously, invalid features could be readily identified whenever the three least significant bits were all set to zero. Invalid features occurred extremely rarely – e.g., 872 of the 41,194,051 (~0.002%) of the unique atmospheric features detected during

all of 2012 we designated as invalid. These features were also assigned special cloud-aerosol discrimination (CAD) scores (Liu et al., 2019) to give users some insight into the failure mechanism that cause the layer to be identified as invalid. In V5.0, having the three least significant bits of a feature classification flag evaluate to zero no longer represents an 'invalid' classification of a detected layer, but instead indicates that data in the column has been rejected by the level 2 LEM filter. The spatial and optical properties within these columns (e.g., top and base altitudes, CAD scores, and optical depths) are all set to a special LEM flag value (–111). Features that were previously classified as invalid are now identified as zero-confidence clouds and they retain any of the special CAD scores that were assigned in previous data releases.

- Substantial modifications were required to add "LEM awareness" to the CALIOP boundary layer cloud clearing (BLCC) algorithm. The function of this procedure is to separate boundary layer clouds from the surrounding aerosols at single shot resolution so that the signal-to-noise ratios (SNR) of the aerosol data could be improved by large scale averaging, thus greatly reducing uncertainties in the aerosol extinction retrievals. Problems arose, however, when dense aerosols lay immediately above stratus decks. In these cases, the combined aerosol and cloud signals were often detected as a single, heterogenous feature in CALIOP's initial 5 km (15 shot) resolution profile scan (Vaughan et al., 2009). The clouds in these multi-type features were subsequently identified at single shot resolution. However, when 15 horizontally adjacent clouds were detected at single shot resolution within a single 5 km segment, a logic flaw in the BLCC scheme failed to separate these clouds from the overlying aerosol. This, in turn, introduced significant errors into the affected aerosol extinction retrievals. While this flaw was largely eliminated in the V4.51 data release (Tackett et al., 2022), some residual errors remained. When checking for the presence of 15 horizontally adjacent clouds detected at single shot resolution, the V4.51 BLCC did not account for contamination by low energy laser pulses. Since no clouds would be detected in these single shot low energy pulses, the count of horizontally adjacent clouds within the 5 km segment would fall below 15. As a consequence, these clouds were not properly separated from the overlying aerosol prior to calculating an extinction solution. Not surprisingly, the extinction solutions retrieved for these unintentionally heterogenous layers were completely unreliable. This situation is ameliorated in V5.00 by excluding low energy pulses when calculating the fraction of laser shots in which clouds are detected at single shot resolution.
- In concert with the above change to the BLCC, the calculation of the 'Single Shot Cloud Cleared Fraction' parameter reported in all 5 km averaged products was updated to correctly account for LEM-rejected profiles. As an example, if 5 clouds are detected at single shot resolution within a 5 km along-track segment that also included 5 low energy laser pulses, the single shot cloud cleared fraction in the V5.00 data products will be 5/10 = 0.5 versus 5/15 = 0.333 in the V4.51 products.
- To improve the accuracy of the surface detection algorithm and prevent spurious detections, the surface detection algorithm was modified to recognize and exclude LEM-rejected data.
- To accommodate the addition of the special LEM flag value (-111), the data types of the 'Cloud Layer Fraction', 'Aerosol Layer Fraction', and 'High Resolution Layers Cleared' scientific data sets (SDSs) were changed from unsigned integers to signed integers.

Changes to Aerosol Lidar Ratios

The Models, In situ, and Remote sensing of Aerosols (MIRA) working group was formed by an international team of aerosol research scientists shortly after the successful conclusion of the CALIPSO Version 5 Aerosol Lidar Ratio Workshop in March 2021. One of the primary goals initially identified by the MIRA group was to develop a solid empirical basis for creating a global set of spatially and seasonally varying Maps of Aerosol lidar ratios for CALIPSO (MAC). Based on in-depth analyses by MAC researchers, CALIPSO's V5.00 data release makes extensive changes to the default lidar ratios used to initiate extinction solutions for the marine and dusty marine aerosol types.

In all previous data releases, the CALIOP aerosol subtypes were characterized by a single, type specific, globally applicable lidar ratio, together with an associated uncertainty estimate that defined the lidar ratio's expected natural variability (Kim et al., 2018). For example, the lidar ratios for marine and dusty marine types were,

respectively, 23 ± 5 sr and 37 ± 15 sr. In preparation for the V5.00 release, the MAC team constructed tropospheric aerosol optical depth (AOD) estimates by subtracting the stratospheric AODs reported in CALIPSO's level 3 stratospheric aerosol product (Kar et al., 2019) from the full column AODs reported by MODIS. These tropospheric AODs were then paired with collocated CALIOP 532 nm attenuated backscatter profiles to retrieve aerosol lidar ratio estimates from constrained solutions of the lidar equation (Toth et al., 2025). Retrievals were limited to cloud-free scenes over oceans in which CALIPSO identified only a single aerosol type (i.e., either marine or dusty marine) within a 5 km averaged column. Solutions were calculated for all qualifying measurements obtained while CALIPSO flew in the A-Train with MODIS-Aqua, from June 2006 through August 2018. The retrieved lidar ratios for each of the two aerosol types were then aggregated into four seasonal maps (winter, spring, summer, and fall) having a spatial resolution of 2° latitude × 4.8° longitude. Grid cells having fewer than 50 samples – e.g., in polar winters, when the MODIS orbit does not allow for the daytime measurements necessary to retrieve AOD – are excluded from these maps. These excluded values are replaced with an approximation derived from the relationship between the constrained lidar ratios and the sea salt volume fraction (SSVF) computed using parameters reported by the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model (Toth et al., 2025). For any V5.00 CALIOP footprint, the locally appropriate lidar ratio and its associated uncertainty are interpolated in both time and space from the map data.

Figure 12 shows the seasonal maps developed for marine aerosols. Seasonal and regional variations in marine lidar ratios are plainly visible and often show a marked difference from the 23 ± 5 sr value used globally in previous data releases. These differences highlight an important conceptual difference between the V5.00 marine lidar ratios and the marine lidar ratios in previous data releases. Previously, the marine lidar ratio represented our best approximation of "clean marine" conditions (e.g., see Kim et al., 2018). In V5.00, marine lidar ratios are assigned to aerosols detected over oceans that are dominated by marine constituents but that can also be mixed with other, mostly anthropogenic components. As seen in Figure 12, genuinely pristine marine aerosols are largely confined to the remote southern oceans. Coastal regions frequently exhibit a pronounced influence from other aerosol sources that elevates the lidar ratios to values significantly higher than those seen in the southern oceans.

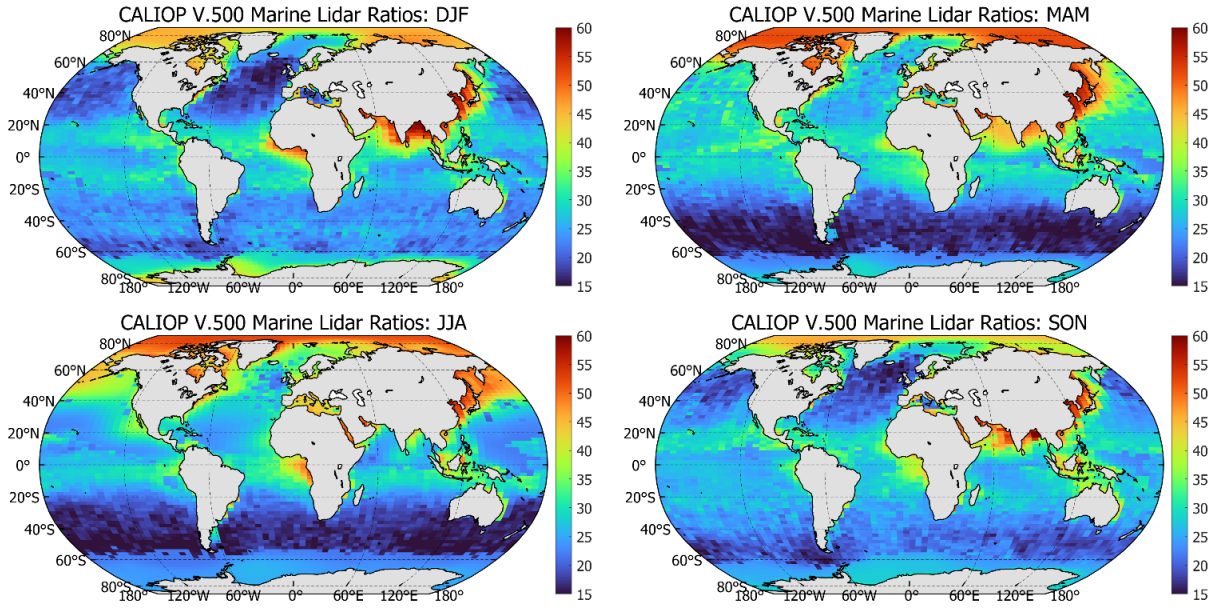


Figure 12: CALIOP initial marine lidar ratios (units = sr) derived from CALIOP 532 nm attenuated backscatter profiles constrained by collocated MODIS aerosol optical depths. The top left shows median values for the winter months (December, January, and February); the top right shows median values for the spring (March, April, and May); the bottom left shows median values for the summer (June, July, and August); and the bottom right shows median values for the fall (September, October, and November).

Our approach for constructing dusty marine maps was, with two notable exceptions, essentially identical to the method we used for marine aerosols. The first exception is that for the dusty marine maps we also included lidar

ratio solutions constrained by CALIOP's ocean-derived column optical depth (ODCOD; Ryan et al., 2024) retrievals. For the marine lidar ratios, ODCOD retrievals were excluded so that they could later be used to validate the MODIS constrained retrievals. However, for the dusty marine type, sun glint contaminating the collocated MODIS measurements precluded their use as an AOD constraint in important dust transport corridors (e.g., during boreal summers in the Caribbean and central western Atlantic). Consequently, we chose to include ODCOD constrained solutions in developing the dusty marine maps. The second exception is that, unlike the marine lidar ratio maps, the seasonal variability in the dusty marine maps is limited to the green shaded region shown in Figure 13. While other parts of the hydrosphere showed little, if any, seasonal variation (perhaps due to temporal changes in dust concentrations), significant seasonal variations in dusty marine lidar ratios were observed over the Mediterranean Sea, the Persian Gulf, the Arabian Sea, and the Bay of Bengal. Furthermore, the seasonal sample counts in these regions were large enough in each season to retrieve confident lidar ratio estimates. Neither of these conditions were met simultaneously in any other extended regions of the globe.

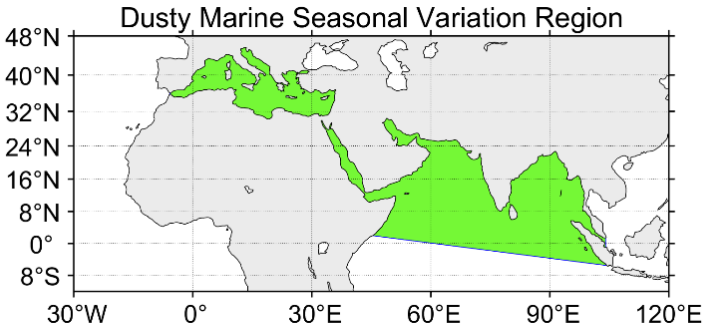


Figure 13: the green shaded area of the map indicates those regions where CALIOP's dusty marine lidar ratios are observed to vary seasonally.

The V5.00 dusty marine maps are shown in Figure 14. As expected, the dusty marine lidar ratios are uniformly larger than the marine lidar ratios. The spatial patterns of the two types are generally similar, simply because anthropogenic aerosol transport patterns are essentially identical for both types and the inclusion of anthropogenic aerosols elevates both sets of values.

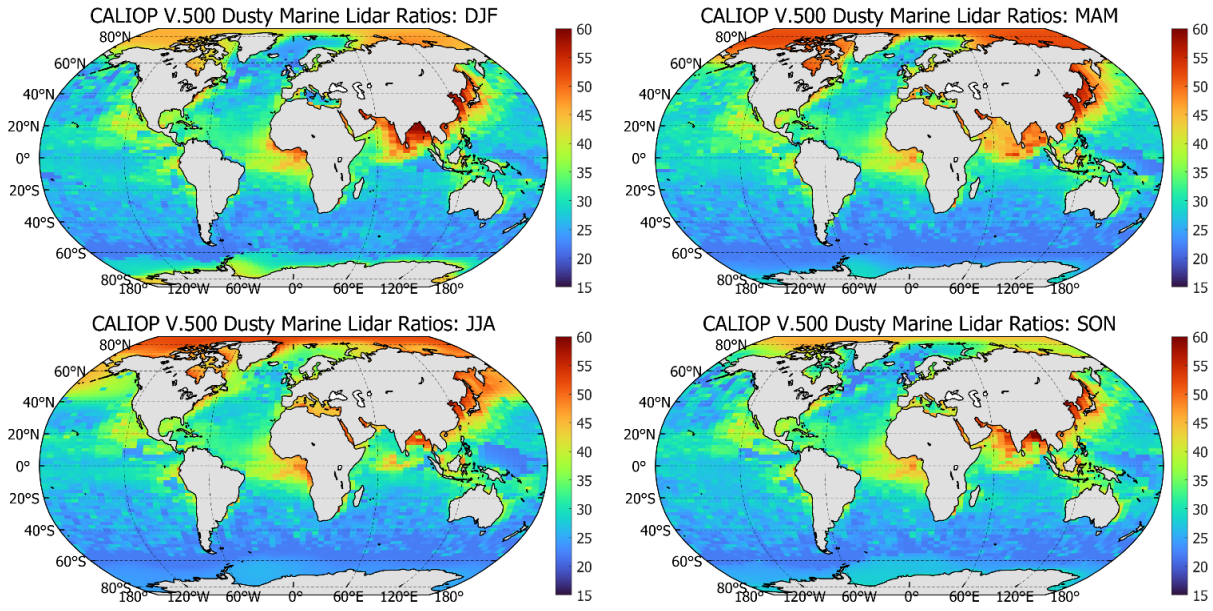


Figure 14: CALIOP initial dusty marine lidar ratios (units = sr) derived from CALIOP 532 nm attenuated backscatter profiles constrained by collocated aerosol optical depth measurements from MODIS and ODCOD. The top left shows median values for the winter months (December, January, and February); the top right shows median values

for the spring (March, April, and May); the bottom left shows median values for the summer (June, July, and August); and the bottom right shows median values for the fall (September, October, and November).

A second change to the CALIOP aerosol lidar ratios is the selection of a new value of 45 ± 24 sr for 1064 nm retrievals of clean continental aerosol. This change is based on field measurements of continental aerosol made by Haarig et al., 2025 using a 1064 nm rotational Raman lidar. The 532 nm lidar ratio for the clean continental type remains unchanged at 53 ± 24 sr.

Changes to Aerosol Type Identification

In addition to the MAC lidar ratio work, the V5.00 data set contains a second fairly substantial (albeit localized) change. CALIOP's V4.51 aerosol subtyping algorithm classified planetary boundary layer (PBL) aerosols located over ocean surfaces as marine, with ocean surfaces being identified using the CALIPSO land-water mask. In Arctic winter however, most of the "ocean surfaces" reported by the land-water mask are actually extended over-ocean ice sheets, and the aerosol above is not marine but instead some species of anthropogenic aerosol arriving in the polar regions via long range transport. In these cases, assigning a marine aerosol lidar ratio to the aerosol layer both misrepresents the true aerosol type and leads to sometimes substantial underestimates of aerosol optical depth. By expanding the surface type identification scheme to also include snow and ice data reported in the CALIOP level 1b files and by making small modifications to the aerosol subtyping flowchart (see Figure 15), the V5.00 CALIOP data product now more correctly classifies a large fraction of these "over frozen ocean" aerosols.

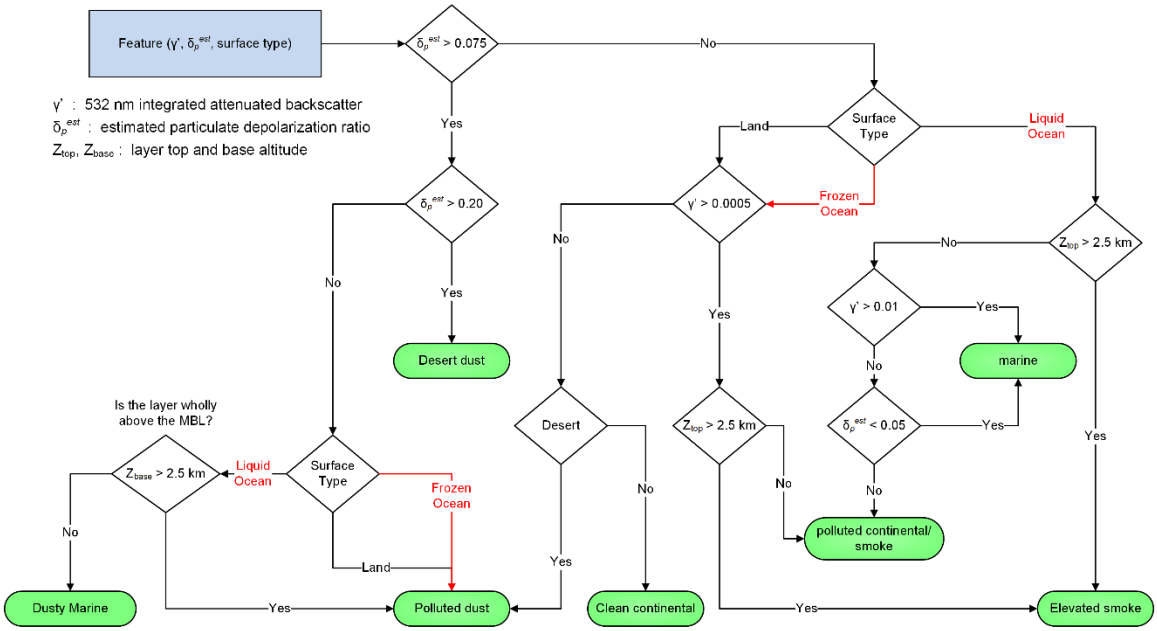


Figure 15: revised tropospheric aerosol subtyping flowchart; paths and labels in red were added for the V5.00 release.

New ODCOD Retrieval Confidence Flag

In response to data user requests, a new confidence flag has been added to the existing 'ODCOD QC Flag' SDS reported in the 'Ocean Derived Column Optical Depth' Vgroup included in all layer products and profile products. Table 22 describes how to interpret all of the bits used in the 'ODCOD QC Flag'. The new confidence flag (shown in orange) is reported in bit 6. In this new configuration, all valid ODCOD retrievals will have an 'ODCOD QC Flag' of 127 or less. All high confidence ODCOD retrievals will have an 'ODCOD QC Flag' value of 63 or less. Any of the following conditions will cause bit 6 to be toggled on, indicating a low confidence solution.

a) Wind speed out of range; the AMSR corrected MERRA-2 wind speed is less than 3 m/s or greater than 15 m/s

b) Surface integrated depolarization ratio is greater than 0.05, suggesting that the surface is not pure sea water

c) Surface integrated attenuated backscatter exceeds the upper bound specified in Table 21.

Table 21: maximum values of surface integrated attenuated backscatter for confident retrievals; higher values may include undetected surface saturation.

	Daytime	Nighttime
532 nm	0.0413 sr^{-1}	0.0353 sr^{-1}
1064 nm	0.0384 sr^{-1}	0.0325 sr^{-1}

d) The number of range bins shifted by the altitude registration algorithm (Powell et al., 2009) is not constant over all profiles used to create the averaged surface return. This is only a concern for averaged profiles where having multiple "bins shifted" will distort the shape of the spatially averaged surface return and degrade the quality of the fit to the CALIOP response model (CRM; Ryan et al., 2024). To aid in diagnosing the presence of non-constant bin shifts, a new SDS ('ssNumber Bins Shift') was added to the 'Single Shot Detection' Vgroup reported in the 5 km layer products.

Table 22: interpretation of the individual bits in the ODCOD QC flags (adapted from Ryan et al., 2024); bit 0 is the least significant bit.

Bit	Interpretation
0	Of the measurements provided to the ODCOD algorithm by the surface detection algorithm, ODCOD adjusted the CRM such that the first data point of the surface detection data was not the first point on the CRM
1	The surface detection algorithm provided surface measurements covering a range greater than 120 m
2	The ODCOD algorithm used range bins above the top of the detected surface altitude
3	The ODCOD algorithm used range bins below the base of the detected surface altitude
4	When solving for the alignment of the CRM, the first measurement that should fall on the CRM curve was not originally provided by the surface detection algorithm
5	ODCOD had to adjust the CRM such that the first measurement provided by the surface detection algorithm was not the first point on the CRM
6	ODCOD retrieval confidence flag; a value of 1 indicates a confident retrieval whereas 0 signals low confidence
7-9	Unused
10	The surface detection algorithm did not find a surface
11	The International Geosphere–Biosphere Programme (IGBP) surface type is not 17 for ocean
12	The depolarization ratio of the surface is greater than 0.15
13	The corrected MERRA-2 wind speed is outside of the inclusive range 0.025 m s^{-1} to 43 m s^{-1}
14	ODCOD has failed to find the time delay of the CRM from the surface measurements provided by the surface detection algorithm
15	The vertical extent of the detected surface contains too few samples to derive a time delay solution
16	When solving for the CRM area, the solution grew unrealistically large
17	A failure occurred while attempting to solve for the scale factor
18	Surface saturation was detected in the surface return
19	Negative signal anomaly was detected in the surface return
20	The detected surface contains negative or invalid backscatter measurements
21	Necessary input data is either fill or invalid
22	The surface detection algorithm had to resort to an alternative method of finding the surface when the surface return was averaged to coarser resolutions that may not be reliable for ODCOD

Bit	Interpretation
23–31	Unused

Wrong profile starts for a new chunk after missing data gap

The CALIOP level 2 processing ingests level 1 data in 80 km "chunks", equivalent to the largest horizontal averaging resolution used in the search for features. Each chunk contains 240 consecutive single-shot profiles distributed uniformly across 16 consecutive 5-km (15 shot) frames. If there are any missing data within these 240 consecutive profiles, the candidate chunk is discarded, and the level 2 processing resumes beginning with the first single shot profile in the first frame available after the data gap. However, the V4.51 and earlier level 2 code contained an "off by one" error that caused this processing restart to begin with the second profile in the frame instead of the first. As a consequence of eliminating this error, in granules containing data gaps – e.g., for planned activities such as boresight alignments and polarization gain ratio calibrations or for unplanned events such as the repeated ground station downlink anomalies that occurred between 21:20:46 UTC on 2020-12-12 and 00:12:12 UTC on 2020-12-13 – the level 2 chunks built following those data gaps will use a somewhat different collection of single shot profiles in V5.00 relative to all previous versions.

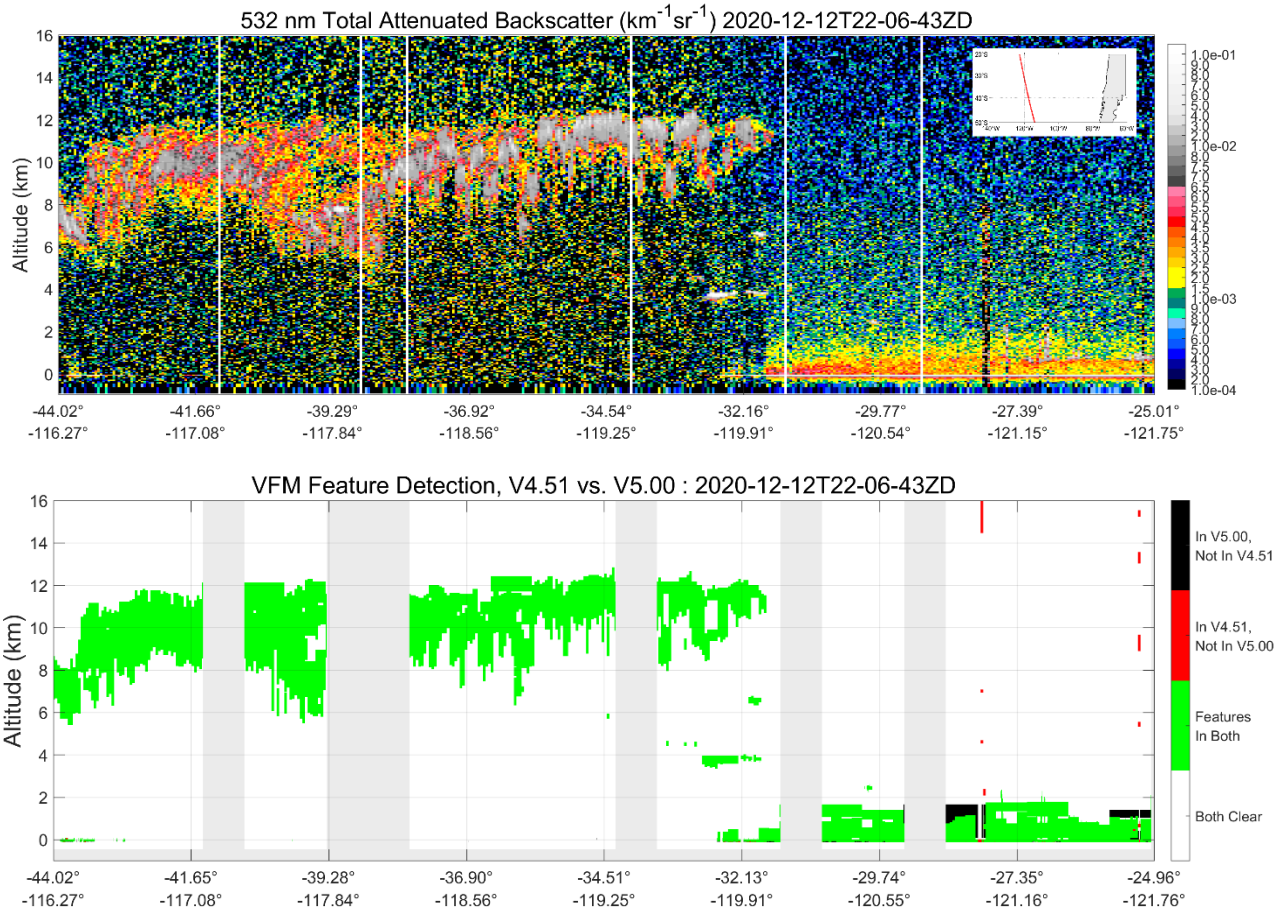


Figure 16: The upper panel shows 532 nm total attenuated backscatter coefficients measured in the southern Pacific Ocean on 2020-12-12 at ~22:27 UTC. The white vertical stripes in the image represent frames (i.e., 15 consecutive shots) containing missing data. The lower panel compares V4.51 and V5.00 VFM feature detections for the scene shown in the upper panel. The gray vertical stripes show where chunks (i.e., 240 consecutive shots) were rejected by the level 2 processing algorithms due to missing data in one or more included frames.

Figure 16 shows an example of the kinds of changes that can occur due to this change in profiles averaged. The upper panel shows daytime data acquired in the southern Pacific Ocean on 2020-12-12 at ~22:27 UTC. On this day, data downlink from the satellite was intermittently interrupted by ground station errors. The white vertical strips

in the upper panel show where individual frames were missing or irretrievably corrupt in the downlink data stream. The lower panel is a "red-green-black" diagram comparing feature detections in V4.51 to those in V5.00. The green areas represent features detected in both data processing versions, while the red areas show features that were detected in V4.51 but not V5.00 and the black areas show the opposite; i.e., features that were detected in V5.00 but not V4.51. Regions where no features were detected in either analysis are shown in white. The vertical gray bars show where chunks have been rejected by the level 2 processing due to the presence of missing data. In some cases – e.g., following the first to data gaps beginning at ~41.5°S and ~37.2°S – there are no layer detection changes. On the other hand, the region following the final data gap shows substantial changes, though not all of them can be attributed to the change in the selection of pseudo-single shot profiles to average. The red range bins not detected in V5.00 were eliminated by the LEM algorithm described above. In this and other segments, the intermittent detection changes in individual range bins adjacent to the Earth's surface occur due to V5.00 changes in VFM altitude registration, as described below in the Revised VFM Layer Detection Altitudes section. The additional aerosol detections highlighted by the black areas in this data segment are a fortuitous result of the combination of a shift in profiles averaged combined with LEM filtering. We note, however, that additional feature detections are not a guaranteed outcome of this code change.

HDF File Changes

The V5.00 CALIPSO data products are distributed as Hierarchical Data Format Version 4 (HDF4) files, consistent with the EOS requirement in effect when CALIPSO launched in 2006. Since launch, there have been substantial technological advances in data discoverability and access to resources. To make CALIPSO data more readily accessible to the scientific community beyond the life of the mission and take advantage of newer data access capabilities, several modifications were made to the format and content of the CALIOP HDF files. These include:

- Updating all units to conform to NetCDF Climate and Forecast (CF) metadata conventions.
- Ensuring that all dimensions are named, to allow HDF to NCDF conversions using commercial off the shelf (COTS) tools that currently exist.
- Creating/expanding SDS attributes and comments to make the data products more self-documenting.

New Data Product: CAL_LID_L2_MLay_Diagnostic-Beta-V5-00*.hdf

The newly released CAL_LID_L2_MLay_Diagnostic-Beta-V5-00*.hdf data products are augmented versions of the standard CAL_LID_L2_05kmMLay_Standard-V5-00*.hdf products. The production strategy for these products is 'Beta', indicating that some of the included parameters have not been fully validated. In particular, these 'not fully validated' parameters include the cloud values reported in the 1064 nm derived optical properties SDSs (e.g., 'Feature Optical Depth 1064'). While the Diagnostic Beta products contain all the information reported in the Standard products, they also contain the five SDSs listed below that are not reported in the Standard products.

1. Low Energy Mitigation Internal Column QC Flag – a 16-bit unsigned integer that offers extra insight into the internal workings of the LEM algorithm. Bit interpretations are given in Table 23.
2. Initial Integrated Attenuated Backscatter 532 – The standard 'Integrated Attenuated Backscatter 532' (γ'_{std}) values reported in all CALIOP lidar level 2 layer files are calculated assuming that the feature attenuated backscatter coefficients have been corrected for signal attenuation due to all layers detected above. The 'Initial Integrated Attenuated Backscatter 532' (γ'_{init}) values are recorded immediately after feature detection during the initial 5 km search for features. At this point in the processing, no single shot cloud clearing has been done nor have the layers been classified by the CAD algorithm. Assuming perfect detection and classification and perfect knowledge of layer lidar ratios and multiple scattering factors, $\gamma'_{init} = T^2_{above} \times \gamma'_{std}$, where T^2_{above} is the two-way transmittance of all overlying particulate layers. In special cases, when particulate layers lie above opaque water clouds, measurements of γ'_{init} can be used to derive estimates of T^2_{above} and hence obtain estimates of the overlying particulate optical depths (Hu et al., 2007). These measurements of γ'_{init} are an essential input required for calculating the 'Column Particulate Optical Depth Above Opaque Water Cloud 532' values reported in all layer products.

3. Initial Integrated Attenuated Backscatter Uncertainty 532 – uncertainty estimates for the 'Initial Integrated Attenuated Backscatter 532' which are subsequently used to calculate the 'Column Particulate Optical Depth Above Opaque Water Cloud Uncertainty 532' values.
4. Initial Integrated Attenuated Backscatter 1064 – this is the 1064 nm analog of the 'Initial Integrated Attenuated Backscatter 532'. Because CALIOP does not make depolarization ratio measurements at 1064 nm, this parameter is unused in the V5.00 analyses. In any (currently unforeseen) future version of the CALIOP data products, the 'Initial Integrated Attenuated Backscatter 1064' values could potentially be used in the algorithm proposed by Chand et al., 2008 to derive column Ångström exponents for particulates above opaque water clouds.
5. Initial Integrated Attenuated Backscatter Uncertainty 1064 – this is the 1064 nm analog of the 'Initial Integrated Attenuated Backscatter Uncertainty 532'. Because CALIOP does not make depolarization ratio measurements at 1064 nm, this parameter is unused in the V5.00 analyses.

Table 23: internal low energy mitigation (LEM) quality control flags set during LEM algorithm operations; bit 0 us the least significant bit

Bit	Interpretation
0	Single shot columns have data rejected in regions 1 and 2 but the 5 km resolution column was NOT rejected by LEM
1	Column has data rejected in altitude region 3 but the 5 km resolution column was not rejected by LEM
2	Column has data rejected in altitude region 4 but the 5 km resolution column was not rejected by LEM
3	Full column is rejected by LEM
4	Column rejected due to altitude region 3
5	Frame rejected due to too many low energy shots
6	Frame rejected due to too many rejected subregions in altitude region 3
7	Frame rejected due to too many rejected subregions in altitude region 4
8	Feature detection was not attempted at 20 km resolution due to too many LEM-rejected columns
9	Feature detection was not attempted at 80 km resolution due to too many LEM-rejected columns
10	Feature detection was not attempted at resolution 6 due to too many LEM-rejected columns
11	Column contains LEM affected data in altitude region 3
12	Column contains LEM affected data in altitude region 4
13-15	Unused

In addition to the new SDSs, the Diagnostic Beta products also contain an extra Vgroup. The '01 km Detection' Vgroup replicates the majority of the SDSs reported in the CAL_LID_L2_01kmCLay_Standard-V5-00*.hdf products. Information in the excluded SDSs is either (a) replicated in the 5 km SDS of the same name (e.g., 'Day Night Flag' and 'IGBP Surface Type') or (b) easily derived from the SDS of the same name reported in the 'Single Shot Detection' Vgroup (e.g., 'Solar Zenith Angle', 'Solar Azimuth Angle', and 'Troposphere Height').

Finally, the 'Ocean Derived Column Optical Depth' (ODCOD) Vgroup in the Diagnostic Beta products augments the information contained in the Standard products by reporting column particulate optical depth retrievals at 1064 nm. In addition, ODCOD Vgroup in the Diagnostic Beta products contains a wealth of diagnostic information and intermediate calculations for the ODCOD retrievals at both wavelengths. Comments in the HDF attributes for each SDS in this Vgroup describe the function of each parameter in the ODCOD calculation chain. Researchers considering using this information should first very thoroughly review the ODCOD retrieval technique described in Ryan et al., 2024.

Unique Layer IDs for Features Detected at 5 km, 20 km, and 80 km Resolutions

Within each granule, unique layer IDs are now assigned for all layers detected at 5 km, 20 km, and 80 km averaging resolutions. Unique layer IDs are unsigned 16-bit integers that enable unambiguous mapping between the

integrated optical properties for the individual layers reported in the 5 km layer products and the corresponding profiles of vertically resolved optical properties reported in the 5 km profile products. Because extinction retrievals are not attempted for layers detected at single shot and 1 km resolutions, unique layer IDs are not assigned for the features reported in the 333 merged layer files and the 1 km cloud layer files.

In the 5 km aerosol, cloud, and merged layer files, unique layer IDs are reported in the same manner as layer optical properties. Recall that layer properties are reported at a uniform 5 km along-track resolution, and that for layers detected at 20 km and 80 km resolution, these properties are replicated over four 5 km columns for 20 km detections and sixteen 5 km columns for 80 km detections. Consequently, while a unique layer ID may appear as many as sixteen times in a single data file, these sixteen instances all reference the same layer, and the layer properties reported for all instances – e.g., top and base heights, integrated attenuated backscatters, and optical depths – will also replicated across all sixteen 5 km columns. In the 5 km profile products, unique layer IDs are reported at the same latitude, longitude, and time coordinates as in the layer products. But, instead of appearing once in each 5 km column, layer IDs are replicated over the full vertical extent of a layer. For example, for a 480 m thick aerosol layer detected at 80 km horizontal averaging in a rectangular region would extend over eight 60 m altitude bins and sixteen 5 km profile segments.

New 'Scene Flag' SDS Added to All Layer Products

To enable rapid identification of all columns containing specific feature types, the CALIPSO layer products now include a 'Scene Flag' SDS. The Scene Flag is implemented as a 32-bit integer with individual bit interpretations as defined in Table 24. Fill values (-9999) are used to identify columns that do not pass the LEM quality checks. Users are therefore warned to check the sign of the Scene Flags and discard/ignore any negative values before querying the flags to determine column feature types.

Table 24: Bit interpretations for positive values of the new 'Scene Flag' SDS. Bit 0 is the least significant bit.

Bit	Interpretation
0	column contains tropospheric aerosol, marine subtype
1	column contains tropospheric aerosol, dust subtype
2	column contains tropospheric aerosol, polluted continental/smoke subtype
3	column contains tropospheric aerosol, clean continental subtype
4	column contains tropospheric aerosol, polluted dust subtype
5	column contains tropospheric aerosol, elevated smoke subtype
6	column contains tropospheric aerosol, dusty marine subtype
7	column contains stratospheric aerosol, polar stratospheric aerosol subtype
8	column contains stratospheric aerosol, volcanic ash subtype
9	column contains stratospheric aerosol, sulfate subtype
10	column contains stratospheric aerosol, elevated smoke subtype
11	column contains stratospheric aerosol unclassified
12	column contains ice clouds composed of randomly oriented crystals
13	column contains ice clouds composed of horizontally oriented crystals
14	column contains water clouds
15	column contains clouds with unknown ice-water phase
16–31	Unused

More Readily Accessible Altitude Information

To make CALIOP altitude information easier to access, new 'Lidar Data Altitudes' SDSs have been added to all level 2 data products. The 'Lidar Data Altitudes' array reports the altitudes, relative to mean sea level, that specify the vertical midpoints of all range bins in the profile measurements downlinked from the CALIPSO satellite. In all previous versions of the CALIPSO lidar data products, the 583-element 'Lidar Data Altitudes' array was stored only

in the file metadata. To ensure backward compatibility with existing software, the metadata copy of the 'Lidar Data Altitudes' is retained in all files.

In the V5.00 vertical feature mask (VFM) product, the altitude array now contains 545 elements instead of the 583 elements reported in all other products. Because the VFM product does not report layer detections in the highest (~30 km to ~40 km) or lowest (~-0.5 km to ~-2.0 km) CALIOP onboard averaging regions, the superfluous altitude bins were eliminated. Making this change now matches the VFM altitude array size to the vertical dimension of the (unpacked) 'Feature Classification Flags' SDS.

Revised VFM Layer Detection Altitudes

Some changes were made to the altitude registration of layer tops and bases in the VFM product to more accurately reflect the top and base altitudes reported in the CALIOP layer products. Differences between the V4.51 and V5.00 VFM files are small, confined to the high resolution (i.e., single shot) data regime between –0.5 km and 8.2 km, and typically no more than ±1 range bin. The images below show a particularly egregious example of altitude mismatches between V4.51 and V5.00. Approximately 1% of all range bins reporting feature detection in either V4.51 or V5.00 were detected only in V4.51. Similarly, approximately 0.4% were detected only in V5.00. This apparent altitude registration anomaly is, fortunately, confined solely to the VFM files. When comparing V4.51 data products versus V5.00, the layer top and base altitudes reported in the 333 m merged layer files, the 1 km cloud layer files, and the 5 km merged layer files are identical for all layers detected in this granule.

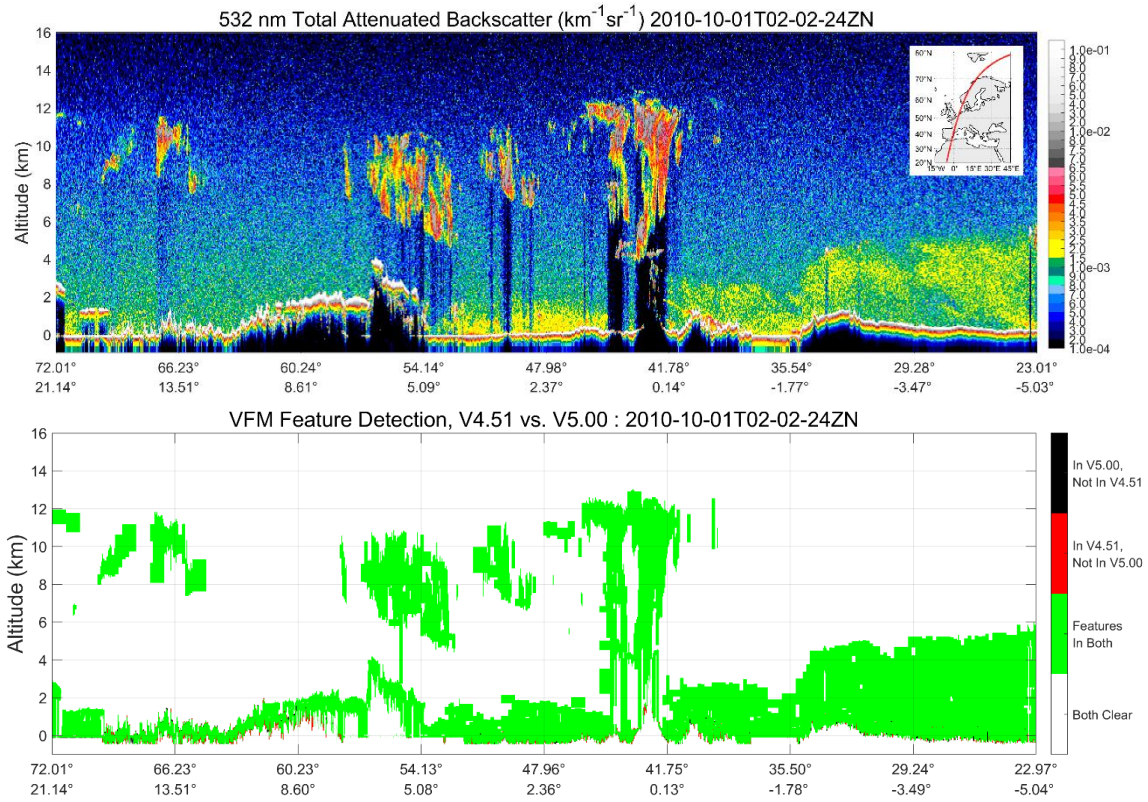


Figure 17: CALIOP nighttime observations of clouds and aerosols over Scandinavia, western Europe, and into northern Africa. These measurements were acquired on 2010-10-01 beginning at 02:02:42 UTC. The top panel shows attenuated backscatter coefficients measured at 532 nm. The bottom panel shows the VFM differences between V4.51 and V5.00. In the majority of the cases, V5.00 extends one range bin below the layer base altitude reported in the V4.51 analyses.

Summary of Newly Added and Removed SDSs

The tables below contain a comprehensive listing of all SDSs that were either added or removed. Included in the listing are SDSs that were renamed in the transition from V4.51 to V5.00.

Table 25: new scientific data sets reported in the CALIOP V5.00 lidar level 2 data products; horizontal resolution suffixes are A = aerosol, C = cloud, and M = merged layers (i.e., both aerosol and cloud).

	VFM	Layer					Profile	
SDS name		**1/3kmM**	**1kmC**	**5kmA**	**5kmC**	**5kmM**	**5kmA**	**5kmC**
Lidar_Data_Altitudes		✓	✓	✓	✓	✓	✓	✓
Low_Energy_Mitigation_Column_QC_Flag		✓	✓	✓	✓	✓	✓	✓
Low_Energy_Mitigation_Feature_QC_Flag			✓	✓	✓	✓	✓	✓
Scene_Flag		✓	✓	✓	✓	✓		
DEM_Surface_Elevation							✓	✓
Unique_Layer_ID				✓	✓	✓	✓	✓
VFM_Feature_Detection_Quality_Flag	✓							
Number_Bins_Shift		✓						
Extinction_QC_Flag_532				✓	✓	✓		
Extinction_QC_Flag_1064				✓	✓	✓		
Single_Shot_Detection::ssLow_Energy_-Mitigation_Column_QC_Flag			✓	✓	✓	✓	✓	✓
Single_Shot_Detection::ssScene_Flag			✓	✓	✓	✓	✓	✓
Single_Shot_Detection::ssNumber_-Bins_Shift			✓	✓	✓	✓	✓	✓

Table 26: scientific data sets removed from the CALIOP V5.00 lidar level 2 data products; horizontal resolution suffixes are A = aerosol, C = cloud, and M = merged layers (i.e., both aerosol and cloud).

	VFM	Layer					Profile	
SDS name		**1/3kmM**	**1kmC**	**5kmA**	**5kmC**	**5kmM**	**5kmA**	**5kmC**
Samples_Averaged							×	×
Surface_Elevation_Statistics							×	×
ExtinctionQC_532				×	×	×		
ExtinctionQC_1064				×	×	×		

Data Quality Statement for CALIPSO's Version 4.51 Lidar Level 2 Data Product Release

Data Version: **4.51**
Data Release Date: **June 1, 2023**
Data Date Range: **June 13, 2006 to June 30, 2023**

The new version 4.51 (V4.51) of the CALIPSO lidar (CALIOP) Level 2 (L2) data products contains a number of improvements and additions over the previous version (V4.2) that was released in October 2018. A summary of the major changes addressed in this release is detailed below, as well as a section highlighting known issues.

Improved Smoke Layer Accuracy above Clouds

The CALIOP L2 feature detection algorithm employs an iterative thresholding technique applied to profiles of 532nm attenuated scattering over differing horizontal resolutions to identify and vertically resolve layers (Vaughan et al., 2009). Rajapakshe et al., 2017 identified instances when CALIPSO was under-reporting the vertical extent of dense smoke layers above low-level clouds when compared with coincident observations made by the Cloud-Aerosol Transport System (CATS) lidar instrument. Figure 1 (a, c) shows one of those specific over-flights, in which the 532nm total attenuated backscatter (a) and vertical feature mask (c) show a clear delineation on the order of 1 km between the base of the upper-level smoke layer and the top of a marine cloud layer below. Further examination of this scene showed that this smoke layer was dense enough to rapidly attenuate the 532nm signal to such a degree that the CALIOP feature detection algorithm artificially elevated the layer base. This is confirmed by seeing that the rapid attenuation is not seen in the 1064 nm signal response (Figure 1b), where the aerosol

extinction is not as great as that seen in the 532nm signal. The 1064 nm response, which appears to close the clear-air gap between the smoke and cloud layer, is also why CATS was able to fully identify the smoke layer as that instrument uses 1064 nm for feature detection. The inability of CALIPSO to fully detect these layers can have the consequence of underestimating global aerosol optical depths, critically important for the derivation and understanding of radiative forcings (Lu et al., 2018).

In the version 3 (V3) release of the CALIOP L2 data a technique was developed for the feature finder algorithm to extend the base of near-surface aerosol layers done to the surface. A variant of this same approach was applied for these overlying dense smoke layer cases. Several criteria need to be met before the layer could be extended, a summary of which are detailed in the V4.51 CALIOP layer level 2 data product description.

Figure 1d shows the application of this new technique, in which the clear-air gap has been closed. The mean 532nm optical depth of these extended smoke layers has also increased, from 0.94 to 0.122, an increase of nearly 30%. A global representation of all extended layers for the month of August 2016, seen in figure 2a, show a concentration in the south-eastern Atlantic during the African burn season, the same region noted in the CALIPSO/CATS comparisons previously documented. Figure 2b shows the impact on the 532nm optical depth for those extended smoke layers, where in a majority of those cases in which the layer was extended the optical depths increase. The few outliers that exist have been traced to instances when the negative values of the 532nm extinction are included, which could happen during rapid attenuation of the signal.

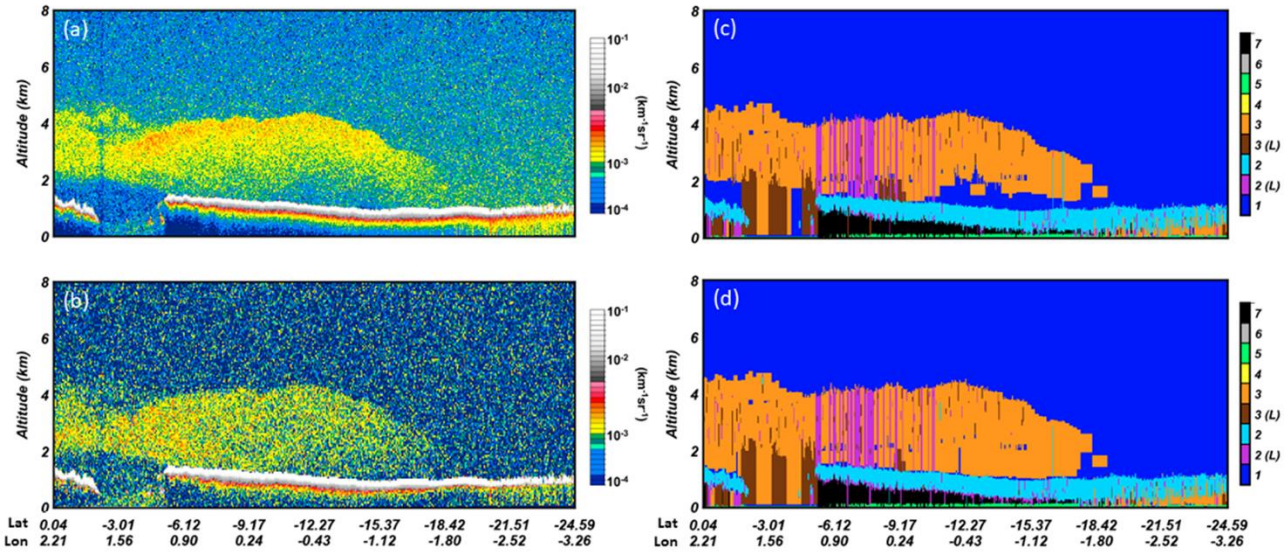


Figure 1: (a) 532 nm total attenuated backscatter from the V4.51 CALIOP Level 1B, (b) 1064 nm attenuated backscatter from the V4.51 CALIOP Level 1B, (c) feature type from the V4.21 CALIOP Level 2 Vertical Feature Mask (VFM), and (d) feature type from the V4.51 CALIOP Level 2 VFM between 1:35:03 and 1:41:44 on August 6th, 2016.

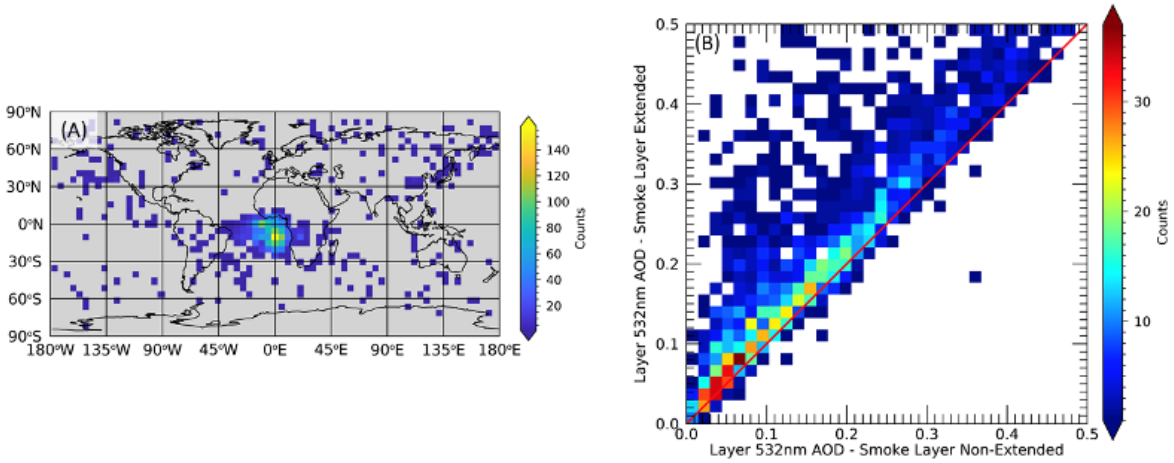


Figure 2: (a) Global distribution of all (day and night) extended smoke layers for August 2016 and (b) 532nm aerosol optical depth changes between extended (y-axis) and non-extended (x-axis) smoke layers. Data used to generate both plots were pulled from the V4.51 Lidar Level 2 5km Merged layer files. Non-extended smoke layers were created by running the V4.51 Lidar Level 2 algorithm but with the smoke extension turned off.

Improved Classifications for Stratospheric Aerosols

The stratospheric aerosol subtyping algorithm has been updated in V4.51 to improve the ability to discriminate between volcanic ash and depolarizing smoke from pyrocumulonimbus (pyroCb) injections, improve the fidelity of the sulfate classification, and to update the 532 nm volcanic ash lidar ratio to the current state of knowledge. The stratospheric aerosol subtypes are now: volcanic ash, sulfate, smoke, polar stratospheric aerosol, and unclassified. A full description of changes with the V4.51 algorithm are documented in Tackett et al., 2023a.

Summary of changes:

- The stratospheric aerosol subtyping algorithm now relies solely on depolarization and integrated attenuated backscatter (γ') to discriminated between volcanic ash, sulfate, and smoke. Removing the reliance on layer color ratio that was included in the V4.2 subtyping algorithm allows more accurate identification of sulfate and less misclassification of sulfate as smoke.
- The estimated particulate depolarization ratio threshold used to discriminate between smoke from pyroCb injections and volcanic ash has been increased from 0.15 to 0.25 based on empirical analysis of recent major pyroCb events.
- The V4.2 "sulfate/other" stratospheric aerosol subtype has been separated into two subtypes in V4.51 - sulfate, and unclassified. Previously, the sulfate/other type included both aerosol layers likely to be sulfate and layers with low values of γ', representing the "other" component of the combined class. In V4.51, layers with low γ' are now reported separately as "unclassified".
- The 532 nm lidar ratio for volcanic ash has been increased to 61 ± 17 sr based on CALIOP-constrained retrievals of volcanic ash plumes.

Figure 3 demonstrates the improvements in stratospheric aerosol subtyping between V4.2 and V4.51. Panels in each row show the frequency of classification of each stratospheric aerosol subtype for manually identified layers associated with a dominant aerosol type. In V4.51, there are fewer misclassifications of volcanic ash as sulfate, more accurate classifications sulfate from true sulfate events, and fewer misclassifications of true smoke as volcanic ash. There remain some misclassifications of sulfate as smoke due to the difficulty of separating these types with CALIOP observables.

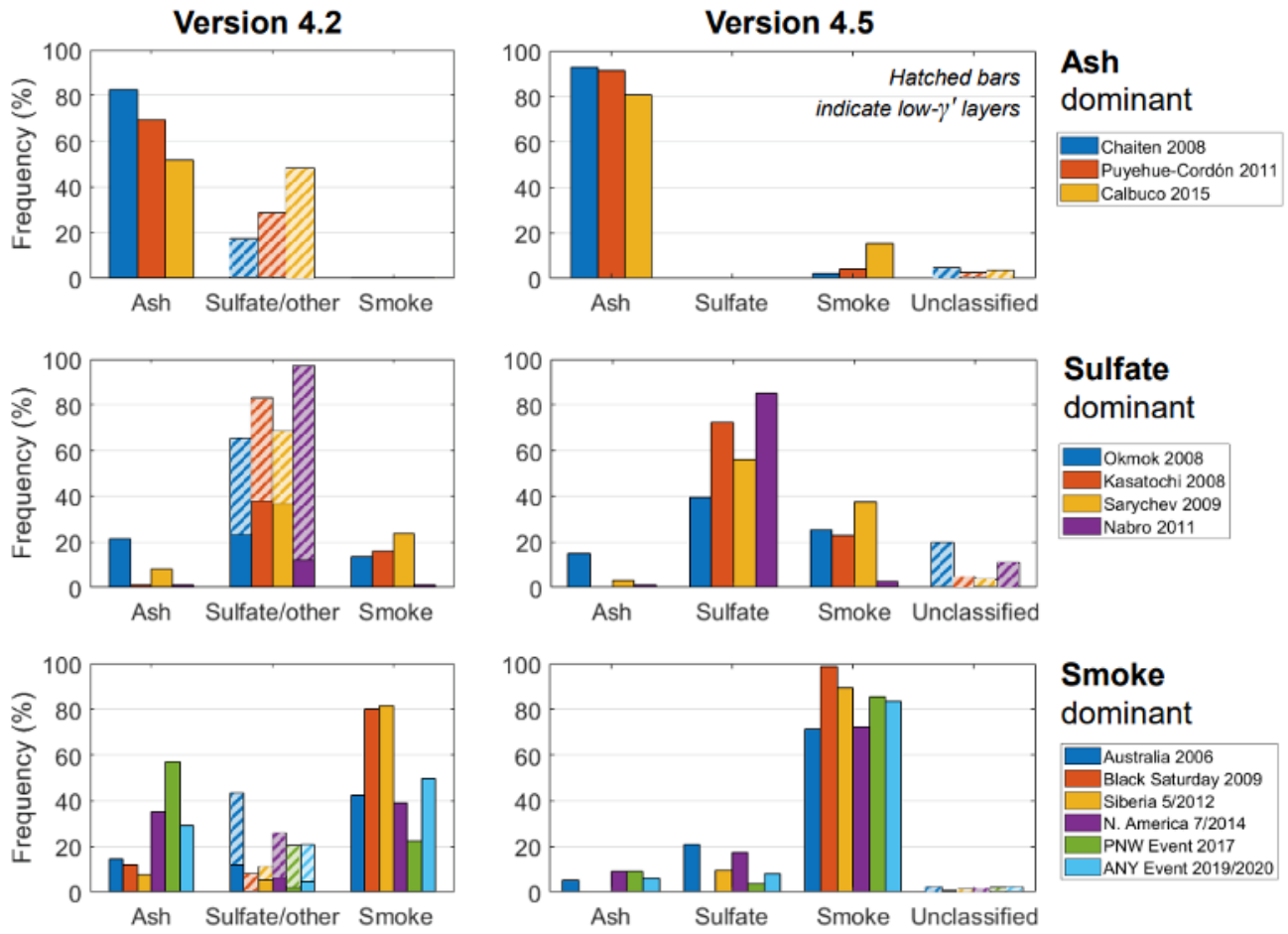


Figure 3: Stratospheric aerosol subtype classification frequency for events dominated by volcanic ash (top row), sulfate (middle row), and smoke (bottom row). V4.2 and V4.51 classifications are in the left and right columns, respectively (Tackett et al., 2023a © Authors 2023, CC BY 4.0).

Total Column Particulate Optical Depths Above Opaque Water Clouds

Hu et al., 2007 developed an approach to estimate 532nm total particulate optical depth above opaque water clouds. The 532nm layer integrated attenuated backscatter (γ′) for an opaque water cloud, assuming clear air-above, could be estimated as a function of the lidar ratio of the water cloud (S_0) and the multiple scattering factor. If there are features above, then γ′ will be reduced based on the overlaying two-way transmittance, which is a function of the optical depth (τ_{above}). Using this information an estimate of the above cloud optical depth can be inferred directly from current CALIPSO measurements, as seen in equation 1. The multiple scattering is a ratio of the layer integrated 532 nm volume depolarization ratio (δ_v), which along with γ′ are parameters that are already computed in the CALIOP L2 data products. An estimate of S_0 for opaque water clouds is assumed to be 19 sr.

$$\tau_{above} = -\frac{1}{2}\ln\left(2\left(\frac{1-\delta_v}{1+\delta_v}\right)^2 S_0\,\gamma'\right) \quad (1)$$

This new column particulate optical depth and uncertainty are detailed further in the V4.51 CALIOP Layer Level 2 data product description. These two new parameters are denoted as 'particulate', and do not differentiate between feature typing as seen in other column optical depth parameters in the CALIOP level 2 data product.

A comparison between this new approach and data collected from NASA Langley's High Spectral Resolution Lidar (HSRL-2) during a leg of the ORACLES-2 field campaign (Redemann et al., 2020) can be seen in Figure 4. This particularly scene, taken from September 18th, 2016, of a CALIPSO transit from 13:33:50 to 13:36: 52 UTC and

coincident with the ER-2 over-flight, is in the south-eastern Atlantic Ocean off the coast of Africa and is characterized by an elevated smoke layer from seasonal biomass burning over a dense marine stratiform cloud layer. The 532nm aerosol optical depth derived from HSRL-2 (black triangles) compares favorably with the new depolarization ratio approach (blue circles) at 5km resolution. The particulate optical depth is also reported at 1km and 333m (single shot) resolution. Profiles of 532nm extinction reported in the Lidar Level 2 5km aerosol profile products were used to compute aerosol optical depths for this scene, using both the V4.51 (red) and V4.20 (green, previous version) data. These two optical depths do not track the depolarization approach as well, given that the profiles of extinction are for aerosol only, as determined by the level 2 typing algorithm, while the depolarization method uses the entire column. The differences between V4.20 and V4.51 aerosol optical depths, in which V4.51 is always greater than/equal to V4.21, is due to the increase in depth of layer associated with smoke base extension.

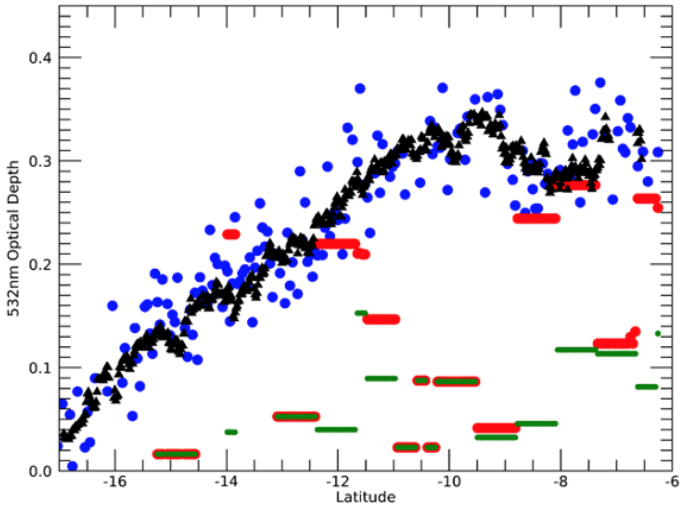


Figure 4: 532nm aerosol optical thickness above clouds from the HSRL-2 (black triangles), 532nm particulate optical depth above opaque water clouds at 5km resolution using the V4.51 Lidar Level 2 depolarization ratio technique (blue circles), and 532nm aerosol optical depth (V4.51 red circles, V4.20 green circles) computed by integration of aerosol profile extinction above opaque water clouds during CALIPSO over-flight of HSRL-2 track on September 18th, 2016, from 13:33:50 to 13:36:52 UTC in the south-east Atlantic Basin during the ORACLES-2 field campaign.

Total Column Particulate Optical Depths from Ocean Surface Returns

The Ocean Derived Column Optical Depth (ODCOD) product is an estimate of total column effective optical depth. ODCOD relates the measured magnitude of the lidar ocean surface return and the two-way transmittance overhead to a modelled ocean returned by the equation 2 below (Venkata and Reagan 2016).

$$T_p^2(z_s) = \frac{c\,A\,T_m^2(z_s)}{2\,R_\lambda} \tag{2}$$

In the relationship, $T^2(z_s)$ is the two-way transmittance at range z_s from the spacecraft where molecular/ozone and particulate components are denoted by the subscripts 'm' or 'p', respectively. Variable c is the speed of light, A is the area of the ocean surface return, and R_λ is the modelled backscatter reflectance from the ocean surface. The retrieval is for the total column, from the CALIOP top of the atmosphere calibration region to the lidar detected ocean surface. The estimate is an *effective* optical depth because no attempt is made to separate multiple scattering and single scattering from the cloud and aerosol particulates in the column (Ryan et al., 2024).

The V4.51 Lidar Level 2 algorithms only estimate 532 nm optical depths in regions of the column where clouds and aerosol are directly detected. The regions where no layers are detected are assumed to be particulate free. This assumption is known to yield an underestimate total column particulate optical depth by on the order of 0.03–0.05 (Toth et al., 2018). ODCOD compliments the CALIPSO data record by providing estimates of optical depth for the entire column including regions CALIOP does not detect particulate.

Figure 5 shows cloud screened seasonal difference of medians (ODCOD - MODIS) of the 5km ODCOD daytime retrieval to MODIS derived Effective Optical Depth Average Ocean interpolated to 532 nm and to the CALIOP 5km footprint midpoint. Generally, ODCOD shows good agreement in each season and over most of the globe with global median differences of 0.019 in December, January, and February; 0.025 in March, April, and May, 0.015 in June July and August; and 0.017 in September, October, November. Regionally, some parts of the globe see larger differences possibly due to uncertainties in wind speed and the surface reflectance model used.

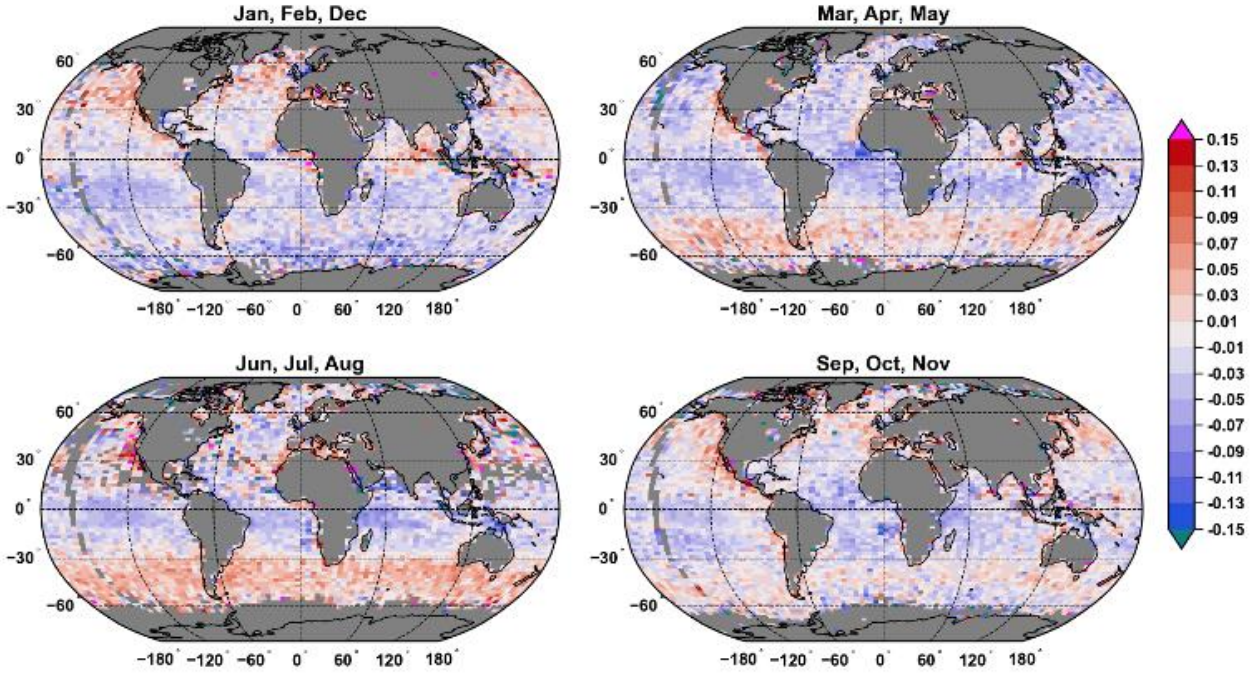


Figure 5: The four panels show cloud screened seasonal difference of medians (ODCOD - MODIS) comparisons of the 5km ODCOD daytime retrieval to MODIS derived Effective Optical Depth Average Ocean interpolated to 532 nm and to the CALIOP 5km footprint midpoint. OD ODCOD higher than MODIS is shown in red and ODCOD lower is shown in blue (Ryan et al., 2024 © Authors 2024, CC BY 4.0).

Corrections to Single-Shot Cloud Clearing

An anomalous condition was identified in the CALIOP Level 2 Selective Iterated Boundary Layer (SIBYL) algorithm, in which 333m single shot clouds identified and retained in the boundary layer could inadvertently be reclassified as an aerosol. The V4.51 layer product description details the error and new approach used, in which a boundary layer scene could be broken apart and these newly defined layers reclassified based on a number of criteria.

Figure 6 illustrates this error and the resultant amelioration. In this scene a dense water cloud underlies a dust plume (Figure 6a). For a majority of this scene the V4.2 algorithm retains the 333m single shot water cloud layer, which occurs when the single shot clouds are detected throughout the entirety of each 5km block of data. The Cloud-Aerosol Discrimination (CAD; Liu et al, 2019) doesn't know that the clouds haven't been removed and reclassifies those bins as having aerosols, as seen in Figure 6b. The consequence of this error can be seen in Figure 6c, as the 532nm 'aerosol' extinction for that layer are clearly biased when compared with the overlying aerosol. When applying the corrected algorithm, the single shot layer remains as a cloud (Figure 6d) and the 532nm aerosol extinction bias has been eliminated (Figure 6e). Tackett et al, 2022 also detailed the consequence of this error, in which nearly 3% of all 5km resolution layers for June-August 2007, 2013, and 2014, clustered around regions of continental aerosol outflow (smoke or desert dust), were impacted.

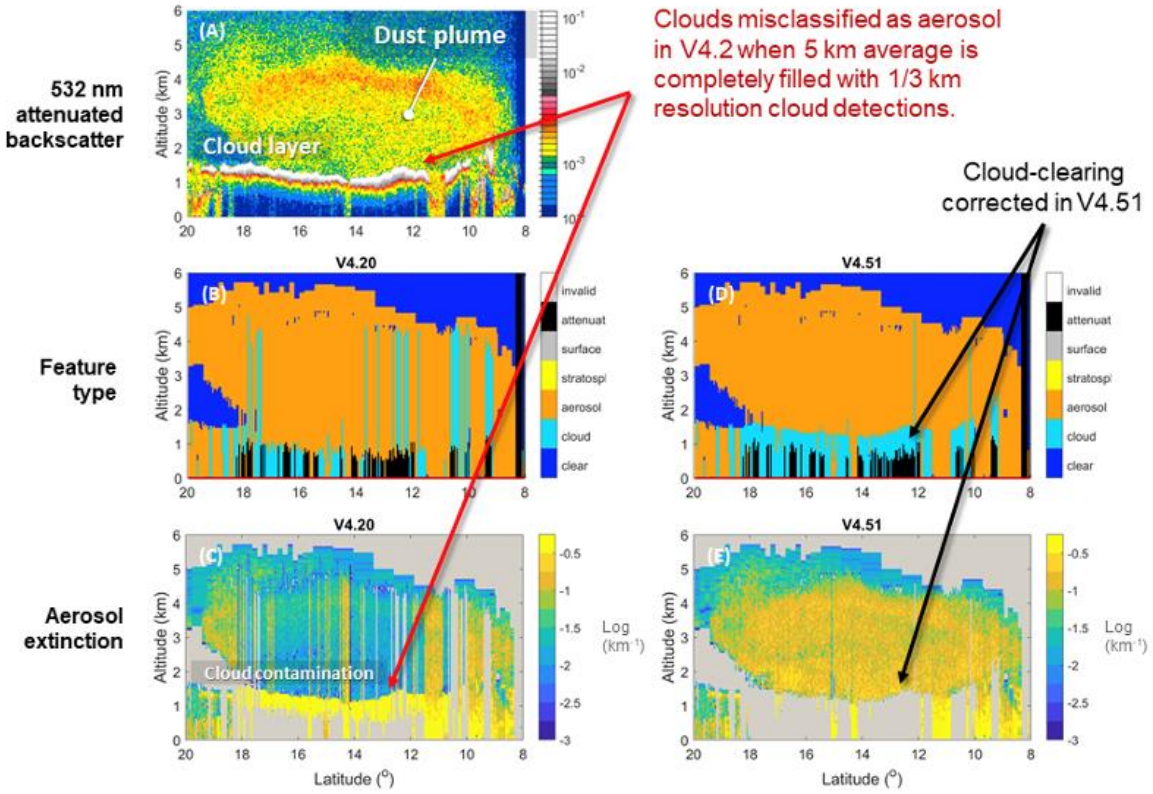


Figure 6: June 15th, 2013, scene in the eastern Atlantic Ocean between CALIOP V4.21 Level 1 and Level 2 data (a-c) and the V4.51 Level 2 data (d-e). The left most column shows the prescience of the low-level water cloud and the implications on extinction of the cloud layer being recast as an aerosol by the Cloud-Aerosol Discrimination algorithm. The right most column shows that the updated V4.51 algorithm now properly accounts for the cloud and the 532nm extinction bias has been eliminated. Taken from Tackett, 2023.

Table 1: new scientific data sets reported in the CALIOP V4.51 L2 data products

	VFM	Layer					Profile	
SDS name		**1/3kmM**	**1kmC**	**5kmA**	**5kmC**	**5kmM**	**5kmA**	**5kmC**
Surface_Wind_Speeds_02m				✓		✓	✓	✓

Data Quality Statement for CALIPSO's Version 4.21 Lidar Level 2 Data Product Release

Version: **4.21**
Data Release Date: **October 02, 2020**
Data Date Range: **October 01, 2020 to present**

A minor version bump (+0.01) has been applied to all CALIPSO data products due to a required upgrade to the operating system on the CALIPSO production cluster. All program executables were re-compiled to process in this new environment with no changes made to the underlying science algorithms or inputs.

Data Quality Statement for CALIPSO's Version 4.20 Lidar Level 2 Data Product Release

Data Version: **4.20**
Data Release Date: **October 10, 2018**
Data Date Range: **June 13, 2006 to September 30, 2020**

The Version 4.20 (V4) CALIOP Level 2 data product is identical to the V4.10 data product with the addition of several HDF science data sets (SDS) that provide information to the user for filtering out low laser energy shots. The technical advisory of this phenomena, shown below, was posted on the CALIPSO website in June of 2018.

Table 2: new scientific data sets reported in the CALIOP V4.20 L2 data products

	VFM	Layer				Profile	
SDS name		**1/3kmM**	**1kmC**	**5kmA**	**5kmC**	**5kmA**	**5kmC**
Minimum_Laser_Energy_532	✓	✓	✓	✓	✓	✓	✓

Data Release: June 12, 2018
CALIOP Instrument Anomaly Advisory: Low Energy Laser Shots
Dates Affected: September 01, 2016 – present
Applies to: CALIOP Level 1B and Level 2 Data Products, V4.10 and earlier

CALIOP (CALIPSO's LIDAR instrument) is experiencing an elevated frequency of low energy laser shots within the South Atlantic Anomaly (SAA) region due to decreased pressure inside the laser canister. The low energy laser shots began in September 2016 and have increased in frequency, particularly since the second half of 2017 (Fig. 1). Science quality of affected profiles within the SAA is degraded and these profiles should be excluded from scientific analyses. The document below provides guidance on how to identify affected profiles in CALIOP level 1B and level 2 products. Science quality of profiles with nominal laser energies is unaffected by this issue. The IIR (Imaging Infrared Radiometer) and the WFC (Wide Field Camera) instruments are also unaffected.

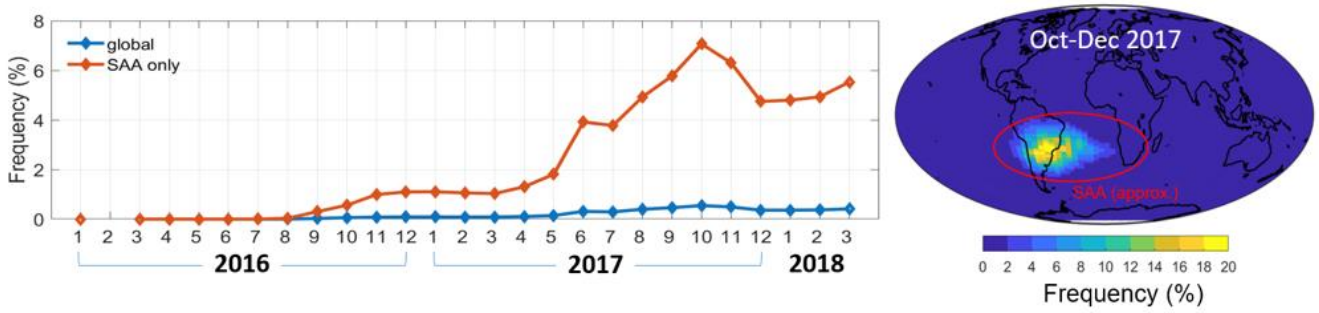


Figure 1 Monthly frequency of low energy laser shots (E_{532} < 80 mJ) from January 2016 to March 2018, globally and in the SAA region (left). Spatial distribution of low energy shot frequency from October to December 2017 (right).

Data Quality Statement for CALIPSO's Version 4.10 Lidar Level 2 Data Product Release

Data Version: **4.10**
Data Release Date: **November 8, 2016**
Data Date Range: **June 13, 2006 to May 31, 2018**

Version 4.10 (V4) is the first wholly new release of the CALIPSO lidar level 2 data products since the initial release of the Version 3 (V3) series of products in May 2010. As expected, V4.10 provides a substantial advance over V3 and earlier releases; known retrieval artifacts have been eliminated and numerous enhancements have been incorporated to increase the accuracy of the science data while simultaneously reducing uncertainties. The most significant code, algorithm, and data product changes include

- a stand-alone surface detection algorithm
- revised probability density functions (PDFs) for the cloud aerosol discrimination (CAD) algorithm
- application of the CAD algorithm to layers detected at single shot resolution and to layers detected in the stratosphere

- a major overhaul of the aerosol subtyping algorithms in which separate algorithms are now used to classify tropospheric and stratospheric aerosols
- improved cloud subtyping and ice-water phase determination
- temperature-dependent determination of multiple scattering factors for ice clouds
- multiple scattering factors for opaque water clouds derived from measured depolarization ratios
- updated extinction retrievals for opaque layers; improved uncertainty estimates and more in-depth quality assurance reporting for all extinction retrievals
- a new algorithm for deriving ice-water content from CALIOP extinction retrievals
- introduction of a new 5-km merged layer product that reports the spatial and optical properties of all cloud and aerosol layers detected in a single file
- new browses image showing cloud subtypes identified within each granule
- the addition of several new parameters in the data products files

Each of these is discussed in some detail in the sections below.

Lidar Surface Detection

In previous versions of the CALIOP level 2 (L2) data products, the altitude of the Earth's surface was determined using a general purpose layer detection scheme that scans lidar profiles from the top of the atmosphere downward, looking for significant positive excursions rising above an expected molecular backscatter signal. In general, this approach works well. However, in multi-layer scenes and/or highly turbid atmospheres the effectiveness of the top-down technique can be degraded by signal attenuation from intervening atmospheric layers that limit its ability to reliably detect surface returns. In the V4.10 data products, detection of the Earth's surface is accomplished using a dedicated, newly developed search routine that scans upward from the bottom of the profile using a derivative-based peak finding algorithm. This new technique demonstrates significant improvement over the V3 method in turbid atmospheres, while maintaining equal or better performance in clear skies. As a result of this improved detection scheme, there are fewer opaque layers identified in the V4.10 data than there were in V3, especially at night. Because regions below layers previously classified as opaque are now scanned for the presence of atmospheric features, there is also a slight increase in the number of cloud and aerosol layers reported. Signal strengths in these 'not previously scanned' regions tend to be quite low, and thus many of these newly detect layers will have low CAD scores.

The V4.10 data products report substantially more surface detection information than was available in V3. The Lidar_Surface_Elevation and Surface_Elevation_Detection_Frequency parameters reported in the V3 data products have been discontinued in V4.10. Instead, surface detection information is recorded in a multi-parameter Lidar_Surface_Detection Vgroup. In addition to the surface detection status (i.e., detected or not detected) and surface altitude information provided in V3, this new Vgroup reports the following parameters at both 532 nm and 1064 nm.

- surface top and base altitudes
- integrated attenuated backscatter
- integrated volume depolarization ratio
- integrated attenuated backscatter color ratio
- where applicable, surface detections at finer spatial resolutions (i.e., 1/3 km and 1 km)

These additional parameters are expected to provide new insights into surface type (e.g., distinguishing between ice, liquid water and land) and, for measurements over oceans, total column optical depths at both 532 nm and 1064 nm.

Cloud and Aerosol Layer Detection

The CALIOP V4.10 level 1 (L1) data, first released in April 2014, significantly improved the calibration of the CALIOP attenuated backscatter coefficients at 532 nm (Kar et al., 2018; Getzewich et al., 2018) and especially at 1064 nm (Vaughan et al., 2019). In particular, the magnitude of the calibration coefficients at 532 nm decreased by ~3% to ~8%, depending on latitude and season, resulting in the concomitant increase in the 532 nm attenuated backscatter coefficients. This increase in backscatter magnitude translates directly into a slightly greater layer detection frequency. When combined with the layer detection increases obtained from the new surface detection algorithm, the V4.10 data products show a net cloud fraction gain of ~5% relative to V3.

Cloud Aerosol Discrimination (CAD)

The magnitude of the calibration changes introduced in the CALIOP V4.10 L1 data necessitated substantial revisions to the V3 level 2 (L2) CAD algorithm. Accordingly, a new set of CAD probability distribution functions (PDFs) was developed and subsequently used to generate the V4.10 L2 data. The V4.10 CAD PDFs are still 5-dimensional, but now have increased latitude resolution (5° intervals vs. 10° in V3) which has led to an overall improvement in CAD reliability. The revised PDFs were specifically designed to be more sensitive to the presence of lofted aerosols. As a consequence, the V4.10 data products show significant improvements in the classification of high-altitude smoke plumes and Asian dust layers, which in earlier versions were often classified as cirrus clouds.

Application of the V4.10 CAD algorithm differs from previous versions in these important aspects:

1. In V3 and earlier, the CAD algorithm was applied only to tropospheric layers, and layers detected above the tropopause were classified as “stratospheric” features. In V4.10 the “stratospheric” feature type has been eliminated. Instead, the CAD algorithm is applied everywhere, to all layers detected. The CAD scores for stratospheric clouds and aerosols are generally robust within a few kilometers of the tropopause. However, at very high altitudes, the general paucity of samples available for the training set as well as falling SNR may affect the reliability of the CAD.

   For in-depth scientific analyses of polar stratospheric clouds (PSCs), users are strongly advised to use CALIPSO’s dedicated PSC products. For less demanding applications, PSCs are also reported in the standard L2 data products. While the V4.10 CAD algorithm classifies the majority of the polar stratospheric layers as clouds, some aerosol layers are also identified. The spatial distribution of these polar stratospheric aerosol layers is similar to the distribution of STS (i.e., the supercooled ternary solution of nitric acid, sulfuric acid and water) obtained from the dedicated CALIPSO PSC product.

2. Unlike V3, the V4.10 CAD algorithm is also applied to those strongly scattering layers that can be detected at single shot resolution (333 m). In the past these layers were classified as clouds by default and were systematically removed before averaging over the weaker signals. In V4.10, layers detected at single shot resolution that are classified as aerosols are no longer removed from coarser resolution averages, and thus can be expected to increase peak aerosol optical depths in the regions where they occur. The bulk of the single shot layers classified as aerosols in V4.10 are found within the dust belt region of the globe. However, it should be noted that optical properties of these layers were not used in building the V4.10 CAD PDFs, which may affect the overall CAD performance when classifying single shot layers.

While the CAD algorithm is applied to all layers detected, there are two anomalous situations where layers are subsequently reclassified using additional analysis. The first occurs when dense smoke plumes extend over stratus decks and other water clouds. The differential attenuation of the signals at 1064 nm and 532 nm by the smoke can lead to very high color ratios in the clouds below, which in turn can result in artificially low CAD scores. In such cases, the color ratio of the underlying clouds is reset to an empirically derived mean value and the CAD scores are recalculated. Both the original CAD score and the revised CAD score are recorded in the layer products. In the second case, the similarity of the scattering signatures of faint, isolated layers of lofted dust and weakly depolarizing cloud fragments makes them largely indistinguishable in the CAD domain. To achieve reliable separation, a sequence of spatial proximity tests is applied to identify those layers which may in fact be “fringes”

of previously identified large scale ice clouds. Layers identified as “cirrus fringes” are assigned a special CAD score of 106.

Figure 1 shows the pattern of changes in CAD scores from V3 to V4.10. Most of the high confidence samples in V3 are also classified as the same type (cloud or aerosol) in V4.10 with similar high confidence. However, as seen in the lower right quadrant of Figure 1, a small fraction of layers classified as clouds in V3 are classified as aerosols in V4.10. Some of these have optical properties that fall in the grey zone between aerosols and clouds and may actually be misclassified clouds. These cases occur most often over the polar regions. Because they typically have low CAD scores, they can be identified and removed at the users’ discretion.

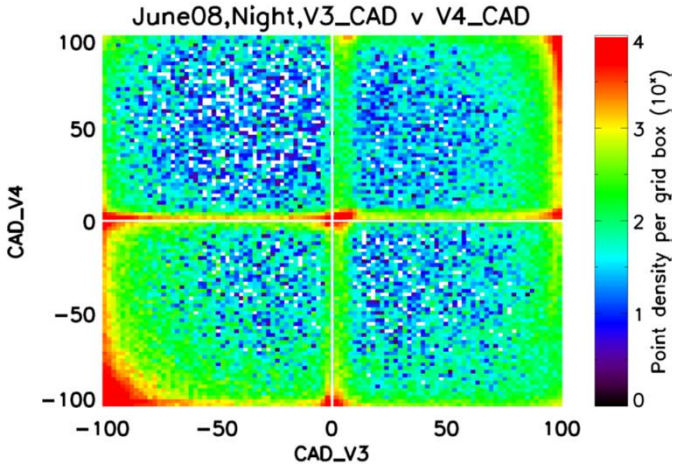


Figure 1: Comparison of CAD scores of the same samples between V3.30 and V4.10 using the L2 profile products.

After acquisition and analysis of over 10 years of space-based lidar data, the CALIPSO team has gained a much improved understanding of the physical and optical properties of the different types of aerosols and clouds that occur in different altitude regimes. As a result, the V4.10 PDFS are more representative and more physically realistic than earlier versions, in which the occurrence frequency of aerosols at higher altitudes was noticeably underestimated. One result of modifying the PDFs to achieve more accurate aerosol identification is an across the board decrease in the magnitude of the CAD scores reported in V4.10; i.e., the mean magnitude of CAD scores for all aerosols detected in V4.10 is lower than the mean magnitude in V3. Likewise, the mean magnitude of the cloud CAD scores is lower in V4.10 than in V3. In retrospect, the higher CAD scores in V3 should be seen as overly optimistic; the slightly lower V4.10 scores now provide a more realistic assessment of CAD classification confidence.

Aerosol Subtyping Changes

Several improvements to aerosol subtyping have been implemented in V4.10. The most fundamental change is that aerosol layers are now classified as either tropospheric aerosol or stratospheric aerosol feature types, depending on the location of the attenuated backscatter centroid relative to the MERRA2 reanalysis tropopause height. In previous versions, aerosol was only identified below the tropopause. Given that the CAD algorithm is applied at all altitudes in V4.10, aerosol layers detected above the tropopause are classified as stratospheric aerosols and are assigned subtypes commonly found in the stratosphere. Figure 2 compares distributions of tropospheric aerosol subtypes between V3 and V4.10.

Tropospheric aerosol subtyping improvements

1. A new “dusty marine” aerosol subtype has been added. Dusty marine layers are mixtures of dust and marine aerosol identified as moderately depolarizing aerosol layers having base altitudes within the marine boundary layer (assumed to be at 2.5 km). In previous versions these layers would have been classified as polluted dust. The dusty marine lidar ratio is more representative of a dust/marine aerosol mixture and its characteristic lidar ratio is ~33% smaller than that of polluted dust. The geographic distribution of dusty marine layers agrees well with known locations of dust subsidence into the marine boundary layer, though some ambiguity occurs in

regions where anthropogenic pollution, dust, and marine aerosol co-exist. These layers are classified as dusty marine, yet it is not always clear whether they should be typed instead as polluted dust.

2. Smoke layer identification and nomenclature has been revised. As in previous versions, elevated non-depolarizing aerosols are assumed to be smoke which is injected above the planetary boundary layer (PBL) due to combustion-induced buoyancy. The definition for "elevated" is revised in V4.10 to mean layers with tops higher than 2.5 km above ground level (i.e., a simple PBL approximation). For clarity, the nomenclature for the smoke aerosol subtype is changed to "elevated smoke". Owing to the revised "elevated" definition and the introduction of a new algorithm to vertically homogenize aerosol subtyping for weakly scattering fringes detected at the base of extended plumes, elevated smoke layers which were misclassified as marine aerosol in V3 are now correctly classified as elevated smoke.
3. Within the PBL, it is difficult to discriminate smoke due to biomass burning from polluted continental aerosol arising from anthropogenic pollution using CALIOP measurements. Therefore, the description of the polluted continental subtype is revised to "polluted continental/smoke" to clarify that either aerosol type could be present.
4. In previous versions, aerosol detected over snow, ice, or tundra were subtyped as either clean continental or polluted continental. Given that transport pathways exist for smoke, dust and other aerosol types to reach the Arctic, this condition has been removed in V4.10 and all species are allowed. Users are cautioned to treat aerosol detected over Antarctica carefully and to exercise prudence when interpreting aerosol subtyping in this region. Often aerosol layers in the Antarctic are classified as dust or polluted dust due to their elevated depolarization. Despite that transport pathways do exist for dust to reach Antarctica (from Patagonia for example), data users are cautioned that layers classified as aerosol may actually be misclassified clouds or blowing snow rather than true dust.
5. Calculation of the particulate depolarization ratio estimates using in the aerosol subtyping scheme now correctly accounts for signal attenuation due to overlying layers. Making this change greatly reduces the over-abundance of polluted dust in identified in V3.
6. Tropospheric aerosol lidar ratios and lidar ratio uncertainties have been updated for the marine, dust, clean continental, and elevated smoke subtypes to reflect the current state of knowledge based on observations by NASA Langley Airborne High Spectral Resolution Lidar, EARLINET, AERONET, CALIPSO and synergistic multi-sensor retrievals.

Stratospheric aerosol subtypes introduced

Stratospheric aerosol subtypes have been introduced in V4.10 for ash, sulfate/other, smoke and polar stratospheric aerosol. The V4.10 stratospheric aerosol subtyping algorithm performs well at identifying volcanic ash and sulfate above the tropopause based on manual verification. Note that below the tropopause, ash and sulfate plumes are given tropospheric aerosol subtypes: volcanic ash is often classified as dust or polluted dust and volcanic sulfate is often classified as elevated smoke. As a result, contiguous aerosol features crossing the tropopause will have aerosol subtypes which switch from tropospheric to stratospheric subtypes, depending on the relationship between the attenuated backscatter centroid altitude of the layer identified by the feature finder and the tropopause altitude. Weakly scattering stratospheric aerosol layers which are not classified as polar stratospheric aerosol are classified as "sulfate/other". Therefore, layers that are, in fact, ash and/or smoke could be misclassified as "sulfate/other" if they are weakly scattering (layer integrated attenuated backscatter less than 0.001 sr^{-1}).

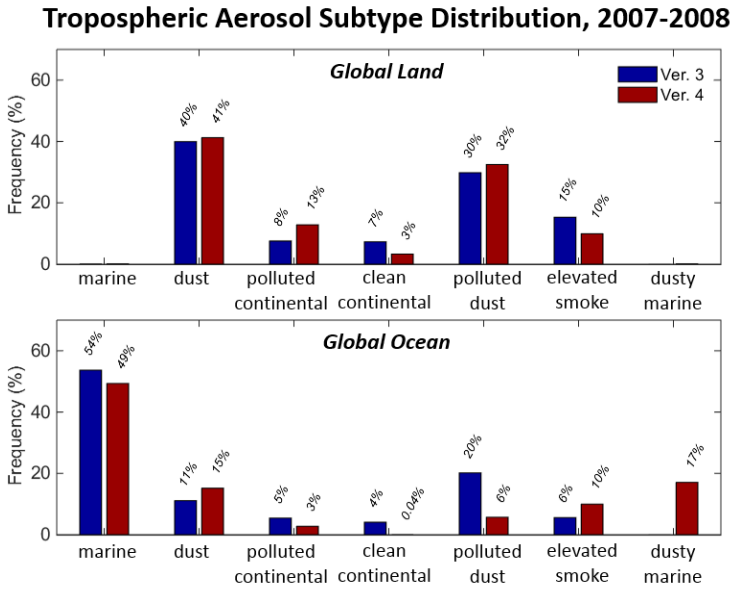


Figure 2: Comparison of global tropospheric aerosol subtype distributions between V3 and V4.10 for 2007-2008, day and night. Over land, elevated smoke is reduced in favor of polluted continental/smoke as a consequence of revised "elevated" definition. Over ocean, dusty marine replaces a substantial portion of polluted dust classifications.

Cloud Subtyping Changes

The CALIPSO cloud subtyping algorithm uses cloud top pressure, cloud opacity and cloud fraction to identify eight cloud types. A bug in the V3 analysis code caused a systematic underestimate of all categories of opaque clouds; in fact, no low overcast opaque clouds were reported in any of the V3 data products. As seen in Figure 3, this defect has been remedied in V4.10, and thus, relative to V3, the V4.10 data products show a large increase in the fraction of low opaque cloud types and a corresponding decrease in the fraction of low transparent clouds.

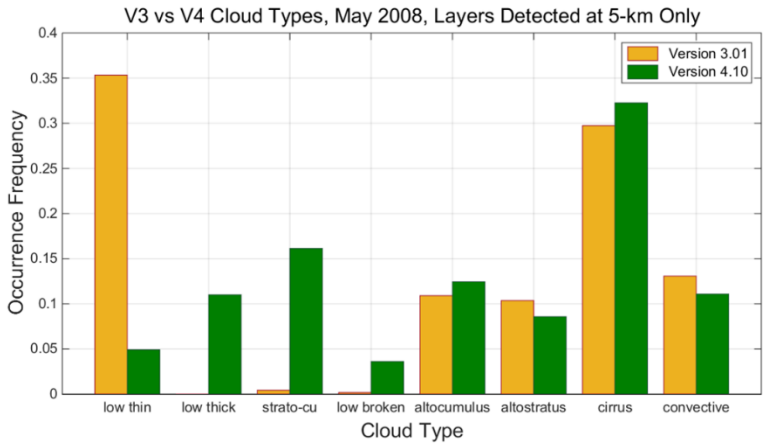


Figure 3: distribution of cloud subtypes in V3 (orange) and V4.10 (green) for all layers detected at a 5-km horizontal averaging resolution during May 2008

Cloud Ice-Water Phase Discrimination Changes

Layers identified as clouds by the CAD algorithm are further classified according to thermodynamic phase as either water, randomly oriented ice (ROI), horizontally oriented ice (HOI) or unknown phase. The phase algorithm primarily uses an objective sorting algorithm based on the clustering of relationships between the layer-integrated attenuated backscatter (IAB) and the layer-averaged depolarization. This technique relies on fundamental CALIOP

L1 measurements, and in V4.10 confidently identifies the thermodynamic phase (ice vs. water) of at least 75% (nadir pointing) and 85% (tilted) of cloud layers, globally.

On November 28, 2007, the initial CALIOP viewing angle of 0.3° (nadir) was permanently changed to 3° (tilted), to repress specular reflections from hexagonal plates. These plates, with crystal faces perpendicular to the CALIOP laser beam, cause specular reflections which can be identified in nadir viewing data by abnormally large integrated attenuated backscatter with essentially zero depolarization. The V3 phase algorithm included a scheme for recognizing HOI that identified numerous instances of these ice clouds. However, in-depth comparisons of the V3 nadir and tilted data determined that very few true specular reflections were occurring in the tilted data, and so HOI testing of ice clouds initially identified as ROI was eliminated. In the V4.10 phase algorithm only water clouds observed in the nadir view are tested. Clouds that mainly consist of ROI may also have water or HOI occurring at warmer temperatures at the bottom of the layer, and these are now identified by the dominant cloud particle phase, which is ROI. Additional details about the change of off-nadir angle from 0.3° to 3.0° are given in Appendix 1 at the end of this document. Table 1 characterizes V3 to V4.10 changes in the volume of cloud phases globally, for both nadir and tilted viewing angles.

Table 1: Cloud volume occurrence frequency (5 km x 60 m bins) in percent; nadir statistics computed using all data from January through November 2007 (excluding off-nadir tests); tilted statistics computed using all data from January through November 2008 (MC = mid-confidence).

Cloud Phase	Phase Flag	Confidence Flag	V3 Nadir (0.3°)	V4.10 Nadir (0.3°)	V3 Tilted (3.0°)	V4.10 Tilted (3.0°)
ROI	1	3	48	53	64	67
Water	2	3	18	20	18	19
HOI	3	3	6	5	< 1	< 1
Unknown	0	0	9	13	6	10
Fringe	1	0	N/A	3	N/A	3
MC-HOI	3	2	17	5	10	N/A
MC-ROI	1	2	3	< 1	< 1	< 1

A non-zero, but negligible, amount of low- and mid-confidence water layers are also identified at each viewing angle. As these phase classifications account for less than 1% of all cloud bins, they are not shown in this table.

In V4.10, between 65-70% of the range bins identified as atmospheric features by CALIOP are classified as clouds. The population of water clouds identified at horizontal averages of 5 km or more remains very stable between V3 and V4.10, at about 18%. The ROI population is larger in tilted data than in nadir data and is 10-15% larger in V4.10 than in V3. The additional V4.10 ROI bins were mainly classified as clear air, mid-confidence HOI or stratospheric features in V3. Unknown phase clouds increase in V4.10 due to generally lower CAD scores and the detection of more thin cloud layers with weak backscatter and depolarization signals. The reduction in V4.10 HOI is due to the elimination of a spatial coherence test in the phase algorithm. About 3% of the cloud population in V4.10 are identified as “cirrus fringes”. Since the composition of cloud populations varies regionally and seasonally, these numbers should be used only for guidance in understanding the changes between V3 and V4.10.

Lidar Ratios and Multiple Scattering Factors for Ice Clouds

In V3 and earlier, ice clouds were assigned a constant multiple scattering factor of $\eta_{532} = 0.6$. In V4.10, the multiple scattering factor is instead implemented as a sigmoid approximation function of the layer attenuated backscatter centroid temperature, with η_{532} increasing from 0.46 at 270 K to 0.76 at 190 K.

This approximation function was derived from extensive analysis of collocated measurements acquired by the CALIPSO lidar and the CALIPSO IIR, which reconciled observed and theoretical ratios of 532 nm optical depths derived from V3 CALIOP measured two-way transmittances to the absorption optical depth retrieved from IIR

measurements at 12.05 μm. The theoretical ratios are computed assuming severely roughened aggregated columns.

In V3 and earlier, ice cloud extinction retrievals that could not be constrained by the direct measurement of the two-way transmittance (i.e., "unconstrained" retrievals) were assigned an initial default lidar ratio of 25 sr. For semi-transparent clouds, comparisons with IIR absorption optical depth at 12.05 μm and a radiative closure experiment using MODIS 11 μm radiances both showed, on average, quite good agreement with V3 CALIOP constrained retrievals, but substantially worse agreement with unconstrained retrievals, thus demonstrating that the initial default lidar ratio was generally too small.

To ensure full consistency between unconstrained and constrained retrievals in semi-transparent clouds, initial ice cloud lidar ratios in V4.10 are derived from the statistical analysis of several years of constrained retrievals, using only those clouds identified as high-confidence randomly oriented ice. Like the multiple scattering factor, the V4.10 initial ice cloud lidar ratio is estimated using a sigmoid approximation function based on the layer attenuated backscatter centroid temperature. Default lidar ratio values decrease from ~35 sr to ~20 sr as the cloud centroid temperature decreases. This initial lidar ratio is only used for semi-transparent ice clouds when constrained retrievals are not possible. For opaque clouds and constrained retrievals of semi-transparent clouds, the extinction retrievals are initialized using a lidar ratio derived directly from the CALIOP L1 measurements and the temperature-dependent multiple scattering factor.

Multiple Scattering Factors for Water Clouds

In V3 and earlier, all water clouds were assigned a constant multiple scattering factor of $\eta_{532} = 0.6$. In V4.10, η_{532} for transparent water clouds remains fixed at 0.6. For opaque layers, however, layer-effective water cloud multiple scattering factors are computed from the measured layer integrated volume depolarization ratios. As a consequence, under the appropriate conditions (e.g., single layer clouds in otherwise clear skies), estimates of water cloud lidar ratios can now be obtained from the multiple scattering factors and the 532 nm layer integrated attenuated backscatter estimates.

Extinction and Optical Depths

The particulate backscatter and extinction profiles and layer optical depths reported in the CALIOP V4.10 data products are produced by a modified and substantially enhanced version of the hybrid extinction retrieval algorithm used in earlier releases. Several developments are particularly noteworthy.

- Analysis of Opaque Layers

  The extinction retrieval used for opaque layers is entirely different. In V3 and earlier, the lidar ratios used for opaque layers were assigned by the scene classification algorithms based on layer type (i.e., cloud vs. aerosol) and subtype (e.g., ice vs. water, dust vs. smoke, etc.). In V4.10, initial estimates of the lidar ratios for opaque layers are computed directly from the measured integrated attenuated backscatter, and refined as necessary within the extinction solver to ensure that extinction coefficients are calculated through the full vertical extent of the layer. This procedure yields highly precise and accurate layer-effective lidar ratios, which translate directly into more realistic extinction coefficient estimates and eliminate many artifacts previously seen in CALIOP optical depth distributions.

- Increased Number of Constrained Retrievals

  Constrained retrievals use measurements of clear air above and below a lofted layer to directly estimate layer optical depth. These optical depths provide a constraint on the solution of the lidar equation, allowing the layer lidar ratio to be retrieved from the data rather than estimated a priori. In retrospect, the approach used in V3 and earlier was perhaps too timid, in that constrained solutions were only attempted for lofted layers with optical depths greater than ~0.3. In V4.10 constrained solutions (and hence lidar ratio estimates) are derived for all layers having valid two-way transmittance measurements. Uncertainties in the lidar ratio estimates are now reported for all retrievals. While lidar ratios derived from layers with small optical depths can have large random uncertainties, the extinction retrievals are unbiased.

- More Precise Reporting of Retrieval Success

  For each extinction profile retrieved, information about the termination state of the extinction algorithm is provided in the extinction QC flags. These flags are implemented as 16-bit unsigned integers. Multiple bits can be toggled within each extinction QC flag, with each bit conveying a specific piece of information. In V3, 10 of the 16 bits were used. The V4.10 algorithm uses 14 bits and is both more verbose and more rigorous in its assessment of retrieval quality. As a consequence, users will encounter many more distinct QC flags in V4.10 than in V3. For example, for all data collected in 2008 V3 reported only 19 different extinction QC flags at 532 nm. By contrast, V4.10 is expected to report in the neighborhood of 75 different values for the same time period. The most reliable retrievals have extinction QC flags of 0, 1, 2, 16 or 18. Data flagged with other values should be treated with varying degrees of suspicion. In aberrant cases, the extinction retrieval can fail. The backscatter and extinction coefficients reported for these failed retrievals are set to a fill value of -333.

- Improved Estimates of Extinction Uncertainties

  During the V4.10 development, considerable attention was given to providing more accurate estimates of the uncertainties reported for the CALIOP extinction coefficients and optical depths. In particular, V4.10 lidar ratio uncertainty estimates are now verified and adjusted as appropriate on a layer-by-layer basis. In the V3 processing these uncertainty estimates were always specified a priori and never varied thereafter.

  While layer-effective multiple scattering factors for opaque water clouds can be reliably estimated, the known range dependence of water cloud multiple scattering is not accounted for in the V4.10 CALIOP retrieval algorithm. Furthermore, the V4.10 extinction retrieval does not attempt to compensate for the loss of ranging information introduced by pulse stretching. As a result, beyond the first range bin (and frequently within the first range bin) the V4.10 CALIOP extinction retrievals in opaque water clouds should be considered entirely unreliable. To reinforce this notion, the uncertainties for opaque water clouds are not calculated, but instead are assigned a uniform fill value of -29.

- Summary of Extinction Changes from V3 to V4.10

  The changes described above will have considerable impact on the magnitude of the V4.10 backscatter and extinction coefficients, their attendant uncertainties, and in those parameters subsequently derived from these values. However, it is critically important for data users to understand that the extinction and optical depth changes from V3 to V4.10 cannot be attributed wholly to changes in the extinction algorithm. Changes in the a priori specifications of layer multiple scattering factors and/or layer lidar ratios can by themselves introduce considerable changes in the retrieved values of extinction and optical depth. Similarly, changes in the calibration coefficients from V3 to V4.10 result in small increases in the L1 attenuated backscatter coefficients, which in turn yield concomitant but nonlinear increases in particulate backscatter and extinction coefficients.

Revised Ice-Water Content Algorithm

Cloud ice water content (IWC) is reported for all ice clouds detected by CALIOP. As in V3, IWC is a provisional data product calculated as a parameterized function of the CALIOP 532 nm extinction coefficients retrieved within ice clouds. In V4.10, the parameterization has been modified to include a temperature-dependent particle size relationship that approximates the observations compiled in an expanded set of aircraft microphysical data. Due to the cumulative impact of multiple factors, including the new ice mass parameterization and improved V4.10 calibration and extinction coefficient retrievals, users can expect V4.10 IWC to be significantly larger than V3 IWC; e.g., up to 6-8 times as large for thick ice clouds at warm temperatures. At cold temperatures, the V4.10 IWC calculated for a given extinction coefficient is smaller than in V3. However, since the extinction coefficients are larger in V4.10, the resulting change in IWC from V3 to V4.10 is relatively small.

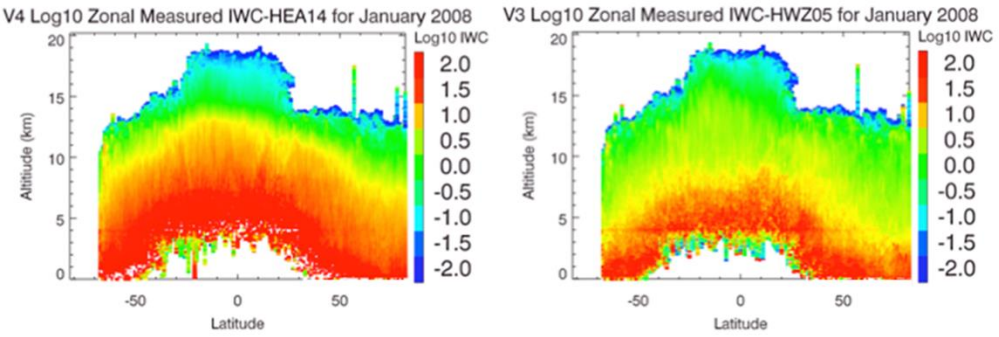


Figure 4: zonal mean IWC for January 2008 calculated from V4.10 data (left panel) and V3.01 data (right panel).

CALIOP IWC is a highly derived data product. Besides the cloud particle area-to-mass parameterization, it relies on cloud feature determination (CAD), cloud phase determination, specification of lidar ratios and multiple scattering factors and the ensuing extinction retrieval. Uncertainty in cloud ice water content is affected by the accuracy of the microphysical parameterization. However, the IWC uncertainty reported in the CALIOP data products reflects only the uncertainty in the extinction coefficient retrieval. Because IWC is parameterized from ice particle extinction, it follows that data screening criteria for valid IWC should be similar to those used to identify valid extinction coefficients. Best results will likely be obtained by using only those data with high CAD scores that have also been classified as ice with high confidence.

New Data Product: 5-km Merged Layer Product

In response to numerous end-user requests, the V4.10 data release includes a new 5-km merged layer product that aggregates all of the information found in the existing 5-km cloud layer product and 5-km aerosol layer product and packages it into a single file. The 5-km merged layer product also contains a comprehensive subset of the data reported in the single shot layer product (as do the V4.10 5-km cloud and aerosol layer products), so that unambiguous cloud clearing information will always be immediately available. This new product offers several advantages to users of the CALIPSO layer products. In particular, (a) the spatial relationships between clouds and aerosols detected at varying averaging resolution in any column are fully specified and (b) the optical influences between layers of different types (e.g., the uncertainties in cloud optical depth retrievals for cirrus clouds lying above aerosol layers) can be readily appreciated and fully characterized. A complete specification of all parameters included in the 5-km merged layer product is given in the latest release of the CALIPSO Data Products Catalog.

Browse Image Improvements

The CALIOP browse images have been augmented in the V4.10 release with new plots showing the cloud subtypes identified within each granule. The aerosol subtyping plots have been updated to reflect changes in aerosol subtyping and the addition of the stratospheric aerosol subtypes.

New Version 4.10 Data File Parameters

Table 2 lists all new parameters that have been added to the V4.10 data products. Of particular note is the inclusion of single shot layer detection information in all 5-km layer products.

Table 2: new scientific data sets reported in the CALIOP V4.10 L2 data products

	VFM	Layer				Profile	
SDS name		**1/3kmM**	**1kmC**	**5kmA**	**5kmC**	**5kmA**	**5kmC**
Lidar Surface Detection (Vgroup)		✓	✓	✓	✓	✓	✓
Single Shot Detection (Vgroup)	✓			✓	✓		
Profile ID	✓					✓	✓
CAD Score		✓					
“Was Cleared” Flag		✓					
Opacity Flag		✓	✓				

	VFM	Layer				Profile	
SDS name		1/3kmM	1kmC	5kmA	5kmC	5kmA	5kmC
Layer Centroid Temperature		✓	✓		✓		
Initial CAD Score		✓	✓		✓		
Attenuated Scattering Ratio Statistics, 532 nm		✓	✓	✓	✓		
High Resolution Layers Cleared				✓	✓		
Final Lidar Ratio Uncertainty, 532 nm				✓	✓		
Final Lidar Ratio Uncertainty, 1064 nm					✓		
Ozone Number Density						✓	✓
IGBP Surface Type						✓	✓

Data Quality Statement for CALIPSO's Version 3.41 Lidar Level 2 Data Product Release

Data Version: **3.41**
Data Release Date: **October 02, 2020**
Data Date Range: **October 01, 2020 to present**

Version 3.41 marks a change in the data products due to a required upgrade of the operating system on the cluster computer used to generate CALIPSO's publicly distributed data products. All program executables were re-compiled to process in this new environment with no changes made to the underlying science algorithms or inputs.

Data Quality Statement for CALIPSO's Version 3.40 Lidar Level 2 Data Product Release

Data Version: **3.40**
Data Release Date: **December 14, 2016**
Data Date Range: **December 01, 2016 to September 30, 2020**

Version 3.40 release reflects an update of the Forward Processing – Instrument Teams (FP-IT) meteorological data provided by the Global Modeling and Assimilation Office (GMAO) from version 5.9.1 to version 5.12.4. No changes were made to the program executables.

Data Quality Statement for CALIPSO's Version 3.30 Lidar Level 2 Data Product Release

Data Version: **3.30**
Data Release Date: **April 19, 2013**
Data Date Range: **March 1, 2013 to November 30, 2016**

Version 3.30 data products incorporate new versions of two important ancillary data products: (1) updated GMAO FP-IT meteorological data and (2) the enhanced Air Force Weather Authority (AFWA) snow and ice data set as ancillary inputs. Based on comparisons of CALIOP V3.02 data to the newly generated CALIOP V3.30 products, the transition to the new GEOS-5 FP-IT data is predicted to have only minimal effects on the science data products. Production of the V3.30 data set began with March 1, 2013.

Layer Detection

GEOS-5 molecular number densities increased in the CALIOP night and day calibration regions, which lowered the calibration coefficients and increased the attenuated backscatter coefficients, which in turn caused the number of layers detected to increase slightly. A two month of V3.02 to V3.30 showed that the number of aerosol and cloud layers increased by < 0.8% and < 0.2%, respectively.

Layer Classification

The GEOS-5 tropopause heights decreased by ~1 km between 30°S and 40°N. Since CALIOP classifies layers detected above the tropopause as stratospheric features, 3%–5% of the features formerly classified as stratospheric were reclassified as either cloud or aerosol. Over Antarctic in September 2011, tropopause heights decreased between 1.0 km and 1.5 km, causing all of cloud and aerosol layers with bases above the revised tropopause heights to be reclassified as stratospheric features. This effect may occur seasonally over Antarctica.

Data Quality Statement for CALIPSO's Version 3.02 Lidar Level 2 Data Product Release

Data Version: **3.02**
Data Release Date: **December 12, 2011**
Data Date Range: **November 1, 2011 to February 28, 2013**

Version 3.02 represents a transition of the Lidar, IIR, and WFC science data processing and browse image production to a new cluster computing system. No algorithm changes were introduced, and only very minor changes were observed between V3.01 and V3.02 as a result of the compiler and computer architecture differences.

Data Quality Statement for CALIPSO's Version 3.01 Lidar Level 2 Data Product Release

Data Version: **3.01**
Data Release Date: **April 28, 2010**
Data Date Range: **June 13, 2006 to October 31, 2011**

Version 3.01 of the Lidar Level 2 data products is a significant improvement over the previous version. Major code and algorithm improvements include:

- the elimination of a bug in the cloud clearing code that caused a substantial overestimate of low cloud fraction in earlier data releases (details given in Vaughan et al., 2010);
- enhancements to the cloud-aerosol discrimination algorithm that increase the number of diagnostic parameters used to make classification decisions (details given in Liu et al., 2010);
- improved daytime calibration procedures, resulting in more accurate estimates of layer spatial and optical properties (details given in Powell et al., 2010); and
- an entirely new algorithm for assessing cloud thermodynamic phase (details given in Hu et al., 2009).

Layer Detection

As in previous versions, the layer boundaries reported in the Lidar Level 2 Cloud and Aerosol Layer Products appear to be quite accurate. Some false positives are still found beneath optically thick layers; these, however, can generally be identified by their very low CAD scores (e.g., |CAD score| ≤ 20). In opaque 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. Opaque layers are denoted by an opacity flag. 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

Figure 1 (below) compares the distributions of CAD scores derived from four months of version 3 test data to the corresponding version 2.01 data. The V3 curve shows a smoother distribution and generally has fewer low CAD values (i.e., values less than ~|95|), reflecting the better separation of clouds and aerosols when using the version 3 5-D PDFs as compared to the separation provided by 3-D PDFs in previous versions. One notable exception to this observation is the bump between -10 and 20 in the V3 test curve, which accounts for ~ 6% of the total features.

The CAD scores in this region identify both outlier features whose optical/physical properties are not correctly measured or derived, and those features whose attributes fall within the overlap region between the cloud and aerosol PDFs. In contrast, these outliers are populated over the entire CAD span in the V2 release.

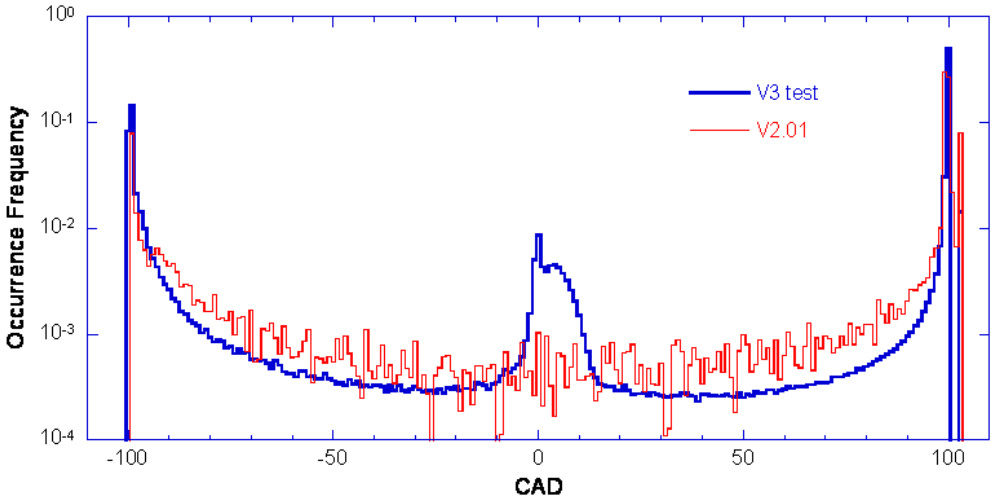


Figure 1: Histograms of CAD scores for Version 2 (red) and Version 3 (blue).

Figure 2 (below) presents the relationship between the CAD score and the layer IAB QA factor, which provides a measure of the integrated attenuated backscatter overlying a cloud or an aerosol layer. A layer IAB QA factor close to 1 indicates that the atmosphere above the layer under is clear. Decreasing values indicate the increasing likelihood of overlying layers that have attenuated the signal within the layer under consideration, and thus decreased the SNR of the measurement. A layer IAB QA factor of 0 would indicate total attenuation of the signal. As seen in the figure, the IAB QA is highest for high magnitude CAD scores and slopes down gradually for small CAD score magnitudes. This relationship reflects the fact that the presence of overlying features tends to add difficulty to the cloud-aerosol classification task, and therefore reduces the confidence of the classifications made. The dip between -10 and 20 represents features that are outliers in the 5-D CAD PDFs and indicates that these outliers most often lie beneath other relatively dense features. The cloud layers with special CAD scores (103 and 104) have the smallest IAB QA values. The relatively big value at CAD = 0 corresponds to the features having zero CAD values at high altitudes where the probability of the presence of overlying features is low. At high altitudes the separation of clouds and aerosols is not as good as at low altitudes because of the presence of subvisible cirrus clouds.

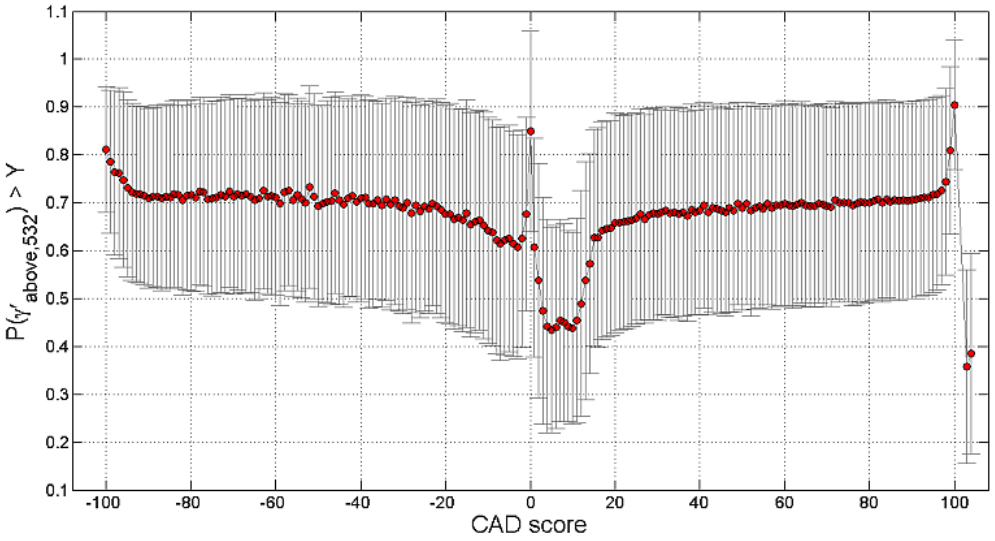


Figure 2: Relationship between CAD score and Layer IAB QA Factor

Overall, because of the better separation between clouds and aerosols in the 5D space, the 5D CAD algorithm significantly improves the reliability of the CAD scores. These improvements include:

1. Dense aerosol layers (primarily very dense dust and smoke over and close to the source regions), which are sometimes labeled as cloud in the V2 release, are now correctly identified as aerosol, largely because of the addition of the integrated volume depolarization ratio to the diagnostic parameters used for cloud-aerosol discrimination. In addition, in the open oceans, dense aerosols that were previously classified as clouds are now frequently observed in the marine boundary layer. Improvements are also seen for these maritime aerosols. Note, however, dense dust/smoke layers found at single-shot (0.333 km) resolution will be classified as cloud by default. This issue will be revisited for post-V3 releases.
2. Because the V2 CAD algorithm used a latitude-independent set of 3D PDFs, a class of optically thin clouds encountered in the polar regions that can extend from the surface to several kilometers were sometimes misclassified as aerosols. In version 3, these features are now correctly classified as cloud.
3. Correct classification of heterogeneous layers is always difficult. 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 in the V2 release. By convention, heterogeneous layers should be classified as clouds. The version 3 feature finding algorithm has also been improved greatly and can now much better separate the embedded or adjacent single-shot cloud layers from the surrounding aerosol. This improvement in layer detection contributes significantly to the improvement of the CAD performance.
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 overestimate of the attenuation due to overlying layers. These erroneous "pseudo-features" are neither cloud nor aerosol and are distributed outside of the cloud and aerosol clusters in the PDF space. The V3 CAD algorithm can better identify these outlier features by assigning a small CAD score (the bump between -10 and 20 in the V3 CAD histogram) and classify most of them as cloud by convention. A CAD threshold of 20 can effectively filter out these outliers.

Some misclassifications may still occur with the 5D algorithm. For example, dust aerosols can be transported long distance to the Arctic. When moderately dense dust layers are occasionally transported to high latitudes, where cirrus clouds can present even in the low altitudes, they may be misclassified. This is also the case for moderately dense smoke aerosols occasionally transported to the high latitudes. Smoke can be mixed with ice particles during the long range transport, which makes the smoke identification even more difficult. When moderately dense dust and smoke are transported vertically to high altitudes, even at low latitudes, misclassifications can occur due to the presence of cirrus clouds. Volcanic aerosol that is newly injected into the high altitudes may have a large cross-polarized backscatter signal and thus may be misclassified as cloud.

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

The cloud phase algorithm used in Version 2 has been replaced with a new, completely different approach. The V3 algorithm classifies detected cloud layers as water, randomly-oriented ice (ROI), or horizontally-oriented ice (HOI) based on relations between depolarization, backscatter, and color ratio (Hu et al. 2009). These classifications have not yet been rigorously validated, which is difficult, but many of the obvious artifacts found in the V2 data have been eliminated.

The V2 algorithm included a rudimentary ability to identify a specific subset of high confidence instances of HOI. These clouds were classified as ice clouds and flagged with a 'special CAD score' of 102, indicating that they had been further classified as HOI. The new version 3 algorithm implements a much more sophisticated scheme for recognizing HOI that correctly identifies many more instances of these clouds. The special CAD score of 102 is no longer used to identify these layers. Instead, the "ice cloud" and "mixed phase cloud" classifications have been eliminated and replaced as shown in Table 1 below. Associated QC flags are given in Table 2.

Table 1: Updated Ice/Water Phase Flags

Value	V2 Interpretation	V3 Interpretation
0	unknown/not determined	unknown/not determined
1	ice	randomly oriented ice (ROI)
2	water	water
3	mixed phase	horizontally oriented ice (HOI)

Table 2: Updated Ice/Water Phase QC Flags

Value	V2 Interpretation	V3 Interpretation
0	no confidence	no/low confidence
1	low confidence	phase based on temperature only
2	medium confidence	medium confidence
3	high confidence	high confidence

A confidence flag of QA=1 indicates the phase classification is based on temperature. Initial classification tests are based on layer depolarization, layer-integrated backscatter, and layer-average attenuated backscatter color ratio. Layers classified as water with temperatures less than -40 C are forced to ROI and given confidence flags of QA=1. Layers classified as ROI or HOI with temperatures greater than 0 C are forced to water and also given confidence flags of QA=1. Clouds for which the phase is 'unknown/not determined' are assigned confidence values of 0 (no/low confidence).

Layers classified as HOI based on anomalously high backscatter and low depolarization are assigned QA=3. These layer characteristics are rarely detected after the CALIOP viewing angle was changed to 3° in November 2007. The V3 algorithm computes the spatial correlation of depolarization and integrated backscatter and uses this as an additional test of cloud phase. Layers classified as HOI using this test are assigned QA=2. The spatial correlation test is responsible for the majority of the layers classified as HOI. These layers typically have higher backscatter than ROI but similar depolarization and are common even at a viewing angle of 3°. We interpret this as clouds with

significant perpendicular backscatter from ROI but containing enough HOI to produce enhanced backscatter. These layers tend to be found at much colder temperatures than the high confidence HOI (see Hu et al. 2009).

Cloud and Aerosol Optical Depths

The reliability of cloud and aerosol optical depths reported in the V3 data products is considerably improved over the V2 release. Whereas the V2 optical depths were designated as a beta quality product, and not yet suitable for use in scientific publications, the maturity level of the V3 optical depths has been upgraded to provisional. Several algorithm improvements and bugs fixes factored into the decision to upgrade the maturity level. Among these were the addition of the aerosol layer base extension algorithm, which greatly improves AOD estimates in the planetary boundary layer (PBL), and several significant improvements to the code responsible for rescaling the attenuated backscatter coefficients in lower layers to compensate for the beam attenuation that occurs when traversing transparent upper layers.

Data Quality Statement for CALIPSO's Version 2.02 Lidar Level 2 Data Product Release

Data Version: **2.02**
Data Release Date: **October 2008**
Data Date Range: **September 14, 2008 to October 29, 2009**

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 Cabanes 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 in 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 V2.01 code versus 9676 layers detected by the V2.02 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, 2006 test case, corrections to the interpolation algorithms applied to the GMAO data result in a slight upward shift in the tropopause heights and hence more clouds and fewer stratospheric layers are identified in the V2.02 results.

Cloud Ice/Water Phase Discrimination

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 between V2.02 and V2.01 increase with increasing penetration depths, and can grow large when optical depths are large (i.e., > 3).

Data Quality Statement for CALIPSO's Version 2.01 Lidar Level 2 Data Product Release

Data Version: **2.01**
Data Release Date: **January 2008**
Data Date Range: **June 13, 2006 to September 13, 2008**

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 are still occurred with some frequency. Among these, the most prevalent are:

- 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 PDF overlap region. 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.
- 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.
- 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.

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 overestimate 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).

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. 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.

Aerosol Subtype Identification

The main objective of the aerosol subtyping scheme is to estimate the appropriate value of the aerosol extinction-to-backscatter ratio (Sa) to within 30% of the true value. Sa is an important parameter used in the determination of the aerosol extinction and subsequently the optical depth from CALIOP backscatter measurements. Sa 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 Sa 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 Sa 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 Sa . 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 Sa 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 and aerosol extinction profiles

The CALIOP cloud and aerosol profiles of extinction and backscatter are released as beta products. Cloud profiles are reported at a horizontal resolution of 5 km; aerosol profiles are reported at a horizontal resolution of 40 km. These products contain a number of known errors and, in their current form, cannot be used as standalone products. The current products contain no data quality information, and hence must be used in conjunction with the Cloud and/or Aerosol Layer Products and/or the Vertical Feature Mask Product, which contain data quality parameters and confidence flags. Data assessment and screening procedures have not yet been developed. Because of this, the profile data product is considered to be not appropriate for scientific publication but is released to users for evaluation 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 appropriate 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.

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 V2.01 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.

Data Quality Statement for CALIPSO's Version 1.10 Lidar Level 2 Data Product Release

Data Version: **1.10**
Data Release Date: **December 8, 2006**
Data Date Range: **June 13, 2006 to November 11, 2007**

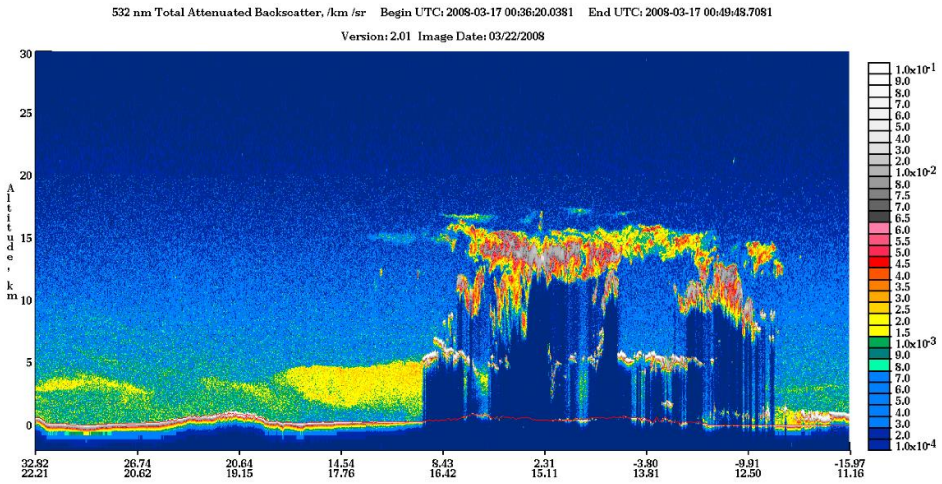
The CALIPSO vertical feature mask (VFM) data product reports a single 16-bit integer for each lidar altitude resolution element in the data stream downlinked from the satellite. Upon decoding each of these bit-mapped integers, users will obtain information describing layer location (both vertically and horizontally), layer type, and the amount of horizontal averaging required for the layer to be detected.

Given the accuracy of the CALIPSO altitude registration, the layer heights reported in the Lidar Level 2 Cloud and Aerosol Layer Products and the VFM appear to be quite accurate. In optically dense layers, the lowest altitude where signal is observed is reported as the base. In actuality, this point 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.

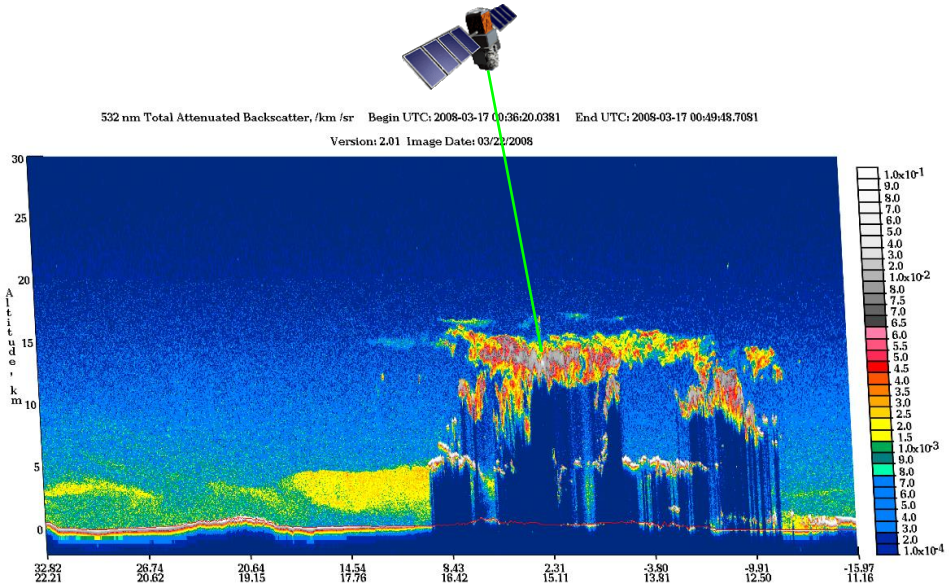
A preliminary version of the algorithm to discriminate cloud and aerosol has been used in this release. Overall, the algorithm performance is fairly good at labeling cloud as cloud and somewhat less successful in labeling aerosol as aerosol. Several types of misclassifications are fairly common and should be watched for. The most common misclassification is portions of dense aerosol layers being labeled as cloud. The algorithm operates on individual profiles, so small regions within an aerosol layer are sometimes labeled as cloud. These misclassifications are often apparent from study of Level 1 browse images. Actual clouds occurring within aerosol layers appear to be correctly classified as cloud most of the time. Additionally, portions of the bases of some cirrus clouds are mislabeled as aerosol, and some tropospheric polar clouds are erroneously labeled as aerosol. Improvements to the cloud/aerosol discrimination algorithm are underway and misclassifications should be greatly reduced in future data releases.

Appendix 1

Prior to November 28, 2007, CALIPSO was nominally pointed in a 'near nadir' direction (actually at ~0.3° off nadir, to avoid the full force of specular reflections from still waters and horizontally oriented ice crystals). In that pointing configuration, the CALIOP browse images – as shown below – correctly represented both the spatial distribution of clouds and aerosols, and the optical effects of overlying layers on the atmosphere below.



As a result of the change to pointing 3° off-nadir, with the tilt being in the same plane as the satellite velocity vector (i.e., along track, rather than cross track), the spatial distribution of layers is not exactly what is shown in the standard browse images. A more faithful representation might look like this…



There are several aspects of this change that could be of interest to data users:

1. While there is now a 3° vertical shear in the spatial distribution of clouds and aerosols (i.e., as shown above), the optical effects of overlying layers are still correctly represented by the original, upright images that are being shown on our web pages.
2. The altitudes reported in all products are still measured with respect to a nadir viewing instrument, and thus users need not (and should not) make any corrections for pointing angle.
3. The laser footprint locations reported in the data products give the latitude and longitude coordinates of the beam location on the Earth at mean sea level.
4. The horizontal offset between any two points in the vertical can be computed using 'triangle trig'. If the vertical distance is V, the horizontal offset, Δh, between the upper and lower points is $\Delta h = V \cdot \tan(3°)$.
5. Version 2 of the data products includes a new 'Spacecraft_Position' SDS that contains all of the spacecraft attitude information required to vertically collocate the CALIPSO and CloudSat profiles. This was included at the request of the CloudSat team, specifically for the continued production of their GeoProf products.

6. In general, the optical properties reported for measurements of aerosols and water clouds are not expected to change as a function of the change in pointing angle. However, the properties reported for individual ice clouds will change by some varying amount, depending on the concentration of horizontally aligned ice crystals present in the cloud. Among the changes that can be anticipated are
    - a reduction in the maximum values of integrated attenuated backscatter measured at both wavelengths;
    - an increase in the minimum depolarization ratios associated with strongly scattering ice clouds;
    - a small change in the proportion of clouds classified as ice versus those classified as water; and
    - an increase in the minimum lidar ratios retrieved for strongly scattering ice clouds.

References

Avery, M. A., R. A. Ryan, B. J. Getzewich, M. A. Vaughan, D. M. Winker, Y. Hu, A. Garnier, J. Pelon, and C. A. Verhappen, 2020: CALIOP V4 Cloud Thermodynamic Phase Assignment and the Impact of Near-Nadir Viewing Angles, *Atmos. Meas. Tech.*, **13**, 4539–4563, https://doi.org/10.5194/amt-13-4539-2020.

Burton, S. P., R. A. Ferrare, M. A. Vaughan, A. H. Omar, R. R. Rogers, C. A. Hostetler, and J. W. Hair, 2013: Aerosol Classification from Airborne HSRL and Comparisons with the CALIPSO Vertical Feature Mask, *Atmos. Meas. Tech.*, **6**, 1397–1412, https://doi.org/10.5194/amt-6-1397-2013.

Chand, D., T. L. Anderson, R. Wood, R. J. Charlson, Y. Hu, Z. Liu, and M. Vaughan, 2008: Quantifying above-cloud aerosol using spaceborne lidar for improved understanding of cloudy-sky direct climate forcing, *J. Geophys. Res.*, **113**, D13206, https://doi.org/10.1029/2007JD009433.

Dai, G., S. Wu and X. Song, 2018: Depolarization Ratio Profiles Calibration and Observations of Aerosol and Cloud in the Tibetan Plateau Based on Polarization Raman Lidar, *Remote Sensing*, **10**, 378, https://doi.org/10.3390/rs10030378.

Garnier, A., J. Pelon, M. A. Vaughan, D. M. Winker, C. R. Trepte, and P. Dubuisson, 2015: Lidar multiple scattering factors inferred from CALIPSO lidar and IIR retrievals of semi-transparent cirrus cloud optical depths over oceans, *Atmos. Meas. Tech.*, **8**, 2759–2774, https://doi.org/10.5194/amt-8-2759-2015.

Garnier, A., J. Pelon, N. Pascal, M. A. Vaughan, P. Dubuisson, P. Yang and D. L. Mitchell, 2021: Version 4 CALIPSO Imaging Infrared Radiometer ice and liquid water cloud microphysical properties – Part I: The retrieval algorithms, *Atmos. Meas. Tech.*, **14**, 3253–3276, https://doi.org/10.5194/amt-14-3253-2021.

Haarig, M., R. Engelmann, A. Ansmann, I. Veselovskii, D. N. Whiteman and D. Althausen, 2016: 1064 nm rotational Raman lidar for particle extinction and lidar-ratio profiling: cirrus case study, *Atmos. Meas. Tech.*, **9**, 4269–4278, https://doi.org/10.5194/amt-9-4269-2016.

Haarig, M., R. Engelmann, H. Baars, B. Gast, D. Althausen and A. Ansmann, 2025: Discussion of the spectral slope of the lidar ratio between 355 nm and 1064 nm from multiwavelength Raman lidar observations, *Atmos. Chem. Phys.*, **25**, 7741–7763, https://doi.org/10.5194/acp-25-7741-2025.

Haarig, M., R. Engelmann, H. Baars, B. Gast, D. Althausen and A. Ansmann, 2025: Discussion of the spectral slope of the lidar ratio between 355 nm and 1064 nm from multiwavelength Raman lidar observations, *Atmos. Chem. Phys.*, **25**, 7741–7763, https://doi.org/10.5194/acp-25-7741-2025.

Heymsfield, A. J., D. Winker, and G.-J. van Zadelhoff, 2005: Extinction-ice water content-effective radius algorithms for CALIPSO, *Geophys. Res. Lett.*, **32**, L10807, https://doi.org/10.1029/2005GL022742.

Heymsfield, A., D. Winker, M. Avery, M. Vaughan, G. Diskin, M. Deng, V. Mitev, and R. Matthey, 2014: Relationships between Ice Water Content and Volume Extinction Coefficient from In Situ Observations for Temperatures from 0° to -86°C: Implications for Spaceborne Lidar Retrievals, *J. Appl. Meteor. Climatol.*, **53**, 479–505, https://doi.org/10.1175/JAMC-D-13-087.1.

Holz, R. E., S. Platnick, K. Meyer, M. Vaughan, A. Heidinger, P. Yang, G. Wind, S. Dutcher, S. Ackerman, N. Amarasinghe, F. Nagle, and C. Wang, 2016: Resolving ice cloud optical thickness biases between CALIOP and MODIS using infrared retrievals, *Atmos. Chem. Phys.*, **16**, 5075–5090, https://doi.org/10.5194/acp-16-5075-2016.

Hu, Y. M. Vaughan, Z. Liu, B. Lin, P. Yang, D. Flittner, W. Hunt, R. Kuehn, J. Huang, D. Wu, S. Rodier, K. Powell, C. Trepte, and D. Winker, 2007: The depolarization-attenuated backscatter relation: CALIPSO lidar measurements vs. theory, *Opt. Express*, **15**, 5327–5332, https://doi.org/10.1364/OE.15.005327.

Hu, Y., M. Vaughan, Z. Liu, K. Powell, and S. Rodier, 2007: Retrieving Optical Depths and Lidar Ratios for Transparent Layers Above Opaque Water Clouds From CALIPSO Lidar Measurements, *IEEE Geosci. Remote Sens. Lett.*, **4**, 523–526, https://doi.org/10.1109/LGRS.2007.901085.

Hunt, W. H, D. M. Winker, M. A. Vaughan, K. A. Powell, P. L. Lucker, and C. Weimer, 2009: CALIPSO Lidar Description and Performance Assessment, *J. Atmos. Oceanic Technol.*, **26**, 1214–1228, https://doi.org/10.1175/2009JTECHA1223.1.

Josset, D., J. Pelon, Y. Hu, R. Rogers, Z. Liu, A. Omar, M. Vaughan, P.-W. Zhai, and the ICARE team, 2012: Global Scale Lidar Ratio Retrieval Over the Ocean, 26th International Laser Radar Conference (ILRC), 25–29 June 2012, Porto Heli, Greece, https://www.researchgate.net/publication/281802691_Global_scale_lidar_ratio_retrieval_over_the_ocean

Kahnert, M. and R. Scheirer, 2019: "Multiple scattering by aerosols as seen from CALIPSO — a Monte-Carlo modelling study", *Opt. Express*, **27**, 33683–33699, https://doi.org/10.1364/OE.27.033683.

Kar, J., K.-P. Lee, M. A. Vaughan, J. L. Tackett, C. R. Trepte, D. M. Winker, P. L. Lucker, and B. J. Getzewich, 2019: CALIPSO Level 3 Stratospheric Aerosol Product: Version 1.00 Algorithm Description and Initial Assessment, *Atmos. Meas. Tech.*, **12**, 6173–6191, https://doi.org/10.5194/amt-12-6173-2019.

Kim, Y., S.-W. Kim, M.-H. Kim and S.-C. Yoon, 2014: Geometric and optical properties of cirrus clouds inferred from three-year ground-based lidar and CALIOP measurements over Seoul, Korea, *Atmos. Res.*, **139**, 27–35, https://doi.org/10.1016/j.atmosres.2013.12.016.

Kim, M.-H., A. H. Omar, J. L. Tackett, M. A. Vaughan, D. M. Winker, C. R. Trepte, Y. Hu, Z. Liu, L. R. Poole, M. C. Pitts, J. Kar, and B. E. Magill, 2018: The CALIPSO Version 4 Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *Atmos. Meas. Tech.*, **11**, 6107-6135, https://doi.org/10.5194/amt-11-6107-2018.

Liu, Z., R. Kuehn, M. Vaughan, D. Winker, A. Omar, K. Powell, C. Trepte, Y. Hu, and C. Hostetler, 2010: The CALIPSO Cloud And Aerosol Discrimination: Version 3 Algorithm and Test Results, 25th International Laser Radar Conference (ILRC), 5–9 July 2010, St. Petersburg, Russia; see https://www.researchgate.net/publication/285685007.

Liu, Z., D. Winker, A. Omar, M. Vaughan, C. Trepte, Y. Hu, K. Powell, W. Sun, B. Lin, 2011: Effective lidar ratios of dense dust layers over North Africa derived from the CALIOP measurements, *JQSRT*, **112**, 204–213, https://doi.org/10.1016/j.jqsrt.2010.05.006.

Liu, Z., J. Kar, S. Zeng, J. Tackett, M. Vaughan, M. Avery, J. Pelon, B. Getzewich, K.-P. Lee, B. Magill, A. Omar, P. Lucker, C. Trepte, and D. Winker, 2019: Discriminating Between Clouds and Aerosols in the CALIOP Version 4.1 Data Products, *Atmos. Meas. Tech.*, **12**, 703–734, https://doi.org/10.5194/amt-12-703-2019.

Lopes, F. J. S., E. Landulfo, and M. A. Vaughan, 2013: Evaluating CALIPSO's 532 nm lidar ratio selection algorithm using AERONET sun photometers in Brazil, *Atmos. Meas. Tech.*, **6**, 3281–3299, https://doi.org/10.5194/amt-6-3281-2013.

Lu, X., Y. Hu, Z. Liu, S. Rodier, M. Vaughan, C. Trepte, and J. Pelon, 2017: Observations of Arctic snow and sea ice cover from CALIOP lidar measurements, *Remote Sens. Environ.*, **194**, 248–263, https://doi.org/10.1016/j.rse.2017.03.046.

Lu, Z., X. Liu, Z. Zhang, C. Zhao, K. Meyer, C. Rajapakshe, C. Wu, Z. Yang and J. E. Penner, 2018: Biomass smoke from southern Africa can significantly enhance the brightness of stratocumulus over the southeastern Atlantic Ocean, *PNAS*, **115**, 2924–2929, https://doi.org/10.1073/pnas.1713703115.

Miller, S. D., and G. L. Stephens (1999), Multiple scattering effects in the lidar pulse stretching problem, *J. Geophys. Res.*, 104(D18), 22205–22219, https://doi.org/10.1029/1999JD900481.

Mioche, G., D. Josset, J.-F. Gayet, J. Pelon, A. Garnier, A. Minikin, and A. Schwarzenboeck, 2010: Validation of the CALIPSO-CALIOP extinction coefficients from in situ observations in midlatitude cirrus clouds during the CIRCLE-2 experiment, *J. Geophys. Res.*, **115**, D00H25, https://doi.org/10.1029/2009JD012376.

Müller, D., A. Ansmann, I. Mattis, M. Tesche, U. Wandinger, D. Althausen, and G. Pisani, 2007: Aerosol-type-dependent lidar ratios observed with Raman lidar, *J. Geophys. Res.*, **112**, D16202, https://doi.org/10.1029/2006JD008292.

O'Connor, E. J., A. J. Illingworth and R. J. Hogan, 2004: A Technique for Autocalibration of Cloud Lidar, *J. Atmos. Oceanic Technol*., **21**, 777–786, https://doi.org/10.1175/1520-0426(2004)021<0777:ATFAOC>2.0.CO;2.

Omar, A. H., J.-G. Won, D. M. Winker, S.-C. Yoon, O. Dubovik, and M. P. McCormick, 2005: Development of global aerosol models using cluster analysis of Aerosol Robotic Network (AERONET) measurements, *J. Geophys. Res.*, **110**, D10S14, https://doi.org/10.1029/2004JD004874.

Omar, A., Z. Liu, M. Vaughan, K. L. Thornhill, C. Kittaka, S. Ismail, Y. Hu, G. Chen, D. Winker, C. Trepte, E. L. Winstead, and B. E Anderson, 2010: Extinction-to-backscatter ratios of Saharan dust layers derived from in situ measurements and CALIPSO overflights during NAMMA, *J. Geophys. Res.*, **115**, D24217, https://doi.org/10.1029/2010JD014223.

Omar, A., D. Winker, C. Kittaka, M. Vaughan, Z. Liu, Y. Hu, C. Trepte, R. Rogers, R. Ferrare, R. Kuehn, and C. Hostetler, 2009: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *J. Atmos. Oceanic Technol*., **26**, 1994–2014, https://doi.org/10.1175/2009JTECHA1231.1.

Papagiannopoulos, N., L. Mona, L. Alados-Arboledas, V. Amiridis, H. Baars, I. Binietoglou, D. Bortoli, G. D'Amico, A. Giunta, J. L. Guerrero-Rascado, A. Schwarz, S. Pereira, N. Spinelli, U. Wandinger, X. Wang and G. Pappalardo, 2016: CALIPSO climatological products: evaluation and suggestions from EARLINET, *Atmos. Chem. Phys.*, **16**, 2341–2357, https://doi.org/10.5194/acp-16-2341-2016.

Pauly, R. M., J. E. Yorks, D. L. Hlavka, M. J. McGill, V. Amiridis, S. P. Palm, S. D. Rodier, M. A. Vaughan, P. A. Selmer, A. W. Kupchock, H. Baars, and A. Gialitaki, 2019: Cloud Aerosol Transport System (CATS) 1064 nm Calibration and Validation, *Atmos. Meas. Tech.*, **12**, 6241–6258, https://doi.org/10.5194/amt-12-6241-2019.

Platt, C. M. R., 1973: Lidar and radiometric observations of cirrus clouds, *J. Atmos. Sci.*, **30**, 1191–1204, https://doi.org/10.1175/1520-0469(1973)030<1191:LAROOC>2.0.CO;2.

Platt, C. M. R., D. M. Winker, M. A. Vaughan and S. D. Miller, 1999: Backscatter-to-Extinction Ratios in the Top Layers of Tropical Mesoscale Convective Systems and in Isolated Cirrus from LITE Observations, *J. Appl. Meteor. Climatol.*, **38**, 1330–1345, https://doi.org/10.1175/1520-0450(1999)038<1330:BTERIT>2.0.CO;2.

Powell, K. A., C. A. Hostetler, Z. Liu, M. A. Vaughan, R. E. Kuehn, W. H. Hunt, K. Lee, C. R. Trepte, R. R. Rogers, S. A. Young, and D. M. Winker, 2009: CALIPSO Lidar Calibration Algorithms: Part I - Nighttime 532 nm Parallel Channel and 532 nm Perpendicular Channel, *J. Atmos. Oceanic Technol*., **26**, 2015–2033, https://doi.org/10.1175/2009JTECHA1242.1.

Powell, K. A., M. A. Vaughan, R. R. Rogers, R. E. Kuehn, W. H. Hunt, K-P. Lee, and T. D. Murray, 2010: The CALIOP 532-nm Channel Daytime Calibration: Version 3 Algorithm, 25th International Laser Radar Conference (ILRC), 5–9 July 2010, St. Petersburg, Russia; see https://www.researchgate.net/publication/391011856 : The CALIOP 532-nm Channel Daytime Calibration Version 3 Algorithm.

Rajapakshe, C., Z. Zhang, J. E. Yorks, H. Yu, Q. Tan, K. Meyer, S. Platnick and D. M. Winker, 2017: Seasonally Transported Aerosol Layers over Southeast Atlantic are Closer to Underlying Clouds than Previously Reported, *Geophys. Res. Lett.*, **44**, https://doi.org/10.1002/2017GL073559.

Redemann, J. and 68 coauthors, 2021: An overview of the ORACLES (ObseRvations of Aerosols above CLouds and their intEractionS) project: aerosol–cloud–radiation interactions in the southeast Atlantic basin, *Atmos. Chem. Phys.*, **21**, 1507–1563, https://doi.org/10.5194/acp-21-1507-2021.

Rodriguez, J. V., R. C. Verhappen, C. Weimer, C. R. Trepte and T. E. Cayton, 2022: Charged Particle Fluxes Associated with CALIPSO Low Laser Energy Shots, *IEEE T. Nucl. Sci.*, **69**, 2146–2153, https://doi.org/10.1109/TNS.2022.3204715.

Rogers, R. R., M. A. Vaughan, C. A. Hostetler, S. P. Burton, R. A. Ferrare, S. A. Young, J. W. Hair, M. D. Obland, D. B. Harper, A. L. Cook, and D. M. Winker, 2014: Looking Through the Haze: Evaluating the CALIPSO Level 2 Aerosol Optical Depth using Airborne High Spectral Resolution Lidar Data, *Atmos. Meas. Tech.*, **7**, 4317–4340, https://doi.org/10.5194/amt-7-4317-2014.

Ryan, R. A., M. A. Vaughan, S. D. Rodier, J. L. Tackett, J. A. Reagan, R. A. Ferrare, J. W. Hair, and B. J. Getzewich, 2024: Total Column Optical Depths Retrieved from CALIPSO Lidar Ocean Surface Backscatter, *Atmos. Meas. Tech.*, **17**, 6517–6545, https://doi.org/10.5194/amt-17-6517-2024.

Sayer, A. M., A. Smirnov, N. C. Hsu, and B. N. Holben, 2012: A pure marine aerosol model, for use in remote sensing applications, *J. Geophys. Res.*, **117**, D05213, https://doi.org/10.1029/2011JD016689.

Shcherbakov, V., F. Szczap, A. Alkasem, G. Mioche and C. Cornet, 2022: Empirical model of multiple-scattering effect on single-wavelength lidar data of aerosols and clouds, *Atmos. Meas. Tech.*, **15**, 1729–1754, https://doi.org/10.5194/amt-15-1729-2022.

Tackett, J. L., M. A. Vaughan, K.-P. A. Lee, Kam-Pui J. Kar, and C. R. Trepte, 2021: Improvements in CALIOP Smoke Optical Depth over Clouds, 11th Symposium on Lidar Atmospheric Applications, 101st Annual AMS Meeting 2021, held online, 10-15 January 2021, https://ntrs.nasa.gov/citations/20205011703.

Tackett, J. M. Vaughan, J. Lambeth, and A. Garnier, 2022: Critical Improvements to CALIOP Boundary Layer Cloud-Clearing in Version 4.5, 2023 American Meteorological Society Annual Meeting, 8–12 January 2023, Denver, CO, https://ntrs.nasa.gov/citations/20220019202.

Tackett, J. L., J. Kar, M. A. Vaughan, B. Getzewich, M.-H. Kim, J.-P. Vernier, A. H. Omar, B. Magill, M. C. Pitts, and D. Winker, 2023: The CALIPSO version 4.5 stratospheric aerosol subtyping algorithm, *Atmos. Meas. Tech.*, **16**, 745–768, https://doi.org/10.5194/amt-16-745-2023.

Tackett, J. L., R. A. Ryan, A. E. Garnier, J. Kar, B. Getzewich, X. Cai, M. A. Vaughan, C. R. Trepte, R. Verhappen, D. M. Winker and K.-P. A. Lee, 2025: Mitigating Impacts of Low Energy Laser Pulses on CALIOP Data Products, *EGUsphere* [AMTD], https://doi.org/10.5194/egusphere-2025-2376, accepted for publication.

Tesche, M., A. Ansmann, D. Müller, D. Althausen, I. Mattis, B. Heese, V. Freudenthaler, M. Wiegner, M. Esselborn, G. Pisani, and P. Knippertz, 2009: Vertical profiling of Saharan dust with Raman lidars and airborne HSRL in southern Morocco during SAMUM, *Tellus B*, **61**, 144–164, https://doi.org/10.1111/j.1600-0889.2008.00390.x.

Tesche, M., U. Wandinger, A. Ansmann, D. Althausen, D. Müller and A. H. Omar, 2013: Ground-based validation of CALIPSO observations of dust and smoke in the Cape Verde region, *J. Geophys. Res.*, **118**, 2889–2902, https://doi.org/10.1002/jgrd.50248.

Thorsen, T. J., Q. Fu, J. M. Comstock, C. Sivaraman, M. A. Vaughan, D. M. Winker, and D. D. Turner, 2013: Macrophysical properties of tropical cirrus clouds from the CALIPSO satellite and from ground-based micropulse and Raman lidars, *J. Geophys. Res.*, **118**, 9209–9220, https://doi.org/10.1002/jgrd.50691.

Toth, T. D., J. Zhang, J. R. Campbell, J. S. Reid, Y. Shi, R. S. Johnson, A. Smirnov, M. A. Vaughan, and D. M. Winker, 2013: Investigating Enhanced Aqua MODIS Aerosol Optical Depth Retrievals over the Mid-to-High Latitude

Southern Oceans through Intercomparison with Co-Located CALIOP, MAN, and AERONET Datasets, *J. Geophys. Res.*, **118**, 4700–4714, https://doi.org/10.1002/jgrd.50311.

Toth, T. D., J. R. Campbell, J. S. Reid, J. L. Tackett, M. A. Vaughan, J. Zhang, and J. W. Marquis, 2018: Minimum Aerosol Layer Detection Sensitivities and their Subsequent Impacts on Aerosol Optical Thickness Retrievals in CALIPSO Level 2 Data Products, *Atmos. Meas. Tech.*, **11**, 499–514, https://doi.org/10.5194/amt-11-499-2018.

Toth, T. D., M. Clayton, Z. Li, D. Painemal, S. Rodier, J. Kar, T. Thorsen, R. Ferrare, M. Vaughan, J. Tackett, H. Bian, M. Chin, A. Garnier, E. Welton, R. Ryan, C. Trepte and D. Winker, 2025: Mapping CALIPSO Marine and Dusty Marine Aerosol Lidar Ratios using MODIS AOD Constrained Retrievals and GOCART Model Simulations, *EGUsphere* [AMTD], https://doi.org/10.5194/egusphere-2025-2832, accepted for publication.

Vaughan, M., K. Powell, R. Kuehn, S. Young, D. Winker, C. Hostetler, W. Hunt, Z. Liu, M. McGill, and B. Getzewich, 2009: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, *J. Atmos. Oceanic Technol.*, **26**, 2034–2050, https://doi.org/10.1175/2009JTECHA1228.1.

Vaughan, M. A., Z. Liu, M. J. McGill, Y. Hu, and M. D. Obland, 2010: On the Spectral Dependence of Backscatter from Cirrus Clouds: Assessing CALIOP's 1064 nm Calibration Assumptions Using Cloud Physics Lidar Measurements, *J. Geophys. Res.*, **115**, D14206, https://doi.org/10.1029/2009JD013086.

Vaughan, M., R. Kuehn, J. Tackett, R. Rogers, Z. Liu, A. Omar, B. Getzewich, K. Powell, Y. Hu, S. Young, M. Avery, D. Winker, and C. Trepte, 2010: Strategies for Improved CALIPSO Aerosol Optical Depth Estimates, , 25th International Laser Radar Conference (ILRC), 5–9 July 2010, St. Petersburg, Russia, https://ntrs.nasa.gov/citations/20100026019.

Vaughan, M., Z. Liu, Y. Hu, K. Powell, A. Omar, S. Rodier, W. Hunt, J. Kar, J. Tackett, B. Getzewich, and K.-P. Lee, 2016: Cloud-Aerosol Interactions: Retrieving Aerosol Ångström Exponents from CALIPSO Measurements of Opaque Water Clouds, 27th International Laser Radar Conference, 5–10 July 2015, New York, New York USA, EPJ Web of Conferences, **119**, 11001, https://doi.org/10.1051/epjconf/201611911001.

Vaughan, M., K.-P. Lee, A. Garnier and B. Getzewich, 2016: Improvements to the CALIOP Surface Detection Algorithm, CALIPSO CloudSat Science Team Meeting, 1-3 March 2016, Newport News, VA USA, https://doi/org/10.13140/RG.2.2.21628.04481.

Vaughan, M., A. Garnier, Y. Hu, Z. Liu, S. Young, D. Winker, K. Powell, and B. Getzewich, 2022: Diurnal Differences in Lidar Ratios for Opaque Water Clouds, CALIPSO–CloudSat Science Team Meeting, 12–14 September 2022, Fort Collins, CO, http://doi.org/10.13140/RG.2.2.31662.89920.

Venkata, S. L. and J. A. Reagan, 2016: Aerosol Retrievals from CALIPSO Lidar Ocean Surface Returns, *Remote Sens.*, **8**, 1006, https://doi.org/10.3390/rs8121006.

Winker, D., 2003: Accounting for multiple scattering in retrievals from space lidar, 12th International Workshop on Lidar Multiple Scattering Experiments, edited by: Werner, C., Oppel, U. G., and Rother, T., International Society for Optical Engineering, *SPIE Proc.*, **5059**, 128–139, https://doi.org/10.1117/12.512352.

Winker, D., X. Cai, M. Vaughan, A. Garnier, B. Magill, M. Avery, and B. Getzewich, 2024: A Level 3 monthly gridded ice cloud dataset derived from 12 years of CALIOP measurements, *Earth Syst. Sci. Data*, **16**, 2831–2855, https://doi.org/10.5194/essd-16-2831-2024.

Yang, P., L. Bi, B. A. Baum, K.-N. Liou, G. W. Kattawar, M. I. Mishchenko, and B. Cole, 2013: Spectrally Consistent Scattering, Absorption, and Polarization Properties of Atmospheric Ice Crystals at Wavelengths from 0.2 to 100 µm, *J. Atmos. Sci.*, **70**, 330–347, https://doi.org/10.1175/JAS-D-12-039.1.

Young, S. A., 1995: Analysis of lidar backscatter profiles in optically thin clouds, *Appl. Opt.*, **34**, 7019-7031, https://doi.org/10.1364/AO.34.007019.

Young, S. A. and M. A. Vaughan, 2009: The retrieval of profiles of particulate extinction from Cloud Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) data: Algorithm description, *J. Atmos. Oceanic Technol.*, **26**, 1105–1119, https://doi.org/10.1175/2008JTECHA1221.1.

Young, S. A., M. A. Vaughan, R. E. Kuehn, and D. M. Winker, 2013: The Retrieval of Profiles of Particulate Extinction from Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) Data: Uncertainty and Error Sensitivity Analyses, *J. Atmos. Oceanic Technol.*, **30**, 395–428, https://doi.org/10.1175/JTECH-D-12-00046.1.

Young, S. A., M. A. Vaughan, J. L. Tackett, A. Garnier, J. B. Lambeth, and K. A. Powell, 2018: Extinction and Optical Depth Retrievals for CALIPSO's Version 4 Data Release, *Atmos. Meas. Tech.*, **11**, 5701–5727, https://doi.org/10.5194/amt-11-5701-2018.