CATS Algorithm Theoretical Basis Document

Level 1 and Level 2 Data Products

Primary Authors:

John E. Yorks, Science Systems and Applications, Inc., Lanham, MD, USA

Stephen P. Palm, Science Systems and Applications, Inc., Lanham, MD, USA

Matthew J. McGill, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Dennis L. Hlavka, Science Systems and Applications, Inc., Lanham, MD, USA

William D. Hart, Science Systems and Applications, Inc., Lanham, MD, USA

Patrick A. Selmer, Science Systems and Applications, Inc., Lanham, MD, USA

Edward Nowottnick, GESTAR, Universities Space Research Asso., Columbia, MD USA

Release 1.2

26 February 2016

Cloud-Aerosol Transport System

CATS Algorithm Theoretical Basis Document

Prepared By:

_____

John E. Yorks   Date

_____

Stephen P. Palm   Date

_____

Dennis L. Hlavka   Date

_____

William D. Hart   Date

_____

Patrick A. Selmer   Date

_____

Edward Nowottnick   Date

Approved By:

_____

Matthew J. McGill   Date
CATS Principal Investigator

Table of Contents

1.0 Introduction

1.1 Purpose

This document describes the algorithms that will be used to calibrate the lidar backscatter profiles acquired by the CATS (Cloud-Aerosol Transport System) instrument flown aboard the International Space Station (ISS). The outputs of these algorithms are Level 1 data, consisting of attenuated backscatter coefficient profiles for the two channels (532 and 1064 nm) along with information on the uncertainties in these products. The Level 2 algorithms to produce geophysical parameters such as layer heights and optical depths use the Level 1 data. In addition, calibration files are generated that track the calibration constants that are derived during Level 1 processing. The data used by the Level 1 processing are geolocated prior to calibration.

1.2 Revision History

Issue Date	Release Number	Description	Lead Author	Sections Affected
06/12/15	1.1	Initial Release	John Yorks	1,2,3
02/26/16	1.2	Level 2 Release	John Yorks	1 - 5

1.3 CATS Mission Overview

The Cloud-Aerosol Transport System (CATS), launched on 10 January 2015, is a lidar remote sensing instrument that provides range-resolved profile measurements of atmospheric aerosols and clouds. Data from CATS is used to derive properties of cloud/aerosol layers including: layer height, layer thickness, backscatter, optical depth, extinction, and depolarization-based discrimination of particle type. The instrument is located on the Japanese Experiment Module – Exposed Facility (JEM-EF) on the International Space Station (ISS). The ISS orbit is a 51-degree inclination orbit at an altitude of about 405 km. This orbit provides more comprehensive coverage of the tropics and mid-latitudes than sun-synchronous orbiting sensors, with nearly a three-day repeat cycle. CATS is intended to operate on-orbit for at least six months, and up to three years. The CATS payload is designed to provide a combination of long-term operational science, in-space technology demonstration, and technology risk reduction for future Earth Science missions.

The measurements of atmospheric clouds and aerosols provided by the CATS payload are used for three main science objectives. 1) One important aspect of the CATS on-orbit science is to provide real-time observations of aerosol vertical distribution as inputs to global models. The vertical profile information obtained by CATS, particularly at multiple wavelengths and with depolarization information, provides height location of

cloud and aerosol layers, as well as information on particle size and shape. 2) Another important aspect of the CATS on-orbit science is to extend the space-based lidar record for continuity in the lidar climate observations. The CATS instrument provides measurements of cloud and aerosol profiles similar to CALIPSO, filling in the data gap, so this information can continually be used to improve climate models and our understanding of the Earth system and climate feedback processes. 3) Finally, CATS advances technology in support of future space-based lidar mission development by demonstrating the ability to retrieve vertical profiles using a high rep-rate laser and photon counting detection, as well as the testing of component for the High Spectral Resolution Lidar (HSRL) technique and 355 nm wavelength.
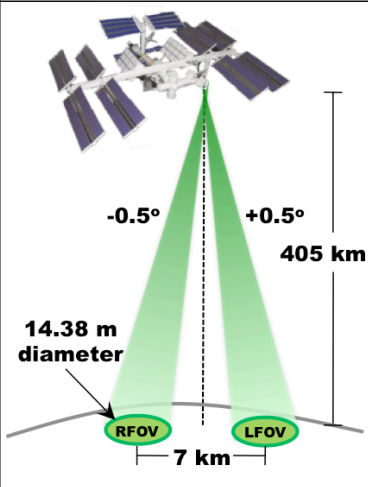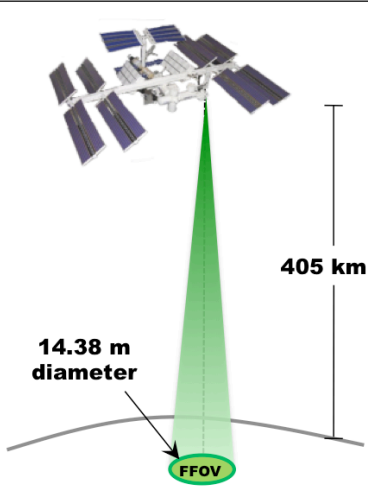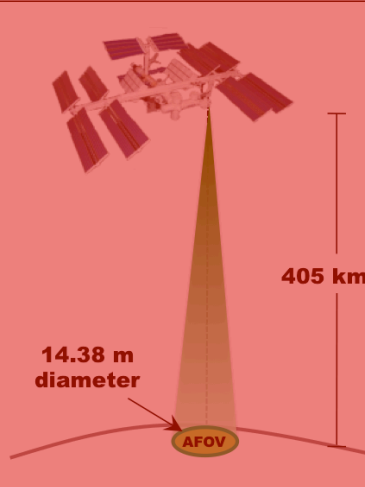
Mode 7.1: Multi-Beam	Mode 7.2: HSRL Demo	Mode 7.3: UV Demo
Backscatter: 532, 1064 nm No HSRL Depolarization: 532, 1064 nm	Backscatter: 532, 1064 nm HSRL: 532 nm Depolarization: 1064 nm	Backscatter: 355, 532, 1064 nm No HSRL Depolarization: 532, 1064 nm

Semi-continuous operation: Feb. 10 – Mar. 21 Failure: under investigation	Semi-continuous operation: Mar. 25 – Present Future Mode of Operation	Failure in laser optics No data available

Figure 1.1 CATS three main Science Modes for operation, with details of each mode's capabilities and operational status.

To meet these three science goals, CATS operates in three different modes using four instantaneous fields of view (IFOV) as shown in Figure 1.1:

- *Mode 7.1: Multi-beam backscatter detection at 1064 and 532 nm, with depolarization measurement at both wavelengths*. The laser output is split into two transmit beams, one aimed 0.5º to the left and one 0.5º to the right, effectively making two tracks separated by 7 km (~4.3 mi) at Earth's surface. This operational mode can no longer be used due to a failure in laser 1 electronics.
- *Mode 7.2: Demonstration of HSRL aerosol measurements*. This mode was designed to use the injection-seeded laser operating at 1064 and 532 nm to demonstrate a high spectral resolution measurement using the 532-nm

2

wavelength. However, this mode has been limited to 1064 nm backscatter and depolarization ratio because issues with stabilizing the frequency of laser 2 prevent collection of science quality HSRL data.

- **Mode 7.3: Demonstration of 355-nm profiling**. This mode was designed to use the injection-seeded laser operating at 1064, 532, and 355 nm to demonstrate 355-nm laser performance. Unfortunately, due to an unexpected failure in the laser optical path, CATS will not collect data in this mode.

1.4 CATS Data Product Levels

The CATS Level 1B and 2 data processing algorithms rely heavily on heritage from existing airborne and space-based lidar systems, such as the Cloud Physics Lidar (CPL, McGill *et al.* 2002), the Airborne Cloud-Aerosol Transport System (ACATS, Yorks *et al.* 2014), and the Cloud-Aerosol Lidar Infrared Pathfinder Spaceborne Observations (CALIPSO) satellite (Winker *et al.* 2009). The HSRL data processing algorithms will be very similar to those used for the ACATS instrument, with much longer turnaround times expected.

The data products generated from the CATS measurements are produced according to a protocol that is similar to that established by NASA's Earth Observing System (EOS), but are not required to meet any specific protocol. The CATS data product levels are defined as follows:

- **Level 0**: reconstructed, unprocessed instrument data at raw resolutions (i.e., the downlinked raw photon counts from the CATS instrument). Any and all communications artifacts (e.g. synchronization of packets, communications headers, duplicate or missing data) are removed in the L0 process.
- **Level 1A**: Level 0 data that is time-referenced, geo-located, corrected for detector nonlinearity and instrument artifacts, normalized to laser energy, and annotated with ancillary information. The CATS Level 1A data (relative normalized backscatter) is an internal product only and is not distributed.
- **Level 1B**: Level 1A data that have been calibrated, annotated with ancillary meteorological data, and processed to sensor units. The CATS Level 1B data (attenuated total backscatter and depolarization ratio) is archived as Level 1 data.
- **Level 2**: Geophysical parameters derived from Level 1 data, such as the vertical feature mask, profiles of cloud and aerosol properties (i.e. extinction, particle backscatter), and layer-integrated parameters (i.e. lidar ratio, optical depth). There will be two CATS Level 2 products:
    - **CATS Heritage L2**: L1B files that are run through the CALIPSO L2 algorithms to provide continuity in the algorithms used for the lidar climate record.
    - **CATS Operational L2**: L1B files that are run through the new operational CATS L2 algorithms, which will include new capabilities that correspond to new instrument technology.

2.0 Instrument Description

The CATS payload is based on existing instrumentation built and operated on the high-altitude NASA ER-2 aircraft. The instrument consists of 2 high repetition rate Nd:YVO4 lasers operating at three wavelengths (1064, 532, and 355 nm) that generate signal photons, a receiver subsystem with a 60 cm diameter telescope to collect photons that backscatter from the atmosphere, and a data system to provide timing of the return photon events. The CATS instrument parameters are given in more detail in Table 2.1.

2.1 Transmitter Subsystems

The CATS laser transmitter that will be used in Mode 7.1, referred to as Laser 1, gets its heritage from the Cloud Physics Lidar (CPL; McGill *et al.* 2002) instrument. It is a Nd: YVO$_4$ laser with a repetition rate of 5000 Hz and an output energy of about 1 mJ per pulse at 532 and 1064 nm. An image of the laser 1 bench assembly is shown in Figure 2.1.

<div align="center">

Table 2.1. CATS Instrument Parameters.

Laser 1 Type	**Nd: YVO$_4$**
Laser 1 Wavelengths	**532, 1064 nm**
Laser 1 Rep. Rate	**5000 Hz**
Laser 1 Output Energy	**~1 mJ/pulse**
Laser 2 Type	**Nd: YVO$_4$, seeded**
Laser 2 Wavelengths	**355, 532, 1064 nm**
Laser 2 Rep. Rate	**4000 Hz**
Laser 2 Output Energy	**~2 mJ/pulse**
Telescope Diameter	**60 cm**
View Angle	**0.5 degrees**
Telescope FOV	**110 microradians**

</div>

The frequency characteristics of pulsed lasers have recently been advanced due to the development of direct detection Doppler lidars and HSRLs. These techniques impose further requirements compared to standard backscatter lidars, such as lasers that are single frequency on a single pulse basis and more stable in time (central frequency drift of less than 1 MHz per minute). An injection-seeded, pulsed Nd: YVO$_4$ laser was developed for CATS, with heritage from a similar laser transmitter built for the Airborne Cloud-Aerosol Transport System (ACATS: Yorks *et al.* 2014), that achieves these frequency characteristics (Hovis *et al.* 2004). This laser, referred to as Laser 2 and shown in Figure 2.2, provides a narrow wavelength distribution suitable for resolving the small frequency shifts due to the Doppler effect. The laser operates at an output power of about 2 mJ per pulse and repetition rate of 4000 Hz. This seeded laser also contains an external frequency-tripling module to provide output at 355 nm for Mode 7.3 that should failure occur, will not negatively impact operations at 532 and 1064 nm.
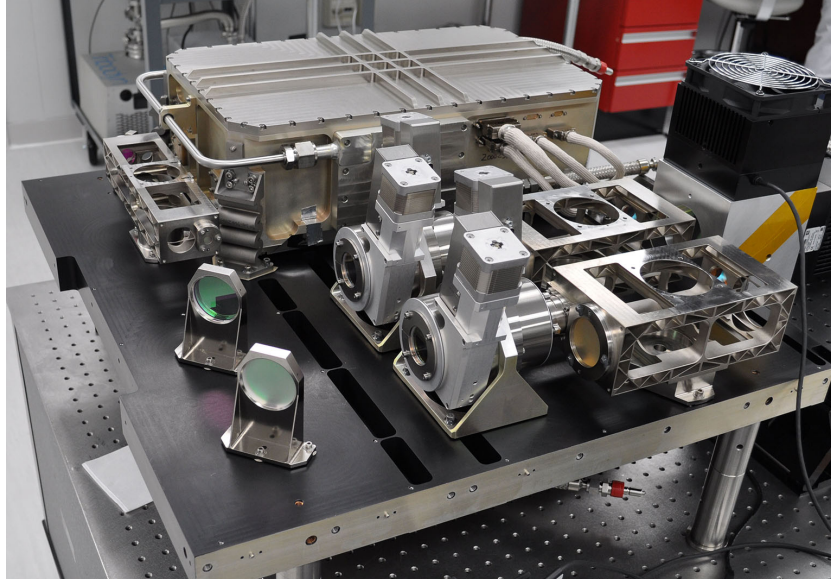
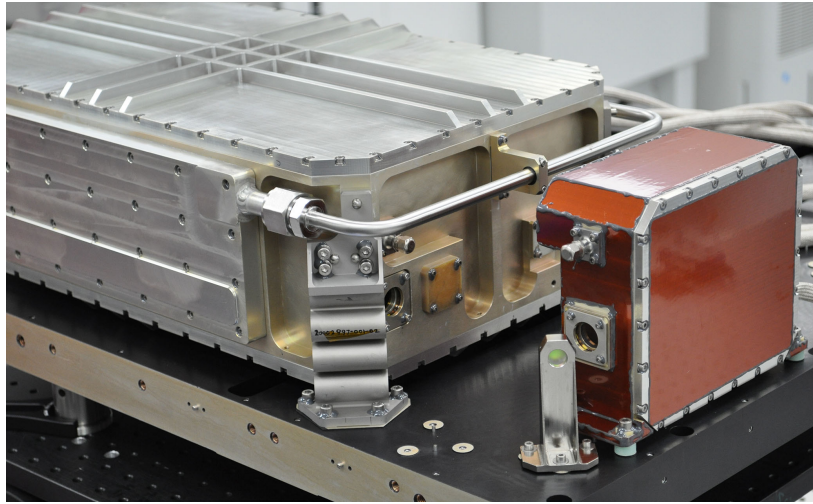**Figure 2.1** The CATS laser 1 bench assembly for 532 and 1064 nm operation.

Figure 2.2 The CATS laser 2 bench assembly with the external frequency-tripling module (orange).

2.2 Receiver Subsystems

CATS employs a 60 cm beryllium telescope that has a 110 microradian field of view, allowing for a 0.5 degree view angle. The telescope, shown in Figure 2.3, is also fiber-coupled to the detector boxes to provide greatest flexibility. CATS contains four detector boxes. Beam splitters are used to measure the parallel and perpendicular polarized return in all four detector boxes. The first two detector boxes are identical and are used for the LFOV and RFOV of Mode 7.1. Each of these detector boxes contain six detector channels for the following detection:

- 532 nm Parallel Backscatter (2)

5

- 532 nm Perpendicular Backscatter (2)
- 1064 nm Parallel Backscatter
- 1064 nm Perpendicular Backscatter

A third and similar detector box is used for Mode 7.3 that contains the same 6 channels as the Mode 7.1 detector boxes, plus 1 additional 355 nm Total Backscatter channel for a total of 7 channels. These three detector boxes are shown in Figure 2.4. The final detector box, referred to as the HSRL detector box, contains 12 detector channels: 10 designated as 532 nm HSRL channels, one 1064 nm Parallel and one 1064 nm Perpendicular backscatter channel.

The heart of the CATS HSRL detector box is an etalon that provides the spectral resolution needed for the HSRL measurement. Backscattered light collected by the telescope is passed through the etalon and an image of the etalon fringe pattern is created. A bandpass filter is used in tandem with the etalon to reject background sunlight, permitting daytime operation. The optical gap of the etalon is 3 cm with a plate reflectivity of 90%. It is critical to maintain the symmetry and shape of the etalon fringe pattern to avoid uncertainty in the measurement. A digital etalon controller was developed by Michigan Aerospace Corporation in which piezoelectric actuators control the etalon electronics to position and maintain the plate parallelism. A holographic circle-to-point converter optic (McGill *et al.* 1997c; McGill and Rallison 2001) is placed in the focal plane of the HSRL receiver to provide the spectral detection. The circle-to-point converter simplifies hardware requirements, improves efficiency of measuring the spectral content in the fringe pattern, and allows CATS to utilize photon-counting detection. The holographic optic is coupled to the 10 individual 532 nm detectors, each representing a small wavelength interval.
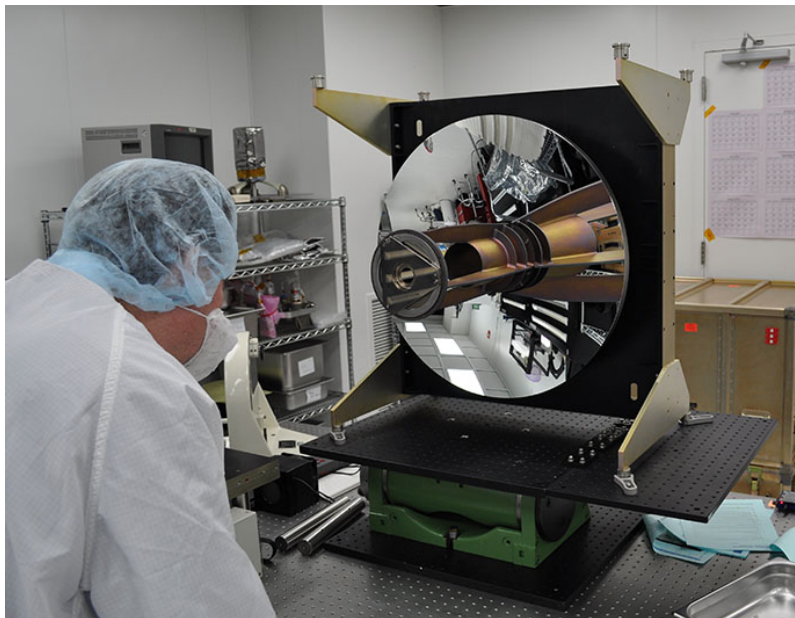

Figure 2.3 The CATS 60 cm beryllium telescope prior to full instrument assembly.

Figure 2.4 The three standard backscatter detector boxes for Modes 7.1 (LSFOV and RSFOV) and 7.3.

2.3 Data Acquisition and Signal Processing

During Level 0 processing, data are time-sorted and corrected for communication artifacts (i.e. duplicate or missing data). Raw CATS data, with 60 m (Mode 7.1; 78 m for Mode 7.2) vertical resolution and 350 m horizontal resolution, are received from the ISS in real-time during acquisition-of-signal (AOS) periods. During loss-of-signal (LOS) periods, which can range from 1 minute to 2 hours, data are not received real-time. CATS data are recorded on-board the ISS during LOS periods then transmitted to the ground during subsequent AOS periods. Due to the LOS periods, data are collected at the CATS ground station in 3-hour segments to ensure all data can be sorted properly. Level 0 files are created once two 3-hour segments have been obtained. Every time a new segment is collected, the previous segment is processed into Level 0 files using the new segment to fill in data gaps caused by LOS periods. Thus, Level 0 files are produced every 6 hours. The ISS time data, which is reported in the CATS Level 0 files, is corrected for drift using a special ISS data stream during Level 0 processing.

Level 0 files are partitioned into either day or night "granule" files based on two criteria. These criteria are:

1) The z-component of the solar line-of-sight unit vector reported in the ISS Broadcast Ancillary Data (BAD) must meet a threshold value of greater than 0.0 in order for a file to be deemed night. A value less than 0.0 would be classified as day.

2) The solar background counts for the given profile must cross a threshold value of 6 counts to be classified as day. A profile is classified as night if the solar background counts are less than 6.

A new granule file is produced when both criteria agree for a given profile and these granules are then labeled correspondingly as either a "day" file or "night" file. It should be noted that there are occasions when the 6-hour sorting window is not large enough to fill in the data gaps caused by out-of-sequence data. In this scenario, two granules may be produced, with 4.5 minutes between the start and end times of the granules, instead of one larger granule.

Some CATS Version 2.06 files (also known as granules) are shorter than the normal approximately 45 minute long files. These files can be as short as 5 minutes of data. There are two explanations for granules sometimes having short lengths.

1) Sometimes data capture begins or ends near the cutoff criteria for granules (i.e. begins at the end of the daytime/night portion of the orbit, ends at the beginning of the granule).

2) A more complex reason is that due to a communications defect between the CATS on-board data system and ISS navigation system, the data capture time in the CATS raw data occasionally deviates from the true time. This occurs randomly and for various durations of time in sequential data records. The way the CATS L0 code processes this phenomenon is by analyzing sequencing numbers to determine if the time is valid. If the time is deemed invalid, the code waits until the sequencing number and data capture time are both in sync for a certain amount of time before producing granules again. This introduces gaps in the data. Once the gaps get big enough to cross a threshold, the code breaks them up into separate files even if they belong to the same "day" or "night."

3.0 Overview of Level 1 Algorithms

3.1 Normalized Relative Backscatter

The CATS Level 1A data is referred to as the Normalized Relative Backscatter (NRB) and is an internal product only that is not distributed. The NRB data is Level 0 data that is geolocated, corrected for detector nonlinearity, range, and the folding of molecular signal from the atmosphere above, normalized to laser energy, and annotated with ancillary information. The ancillary information included in the NRB data is the Broadcast Ancillary Data (BAD) from the ISS that describes the environment in which a payload is operating. The BAD is sent at a rate of 10 Hz and includes the roll, pitch and yaw of the ISS, the quaternion, and the CTRS position information. Since the CATS laser points off-nadir 0.5 degrees in any of the three science modes, and has multiple beams in Mode 7.1, the geolocation of the CATS laser beam is computed for each FOV using the BAD and the CATS relative angles. More information on the algorithm to determine the geolocation of each CATS FOV is provided in section 3.1.1. Once the data is geo-located, the raw photon counts are corrected for detector nonlinearity (D), as described in section 3.1.2. The Normalized Relative Backscatter (β'_{NR}) for each wavelength and range bin is then computed using the equation:

$$\beta'_{NR}(r) = \frac{\{[N(r) * D] - N_B\}r^2}{E}$$

Eq. 3.1

where r is the range at each range bin and E is the laser energy as measured by energy monitors installed on the CATS instrument. The solar background photon counts (N_B) is estimated by averaging the signal below the earth's surface. All the products reported in the CATS L1A data products are used as input to the CATS L1B data products.

3.1.1 Geolocation of CATS Laser Beams

Knowledge of the location of the CATS laser spot on the earth is required for the useful analysis of the CATS backscatter data. The location of the CATS laser spots can be calculated from the position, velocity, and attitude information found in the ISS Broadcast Ancillary Data (BAD) together with the known angular offset of the laser line-of-site (LOS) vector from the instrument's nadir vector. The computation requires a series of coordinate transformations and rotations to find the geodetic latitude and longitude of the laser spot at the height of the Digital Elevation Model (DEM).

The CATS algorithm uses four coordinate systems in the computation of the laser spot location, which include:
1) The CATS instrument body (LOS) reference frame
2) The local vertical local horizontal (LVLH) reference frame
3) The Conventional Terrestrial or geocentric (CTRS, x,y,z) reference frame
4) The geodetic reference frame (longitude, latitude, altitude).

Because angular offsets between the ISS body reference system and the CATS body reference system are unknown, the two systems are considered to be the same. The difference is assumed to be small so the BAD data that is referenced to the ISS body is considered to be reference to the CATS body.

The ISS body position data ($x_{iss}, y_{iss}, z_{iss}$) and velocity data ($v_{xiss}, v_{yiss}, v_{ziss}$) are contained in the ISS BAD. These vectors can be used to construct the LVLH unit vector $\vec{M}_{vh}$, along with the following information.

1) The forward velocity unit vector of the ISS
2) The nadir unit vector
3) The cross product of the forward velocity unit vector of the ISS and the nadir unit vector for a right handed coordinate system
4) The geocentric unit vector of the ISS
5) The speed and distance from the earth's center, respectively

The ISS attitude data is also included in the BAD in the form of quaternion components ($q_{wiss}, q_{xiss}, q_{yiss}, q_{ziss}$), which are the scalar and x, y, z vector quaternion components

referenced to the LVLH reference frame. These quaternion components can be converted to the yaw, pitch, and roll (a_y, a_p, a_r). Yaw, pitch, and roll can be used to construct a matrix to transform the laser LOS vector into the LVLH reference system. The matrix is:

$$M_{ypr} = \begin{pmatrix} C_y C_p & -S_y C_r - C_y S_p S_r & S_y S_r \\ S_y C_p & C_y C_r - S_y S_p S_r & -C_y S_r - S_y S_p C_r \\ S_p & C_p S_r & C_p C_r \end{pmatrix} \qquad \text{Eq. 3.2}$$

where C and S represent functions cosine and sine and subscripts y, p, and r refer to yaw, pitch, and roll, respectively. By using the LVLH unit vector and equation 3.2, the CATS laser pulse location at a sampling bin in geodetic coordinates can be found using the following steps.

1) Compute geodetic coordinates of ISS (a_{tgd} a_{ngd} h_{gd})=V(x_{iss}, y_{iss},z_{iss}), where (a_{tgd} a_{ngd} h_{gd})= geodetic latitude, longitude, and altitude of the ISS and V(x,y,z) is the Vermeille transformation from geocentric to geodetic coordinates (Vermeille 2002).

2) Define the LOS vector for the appropriate CATS laser pointing direction, fore, aft, left, or right. The CATS body reference system is defined by positive x-axis along the forward direction, positive y-axis to the left, and positive z-axis toward the earth. A LOS vector is defined in spherical coordinates by the rotation about the z-axis (f), angle from the x-y plane (q) and range from CATS (R_B). For all four pointing directions, q=(p/2 - 0.008) radians. For fore, aft, left, and right f=0, p, p/2, and 3p/2, respectively. So, in Cartesian coordinates, the laser LOS vector is:

$$(x_{Bis} \quad y_{Bis} \quad z_{Bis}) = (R_B \cos\phi \cos\theta \quad R_B \sin\phi \cos\theta \quad R_B \sin\theta) \qquad \text{Eq. 3.3}$$

3) Use M_{ypr} from equation 3.X to transform the laser LOS vector to LVLH reference frame.

4) Use the LVLH unit vector (M_{vh}) to convert (x_{Bvh} y_{Bvh} z_{Bvh}) to the geocentric reference frame. This conversion results from the following equations.

5) Compute the geocentric coordinates of the laser spot by adding Cartesian coordinates in the geocentric reference to the ISS coordinates.

6) Convert the laser spot coordinates to geodetic reference frame by using Vermeille's transformation.

7) Compute the CATS laser LOS vector in the geodetic reference frame for a series of range bins R_{bi} that will positively pass through the altitude of the local DEM. Select the location where z-component of the CATS laser LOS vector most closely matches the altitude reported by the DEM. The latitude and longitude of that bin will the latitude and longitude of the laser spot.

The DEM values used are average values for 1 km by 1km grid boxes. This will lead to some discrepancies in the latitude and longitude of the CATS laser spot in mountainous terrain. To verify the accuracy of the CATS laserspot location algorithm, spot location results were compared with those derived from an algorithm presented by J. R. Ridgway in the paper "Analysis of Satellite Laser Altimetry Range Measurements Over Land Topography Using Generalized Attitude Parameter Extraction", published in the ISPRS (International Society for Photogrammetry and Remote Sensing) proceedings. These comparisons revealed errors in the CATS algorithm used in version 2.04.
To correct these errors the following changes were made to the CATS algorithm:

1. Signs on elements in the roll and pitch LVLH (local horizontal local vertical) transformation matrix were changed (from positive to negative and vice versa).
2. The order of the matrix multiplication to transform the laser line-of-sight vector to LVLH was reversed.

After the changes were made, the test laser spot locations computed by the CATS footprint geolocation algorithm and the Ridgeway algorithm were the same to within 2.0 meters, demonstrating that the ISS pointing angles are being used properly. Although the updated algorithm in version 2.05 is an improvement, there is still a bias in the geolocation. Over rugged terrain, differences are observed between the DEM and lidar signal. These differences are attributed to two factors.

1) CATS is mounted on the Japanese Experiment Module-Exposed Facility (JEM-EF), which is a distant extremity of the ISS. It is reasonable to assume that an angular offset exists between the ISS central body, where the BAD are measured, and the CATS local reference system on the JEM-EF. However, the angle between the ISS point of reference for the position data and CATS instrument is unknown and assumed to be zero in V2.05. This assumption is likely a main source of error in the CATS V2.05 geolocation algorithm.
2) A communications defect in which a time lag (1-2 seconds in some cases) from when the ISS position data is collected to when it is included into the CATS data stream occurs. Statistically this time lag is found to be 1 second in 80% of CATS data profiles, so the L0 data is corrected this artifact by adjusting the ISS position data by 1 second. However, the time lag is not constant so instances when the time lag is not 1 second do occur and users should be careful when analyzing surface returns over rugged terrain.

The angular offset between the ISS point of reference for the position data and CATS instrument was determined using statistical analysis of the differences between the surface altitude detected using the CATS backscatter data and the DEM of the latitude and longitude computed using the V2.05 algorithm. These differences can be reduced or eliminated by applying a constant angular correction to the yaw, pitch, and roll parameters that are determined from the BAD quaternion values. In each scene, systematic incremental adjustments were made to the yaw, pitch, and roll until the surface height values agree. The offsets to the ISS yaw and roll were constant for the many scenes that were tested but the pitch offset varied. An empirical distribution

function of the pitch offsets was derived from many cases and the mean and median were determined from the distribution. The mean pitch offset was used as the correction.

From this analysis, the angles between the ISS point of reference for the position data and CATS instrument was determined as 2.00 degrees for yaw, -0.50 degrees for roll, and -0.25 for pitch. These "offset" angles have been incorporated into the V2.06 algorithms. Figure 3.1 shows the CATS attenuated total backscatter at 1064 nm for 19 Aug. 2015 (top panel) as the instrument passed over the continent of Africa (ISS track shown in blue in bottom panel, with the red track indicating the scene plotted). The DEM surface altitude is overlaid (green) for the CATS footprint geolocation computed using the V2.06 algorithm in the top panel, which shows very good agreement with the CATS attenuated total backscatter. The variation in the pitch correction could be caused by the uncorrected tag lag, which can shift the laser spot geolocation compared to the CATS backscatter data along the track of the ISS. Future studies will be performed to correct the time lag issue in post-processing.
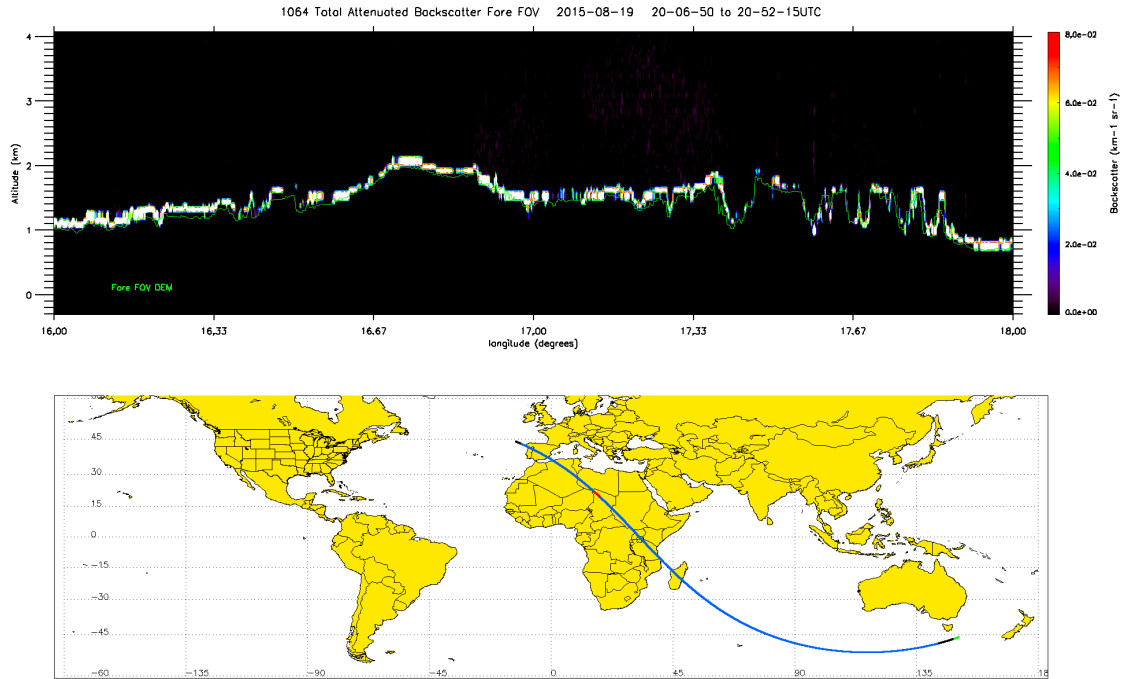

Figure 3.1. Comparison of the CATS attenuated total backscatter at 1064 (top panel) shows that the surface detected by the instrument (white colors) agrees favorably with the DEM values (green).

3.1.2 Detector Nonlinearity

The measured CATS signal can become erroneously low when measuring targets that strongly backscatter light due to detector dead time and must be compensated for. All lidar systems that employ photon-counting detection experience this effect, which is a limitation on the number of photons that can be counted in a given time interval. For

CATS, highly reflective features, such as bright surfaces (desert, sea ice) and water clouds push the detector into a nonlinear counting region. A typical photon counting detector, such as the ones employed in CATS, has a discriminator dead time of 28 to 30 ns for a discriminator maximum count rate on the order of 30 MHz.

The nonlinear effects for this type of detector can be quantified by a detector dead time coefficient. This coefficient represents the fact that only one photon event can be counted at once, and the detector system has a certain time delta, or dead time, before it can count another. If the mean time between events is much greater than the dead time, then it can be shown that

$$N_{a,i} = \frac{N_{m,i}}{1 - \left(\dfrac{N_{m,i}\tau}{\Delta t}\right)}$$

Eq. 3.4

where $N_{m,i}$ is the measured counts on channel i, $N_{a,i}$ is the counts that would be sensed if the detector were completely linear, Δt is the total integration time, and τ is the dead time. These nonlinear effects can be significantly reduced by applying Eq. 3.4 to the measured signal, which allows for a reasonable correction of the atmospheric data bins. The CATS detectors rarely experience count rates higher than 35 MHz in atmospheric bins below 28 km. Therefore, the detector dead time coefficient is less than 1.10 for 99.0% of atmospheric bins. An example of the CATS deadtime correction factors, as a function of photon counts, are shown in Figure 3.2 for detector 1 (RFOV 1064 nm parallel channel).
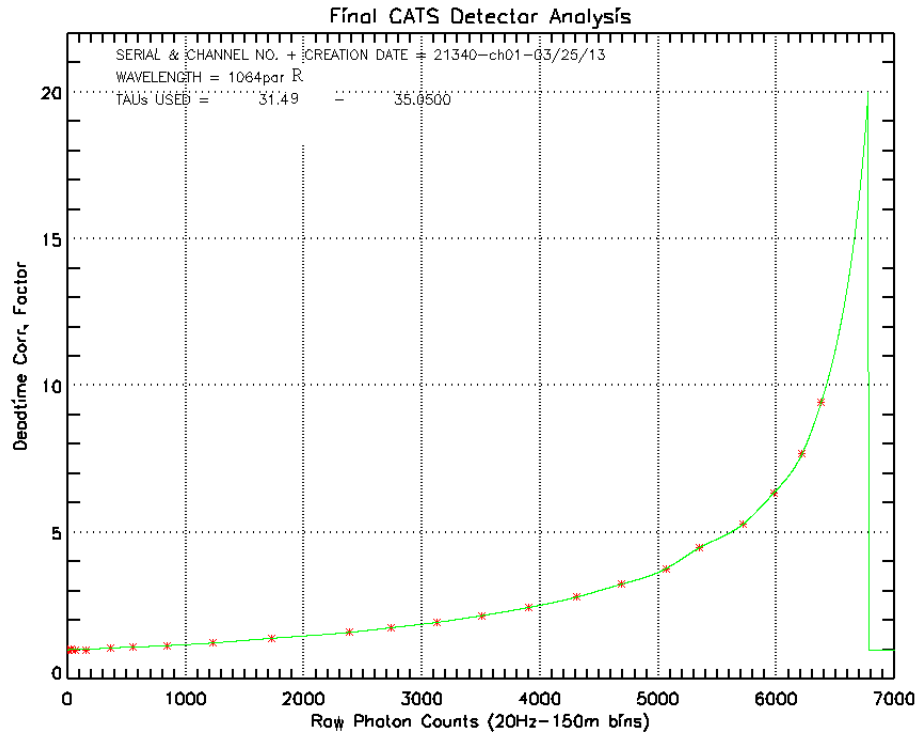

Figure 3.2. The CATS deadtime correction factors for RFOV 1064 nm parallel channel detector as a function of photon counts.

3.1.3 Correction for Molecular Folding

The raw photon data captured by CATS at range r (where r < 28 km) will have contributions from atmospheric scattering at heights z+Nx km, where N=1,2,3, etc. and x = 30 for operating mode 7.1 and 37.5 for mode 7.2. This effect is caused by the high repetition rate of the CATS lasers discussed in section 2.2. In practice, only N=1 is important as scattering above 60 km is negligible. The folding of molecular scattering is important to remove because accurate background cannot be calculated without doing so. Background is computed from the 2 km of data below the ground. However, the data in this region contains molecular scattering from the 28 to 30 km region of the atmosphere. If this signal is not removed, it becomes part of the calculated background. If this background is subtracted from each bin of the profile, it will remove most of the photon counts from true molecular scattering in the calibration zone (23-27 km altitude). This will render accurate calibration from the molecular signal impossible.

The molecular contribution to the measured photon count can be computed from equation 3.3:

$$N_m(r,\lambda) = \frac{N_e(\lambda)}{r^2} \beta_m(r,\lambda) \Delta r A_t T_m^2(r,\lambda) T_o^2(r) Q_e T_{opt} N_a R(r,\lambda) \alpha(\lambda) \qquad \text{Eq. 3.5}$$

Where N_e is the number of photons transmitted by CATS which is defined by the laser energy (E) as:

$$N_e = \frac{E(\lambda)\lambda}{hc} \qquad \text{Eq. 3.6}$$

Where λ is the laser wavelength (532 nm), h is the Planck constant and c is the speed of light. The other terms used in equation 3.5 above are:

r – The range from the satellite to the height z (in m).
$\beta_m(r)$ – the molecular backscatter cross section at range r (m^{-1} sr^{-1}).
Δr – the bin size in meters (60 m)
A_t – Area of telescope (m, effective)
$T_m(r)$ – Molecular atmospheric transmission from top of atmosphere to range r.
$T_o(r)$ – Ozone transmission: top of atmosphere to range r.
Q_e – Quantum efficiency of detector
T_{opt} – Transmission of the receiver system optics
N_a – Number of shots summed (nominally 250)
$R(r)$ – aerosol scattering ratio
$\alpha(\lambda)$ - scaling factor

In equation 3.5, α is used to adjust the computed photon count since the exact values of quantities like system optical transmission (T_{opt}) and detector quantum efficiency (Q_e) are not known exactly and moreover, can change with time. The value of α was computed empirically by adjusting it until the slope of the average NRB signal (for a granule) above 20 km matched the slope of the average modeled molecular backscatter, which is derived using the technique described in section 3.2.2. It was found that too low of an alpha will

produce a slope less than the molecular model, and a value too large will produce too large a slope. Further, it was found that alpha varies only slowly with time and requires only infrequent tuning. Alpha is computed empirically for each wavelength and each field of view. Typical values of alpha range from 0.04 to 0.08 for the 1064 nm channel and 0.02 to 0.07 for the 532 nm channel.

Equation 3.5 is used to compute a profile of molecular scattering contribution from 58 km to 28 km ($N_m(r)$). From that profile, the molecular scattering contribution (folded from above) to the measured CATS photon profile is computed as:

$$N'_m(r) = N_m(r + x)$$ Eq. 3.7

Where x = 30 km for operating mode 7.1 and x = 37.5 for operating mode 7.2. For r between -2 and 28 km. Then the corrected raw photon count profile is:

$$S'(r) = S(r) - N'_m(r)$$ Eq. 3.8

Where $S(r)$ is the raw photon count profile measured by CATS. Note that this process leaves the molecular scattering of the original profile ($S(r)$) intact. It only removes the molecular scattering folded down from above. The NRB corrected for the molecular folding can now be computed as:

$$NRB'(r) = (S'(r) - N_B)/E = C\beta(r)T^2(r)$$ Eq. 3.9

For version V2-06 of the CATS L1A and L1B products, an iterative computation of the alpha parameter is employed. A few hundred nighttime granules were used to derive a relationship between the below ground raw photon count (averaged for the granule) and the alpha parameter. This relationship was found to be a linear dependence. In V2-06 L1A processing this relationship was used to compute an initial value of alpha, which was then applied to the entire granule. The average NRB signal profile between 28 and 20 km is then computed for the granule, as is the average molecular profile. The slope of both of these profiles are computed and compared to each other. If the slopes differ by a pre-defined threshold, the value of alpha is adjusted and the whole granule is reprocessed. Alpha is adjusted based on the sign and magnitude of the slope difference. If, after processing the granule again, the slope difference is still greater than the threshold, the granule is quarantined (put in a holding directory). In the automated processing of over 5,000 granules, only about 50 had to be quarantined. For these files, the value of alpha will be determined manually to assure alpha is of the appropriate magnitude.

3.2 Calibrated Backscatter

The processing algorithms for Level 1B consist mainly of the backscatter and depolarization calibrations. The 532 nm CATS data is calibrated by normalizing the NRB signal to the 532 nm molecular backscatter signal in a set calibration region (Russell *et al.* 1979, Del Guasta 1998, McGill *et al.* 2007, Powell *et al.* 2009). The CATS calibration region is 23-27 km, starting 1 km below the top of the CATS data frame (28 km). The

aerosol loading in this region is computed using CALIPSO data and applied to the calibration. The CATS NRB signal is averaged to 4 minutes at night and 46 minutes during daytime operation to reduce uncertainty in the calculation. During nighttime data collection, the 1064 nm calibration constant can be computed using an identical approach as the 532 nm calculation. However during daytime operation, the 532 and 1064 nm signal is often calibrated using a default value derived from historical data or manual normalization to the Rayleigh backscatter model. The polarization gain ratio, which describes the relative gain between polarization channels, is computed for both 532 and 1064 using the solar radiation scattered by ice clouds (Liu *et al.* 2004).

3.2.1 Ozone Transmission

The ozone transmission, $T^2_o(r)$, is calculated using ozone mass mixing ratios obtained from the GMAO meteorological data set, which contains ozone mass mixing ratios. The ozone mass mixing ratios, $r_O(r)$ are first converted to column density per kilometer (atm-cm/km), $\varepsilon_O(r)$, using the following equation:

$$\varepsilon_O(r) = \frac{r_O(r)\rho(r)}{2.14148 \times 10^{-5}}$$

Eq. 3.10

where r is the range in km, and $\rho(r)$ is the atmospheric density at r and calculated from the meteorological data as:

$$\rho(r) = \frac{P(r)}{RT(r)}$$

Eq. 3.11

The next step is to calculate the ozone transmission term, $T^2_o(\lambda)$, which is computed using the following equation:

$$T^2_O(\lambda,r) = \exp\left[-2c_O(\lambda)\int_H^r \varepsilon_O(r')\ dr'\right]$$

Eq. 3.12

where $c_O(\lambda)$ is the Chappius ozone absorption coefficient in cm^{-1} and λ is 532 nm. The ozone absorption coefficient is obtained at the correct wavelength from a table compiled in Iqbal (1984) using data from Vigroux (1953). The 532 nm Chappius ozone absorption coefficient used is 0.065 cm^{-1}. The 1064 nm coefficient is ~0.0 cm^{-1}. H is nominally 60 km.

The ozone transmission is then modified to account for the off-nadir angle of the CATS laser beam. The off-nadir angle varies by CATS FOV and its computation is detailed in section 3.1.1. If θ is the off-nadir angle of the laser beam, the angle correction to the transmission is:

$$T^2_o(\lambda,r) = T^2_o(\lambda,r)^{\sec(\theta)}$$

Eq. 3.13

The application of the ozone transmission to the CATS calibration method is described in section 3.2.5.

3.2.2 Rayleigh Scattering

This section provides a brief introduction to atmospheric molecular scattering as it relates to CATS and describes the molecular scattering parameters computed for use in CATS data processing. Atmospheric molecular scattering consists of two components. Rayleigh scattering is considered elastic scattering from particles that are very small compared to the wavelength of the scattered radiation, such as molecules. Vibrational Raman scattering has a scattering cross section that is very small compared to Rayleigh scattering, so it is neglected when computing the molecular scattering (Bucholtz 1995; Bodhaine *et al.* 1999; She 2001). For lidar applications, the phrases Rayleigh scattering and molecular scattering are used as synonyms. The main sources of Rayleigh-scattered light are nitrogen and oxygen, since these two gases accounts for about 99% of the Earth's molecular atmosphere. The Rayleigh scattering intensity is related to the wavelength of incident radiation (λ) through the relationship λ^{-4} and dominates backscatter signals from elastic backscatter lidars at short laser wavelengths. For elastic backscatter lidars such as CATS, the Rayleigh scattering signal is used to normalize the total return signal and determine the instrument calibration constant. Additionally, the molecular backscatter coefficient (β_M) and molecular extinction coefficient (σ_M) must be known to reduce the unknown parameters in the standard lidar equation to two.

The molecular backscatter coefficient is determined from Rayleigh scattering theory (Tenti *et al.* 1974; Young 1981) and is proportional to atmospheric density. Thus, the molecular backscatter coefficient can be computed using its relationship to atmospheric temperature and pressure, as demonstrated by Collis and Russell (1976), through the equation:

$$\beta_M = \frac{p}{KT}(5.45 \times 10^{-32})\left(\frac{\lambda}{550}\right)^{-4.09}$$

Eq. 3.14

where T is the atmospheric temperature in units of Kelvin, p is the atmospheric pressure in units of hPa and K is the Boltzmann constant (1.38×10^{-23} J K^{-1}). Furthermore, the molecular extinction coefficient (σ_M) is resolved from the molecular backscatter coefficient though the relationship:

$$\sigma_M = \beta_M \left(\frac{8}{3}\right)\pi$$

Eq. 3.15

NASA Goddard Earth Observing System version 5 (GEOS-5) forecasts provided by the NASA Global Modeling and Assimilation Office (GMAO, Rienecker *et al.* 2008) deliver a forecast of the atmospheric temperature and pressure profiles for 72 vertical levels (0-

85 km AGL) at a horizontal resolution of 10 seconds that is subset along the ISS orbit track. The temperature and pressure from GMAO are interpolated to the CATS vertical bin width of 60 m over a range of 0 to 60 km AMSL to better match the vertical structure of the CATS lidar backscatter data. The molecular backscatter and extinction coefficients are computed using equations 3.14 and 3.15, respectively. These parameters, along with the interpolated temperature, pressure, and other GMAO variables are output in the Level 1B files. Errors in the forecasted GMAO are estimated to be 0.5 K for profiles of temperature (Prive' *et al.* 2012) and 1 hPa for surface pressure (Reinecker *et al.* 2008).

In cases when the parallel-polarized backscatter channel is normalized to the attenuated molecular backscatter, the molecular depolarization ratio (δ_M) must be considered. The depolarization ratio is defined and the ratio of perpendicular to parallel backscatter. The molecular depolarization ratio for the three CATS wavelengths is given in Table 3.1, as provided by the CALIPSO ATBD and Behrendt and Nakamura (2002). These values are used to compute an attenuated molecular backscatter for the parallel-polarized light and reduce error in the CATS calibration constant when normalizing the parallel channel to Rayleigh backscatter.

Table 3.1. CATS molecular depolarization ratios for three operating wavelengths

λ (nm)	δm (%)
355	1.554
532	1.430
1064	1.400

3.2.3 Polarization Gain Ratio

Pulsed lasers, such as the ones used in the CATS instrument, naturally produce linearly polarized light. Using a beam splitter in the receiver optics, the perpendicular and parallel planes of polarization of the backscattered light are measured. The linear volume depolarization ratio is defined as the ratio of perpendicular total (Rayleigh plus particle) backscatter to parallel total backscatter, and has values between 0.2 and 0.6 for non-spherical particles such as ice crystals (Sassen and Benson 2001; Yorks *et al.* 2011a). Deriving accurate depolarization ratios from CATS data requires knowledge of the relative gain between the perpendicular and parallel channels of the CATS receiver, referred to as the polarization gain ratio (PGR).

The CATS operational PGR consists of two terms. The equation for the CATS operational PGR is:

$$PGR_{OP} = PGR_1 \times PGR_2$$

Eq. 3.16

The first term, PGR_1, characterizes the relative gain between the perpendicular and parallel channels at both the 532 and 1064 nm wavelengths. The second term, PGR_2, corrects for poor depolarization purity in the 532 nm measurements.

The algorithm to compute PGR$_1$ uses the ratio of the parallel to perpendicular solar background radiation scattered from dense ice clouds, as outlined by Liu *et al.* (2004). The background light measured by CATS is the scattering of solar radiation by the surface, clouds, aerosols, and molecules in the atmosphere. The difference in solar background counts between the parallel and perpendicular channels is minimal since the solar radiation scattered by dense ice clouds is unpolarized in theory (Liou *et al.* 2000). Dense ice clouds used to compute the first polarization gain ratio term during daytime periods only include only the top cloud layers and are identified using four criteria:

1) **Mid-cloud temperature (T_M): T_M < -35 C**
2) **532 nm (mode 7.1) or 1064 nm (mode 7.2) integrated attenuated total backscatter (γ_{1064}): $0.008 < \gamma_{1064} < 0.044$ sr^{-1}**
   where:

   $$\gamma_{1064} = \int_{bottom}^{top} \beta_{1064} T^2_{1064} dz$$ Eq. 3.17

3) **1064 nm layer-integrated depolarization ratio ($\delta_{\lambda\alpha\psi}$): $0.30 < \delta_{\lambda\alpha\psi} < 0.80$**
   where:

   $$\delta_{1064} = \frac{\sum_{layer} NRB_{perp}}{\sum_{layer} NRB_{par}}$$ Eq. 3.18

4) **532 nm (mode 7.1) or 1064 nm (mode 7.2) optical depth (τ_{1064}): $\tau_{1064} > 1.75$**
   This criteria is actually assessed using the two-way transmission (T^2_C) of the threshold optical depth, which is computed using the equation

   $$T^2_C = e^{-2\tau_c}$$ Eq. 3.19

   and compared to the two-way transmission of the layer (T^2_{lay}) as estimated below:

   $$T^2_{lay} = 1 - S \times \gamma_{1064}$$ Eq. 3.20

   where S is the lidar ratio estimated as 25 sr for dense ice clouds.

The polarization gain ratio can then be derived using the ratios of the parallel and perpendicular background signals summed over the entire granule from all profiles within the granule file that contain these "dense ice clouds". Using a similar procedure, Liu *et al.* (2004) found polarization gain ratios for CALIPSO data that compared favorably with the values measured onboard by inserting a half-wave plate with its optical axis aligned at 22.5 degrees to the transmitted laser polarization direction into the optical path of the transmitter (Spinhirne *et al.* 1982) or the receiver (McGill *et al.* 2002).

The CATS depolarization purity at 1064 nm was measured in the lab at GSFC as greater than 150:1 before launch. However, the depolarization purity at 532 nm was measured at about 7:1. To improve the accuracy of CATS 532 nm depolarization measurements, this data must be corrected for this poor depolarization purity at 532 nm. The low depolarization purity at 532 nm, if uncorrected, causes a high bias in the perpendicular

backscatter and depolarization ratio at 532 nm, as well as a low bias in the 1064/532 backscatter color ratio.

A separate measurement of depolarization ratio is necessary to compare to the CATS measurements at both 1064 and 532 nm and estimate the PGR_2 term. The Cloud Physics Lidar (CPL; McGill *et al.* 2002) is an airborne elastic backscatter lidar system that flies aboard the NASA ER-2 high altitude aircraft and operates at 1064, 532, and 355 nm wavelengths. Depolarization is resolved using the 1064 nm channel and cloud optical properties are retrieved using the 1064 and 532 nm channels (McGill *et al.* 2003). CPL flew aboard the ER-2 during the month of February 2015 out of Palmdale, CA. During this time, the ER-2 flew under the ISS track on four occasions. One such flight on 22 February occurred during local nighttime hours and included observations of dense ice clouds along the ISS track. Figure 3.3 shows the ISS track in black, nearly parallel to the California coastline and the ER-2 track in red, both of which intercept ice clouds (light tan and blue colors). The ER-2 flew about a 30-minute segment below the ISS in which CPL and CATS collected near-coincident data.
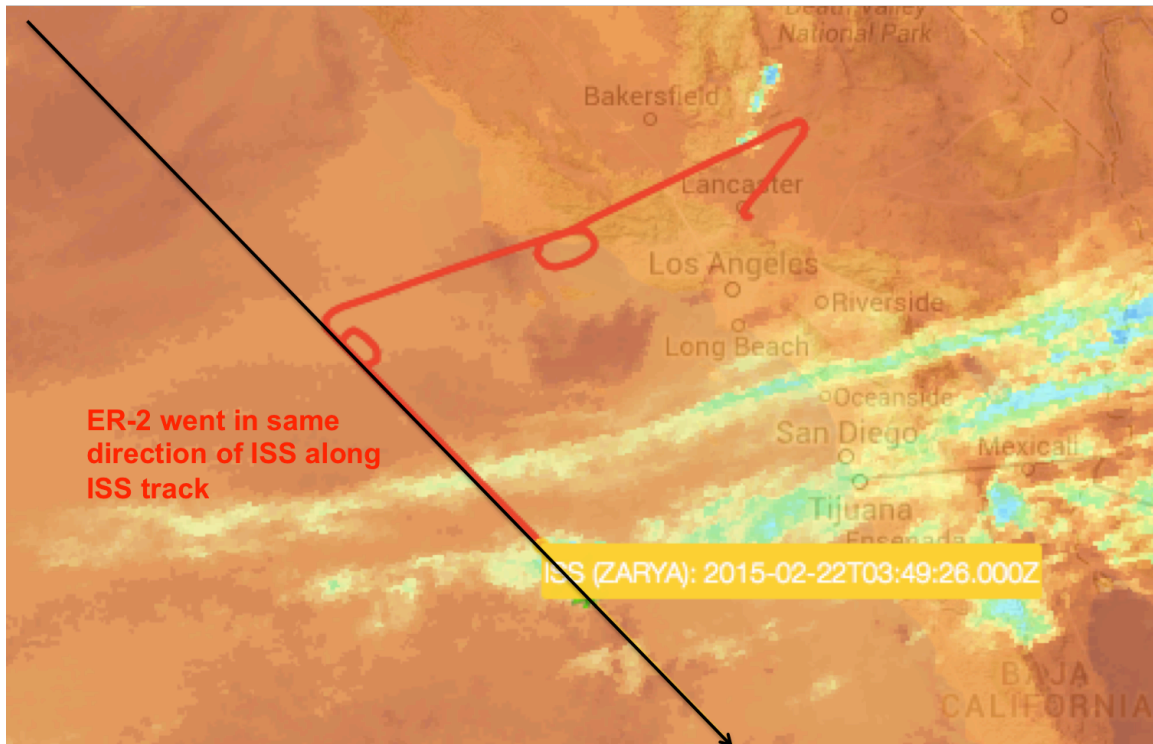

Figure 3.3. The tracks for the ISS and NASA ER-2 near the coast of California on 22 February 2015.

The time of closest coincidence occurs at 03:49:26 UTC near 32.0 degrees latitude as both instruments observe a dense ice cloud, as shown in the 1064 nm attenuated total backscatter data in in Figure 3.4 for both CPL (a) and CATS (b). The estimate of the second PGR term relies on two assumptions:
1) The spectral depolarization ratio, defined as the ratio of depolarization ratio at 1064 nm to depolarization ratio at 532 nm, is unity for dense ice clouds.

2) The CPL 1064 nm depolarization ratios and thus attenuated perpendicular backscatter measurements are accurate to within 3%, which is the estimated crosstalk between polarization detector channels measured for the CPL instrument.

Under these two assumptions, the CPL 1064 nm attenuated perpendicular backscatter should be equivalent to the CATS attenuated perpendicular backscatter measurements at both 532 and 1064 nm for dense ice clouds.
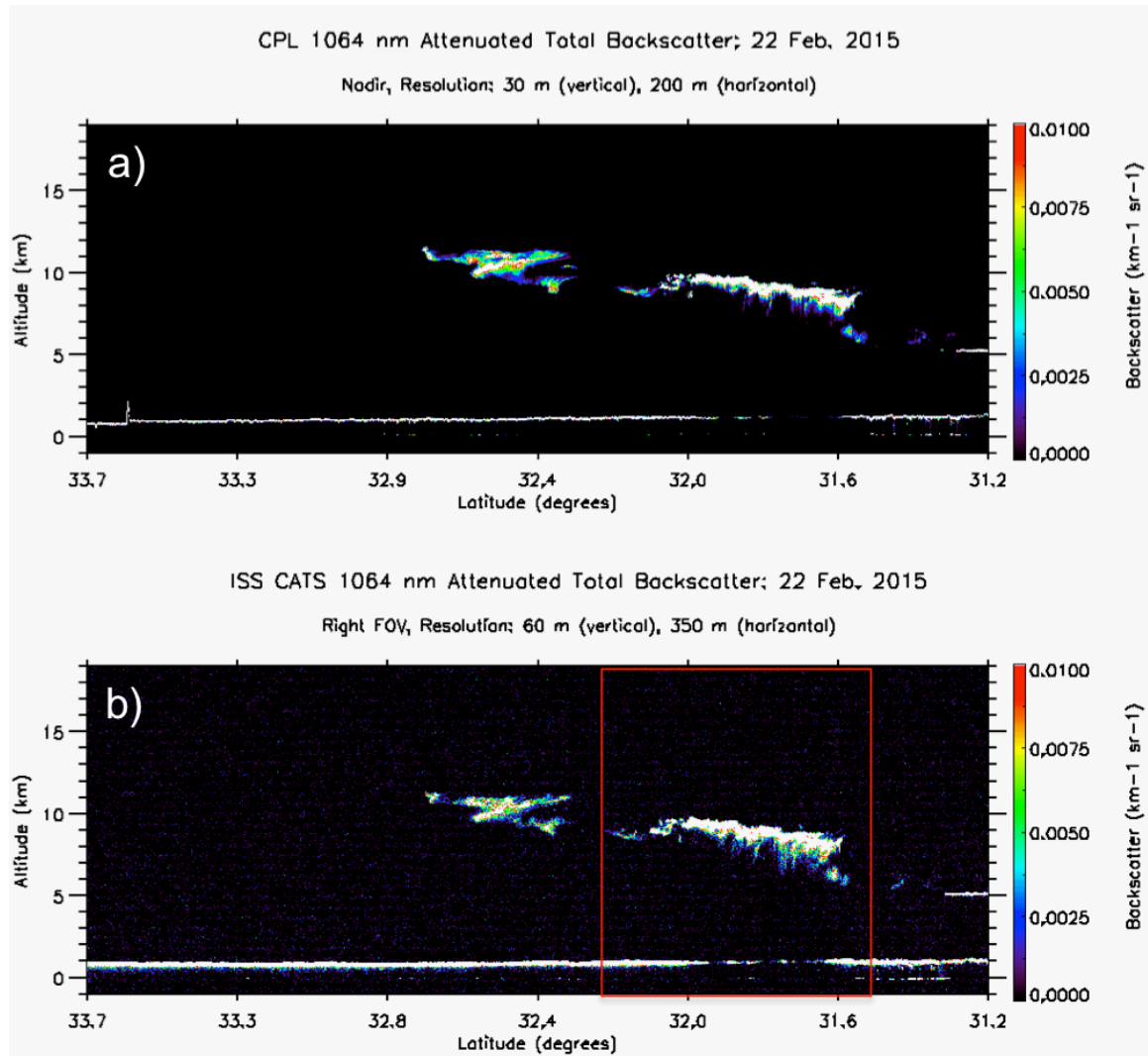

Figure 3.4. The 1064 nm attenuated total backscatter data for both CPL (a) and CATS (b) for the segment of the ER-2 flight along the ISS track on 22 February 2015. The red box denotes the profiles averaged to create Figure 3.3.

Figure 3.5 shows the mean attenuated perpendicular backscatter data for the profiles highlighted in the red box in Figure 3.4 for CPL 1064 nm (blue) and CATS RFOV 532 nm (green) and 1064 nm (red), after the CATS data has been normalized to Rayleigh (Section 3.2.5) and the PGR_1 term has been applied to the data. The CPL 1064 nm (blue) and CATS RFOV 1064 nm (red) profiles agree very well within the dense ice cloud.

However, the CATS RFOV 532 nm (green) profile is higher than both the 1064 nm profiles. A ratio of about 0.646 must be applied to the CATS 532 nm perpendicular backscatter profile (purple) to obtain a spectral depolarization ratio of 1.0 for the dense ice cloud. Thus, the PGR_2 term for RFOV 532 nm is 0.564. Similarly, the LFOV 532 nm PGR_2 term is computed using this near-coincident dataset.

The CATS Version 2-04 data release included these PGR terms, which significantly reduced the high biases in CATS 532 nm attenuated total backscatter and depolarization ratio. However, the statistical analysis of these parameters for cirrus clouds still yielded values 10-20% higher than those observed in the CATS 1064 nm data and over 10 years of CPL data at 1064 nm. This was largely due to uncertainties of as much as 15% in the computed PGR term. These values have been updated for V2-06 based on more statistical samples of coincident CPL and CATS data, reducing the PGR term uncertainties to 4.5% and thus minimizing the higher biases in the 532 nm attenuated total backscatter and depolarization ratio to 5-10%.
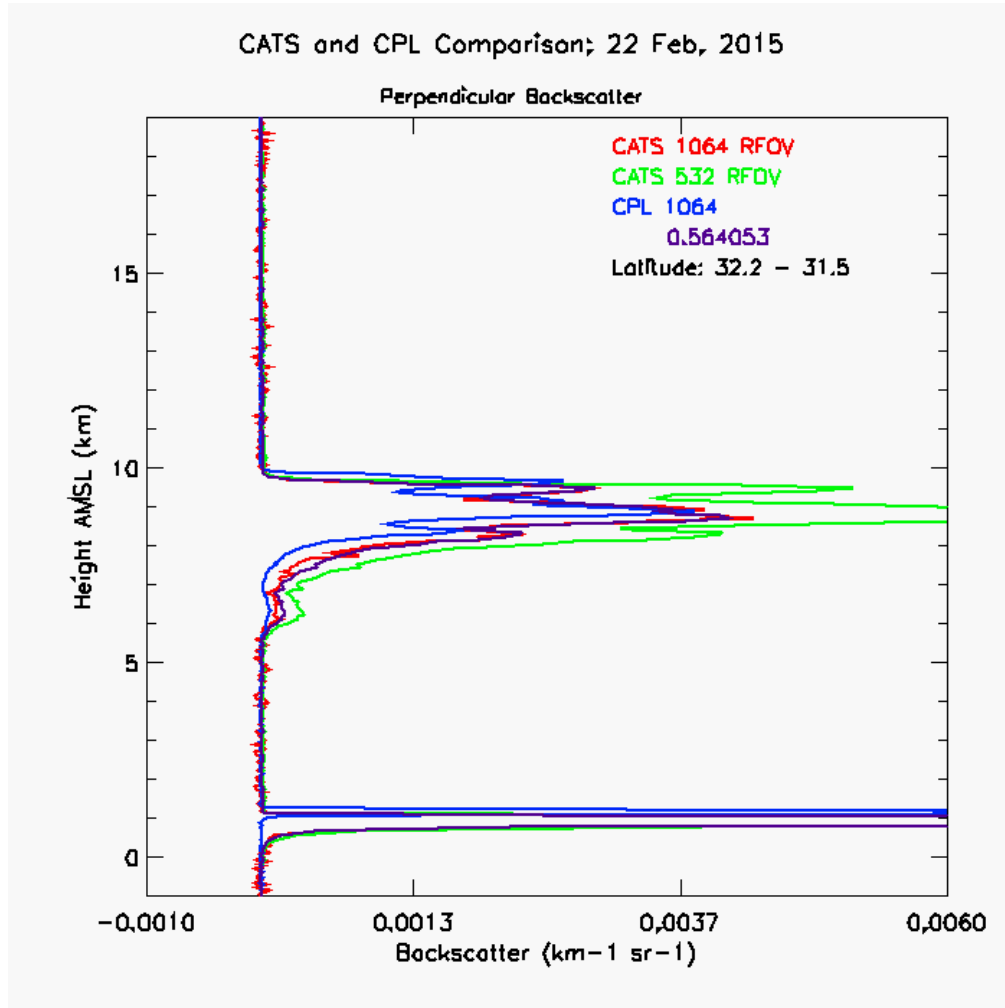

Figure 3.5. The mean attenuated perpendicular backscatter data for the profiles highlighted in the red box in Figure 3.4 for CPL 1064 nm (blue) and CATS RFOV 532 nm (green) and 1064 nm (red). The CATS RFOV 532 nm data multiplied by 0.646 is shown in the purple profile.

3.2.4 Stratospheric Scattering Ratios

The CATS calibration coefficients are computed by normalizing the normalized relative backscatter signal with respect to a modeled molecular backscatter signal over a set altitude range of 23 to 27 km. This altitude regime was selected because the CATS data frame is restricted to an upper limit of 28 km above mean sea level and tropical cirrus and volcanic plumes typically extend as high as 18 and 22 km respectively (although the later depends on the magnitude of each volcanic eruption). Although the aerosol loading in this region is generally less than would be found in the lower stratosphere, it is not free of aerosol contamination. Thus to accurately normalize the NRB signal to the molecular backscatter model, the aerosol loading must be characterized by a spatially and temporally varying ratio of total backscatter to molecular backscatter, referred to as the scattering ratio (R) and defined as:

$$R_\lambda(r) = \frac{\beta_{tot}(r)T^2_{tot}(r)}{\beta_M(r)T^2_M(r)T^2_{O_3}(r)}$$

Eq. 3.21

The 532 nm scattering ratios in the CATS calibration region are estimated using the CALIPSO V4 Level 1 data. Monthly HDF files are provided by the CATS LaRC team every 15 days that include the profiles of the 30-day mean, median, standard deviation, and error of the 532 nm scattering ratio between 22 and 28 km (180 m bins) and between 54 S and 54 N degrees latitude at 2 degree latitude increments. The scattering ratio varies by latitude and altitude, with values as high as 1.22 in the lower stratosphere (22 km) near the equator as shown in Figure 3.6 for February 2007. The most recent monthly HDF file is used to derive the 532 scattering ratio in the CATS calibration algorithm and to estimate the 1064 nm scattering ratio using the equation:

$$R_{1064}(r) = 1 + \frac{\chi_{1064}(\beta_{M,532}(r)R_{532}(r)-1)}{\beta_{M,1064}(r)}$$

Eq. 3.22

where β_m at 532 and 1064 nm are the Rayleigh backscatter model computed as shown in section 3.2.2 and χ_{1064} is the backscatter color ratio defined as (Hair *et al.* 2008):

$$\chi_{1064} = \frac{\beta_{P,1064}(r)}{\beta_{P,532}(r)} = 0.40$$

Eq. 3.23

The application of the 532 and 1064 nm scattering ratios will be discussed in the next section.
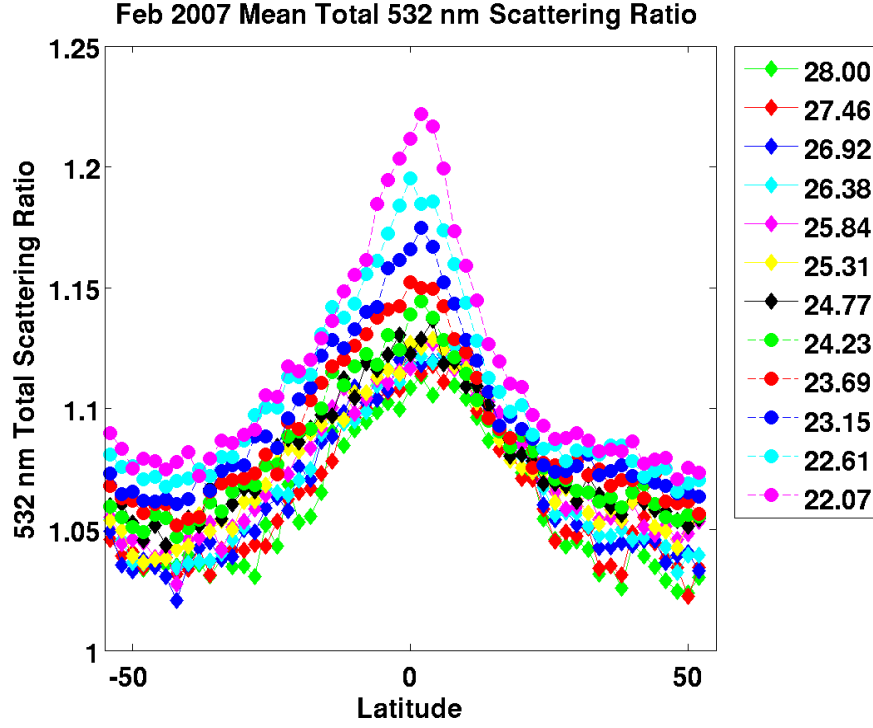
Figure 3.6. The mean CALIPSO scattering ratios for February 2007 for various altitude bins between 22 and 28 km.

3.2.5 Calibration at 532 and 1064 nm Wavelengths

Once the ozone transmission, Rayleigh scattering, polarization gain ratio, and stratospheric scattering ratios have been computed, the next step in the calibration of CATS data is to apply these parameters to the CATS data. The perpendicular NRB at raw resolution is multiplied by the PGR at both the 532 and 1064 nm wavelengths. The total NRB signal at both wavelengths, derived by adding the perpendicular and parallel signals (NRB_{perp} + NRB_{par}), is then divided by both the ozone transmission and stratospheric scattering ratio of the corresponding wavelength as a function of height. Note that the stratospheric scattering ratio is applied to the data only in the calibration routine and not to the data products reported in the L1B files. This CATS calibration-ready NRB signal is averaged typically to 4 minute segments at night and 46 minute segments during daytime operation to create mean profiles of the calibration-ready NRB in the calibration region of 23-27 km.

The 532 and 1064 nm CATS calibration coefficient (C) profiles at each segment are derived by normalizing the mean calibration-ready NRB signal (β_{CN}) to the mean molecular backscatter signal ($\beta_M T^2_M T^2_O$) in the calibration region (Russell *et al.* 1979, Del Guasta 1998, McGill *et al.* 2007, Powell *et al.* 2009), as shown in the equation below:

$$C_\lambda(r) = \frac{\left[\dfrac{NRB(r)}{T^2_O(r)*R}\right]}{\beta_M(r)T^2_M(r)} = \frac{\beta_{CN}(r)}{\beta_M(r)T^2_M(r)}$$

Eq. 3.24

24

The final calibration coefficient at each segment (typically 4 mins for night, 46 mins for day) is simply the mean of the calibration coefficient profile from 23 to 27 km. For nighttime conditions, this provides 7-20 calibration coefficients per granule compared to only 1 per daytime granules. The final calibration constant is computed by either calculating the mean of the calibration coefficient data points in each granule or by a linear fit (pre-determined by Mode and SNR), as shown in Figure 3.7. If the calibration coefficient at a specific segment does not stay within threshold values, it is not used in the average or fit. In daytime granules, it is possible that no calibration coefficients meet these threshold values. When this occurs, a default calibration constant is set for the entire granule based on historical data and/or manual normalization to the modeled Rayleigh signal. It should be noted that the CATS 1064 nm calibration constant is also derived using the 532 nm signal and backscatter from ice clouds, similar to CALIPSO at 1064 nm (Vaughan *et al.* 2010), but is not used operationally. This technique is only used for research purposes by the CATS team to compare the two 1064 nm calibration techniques.
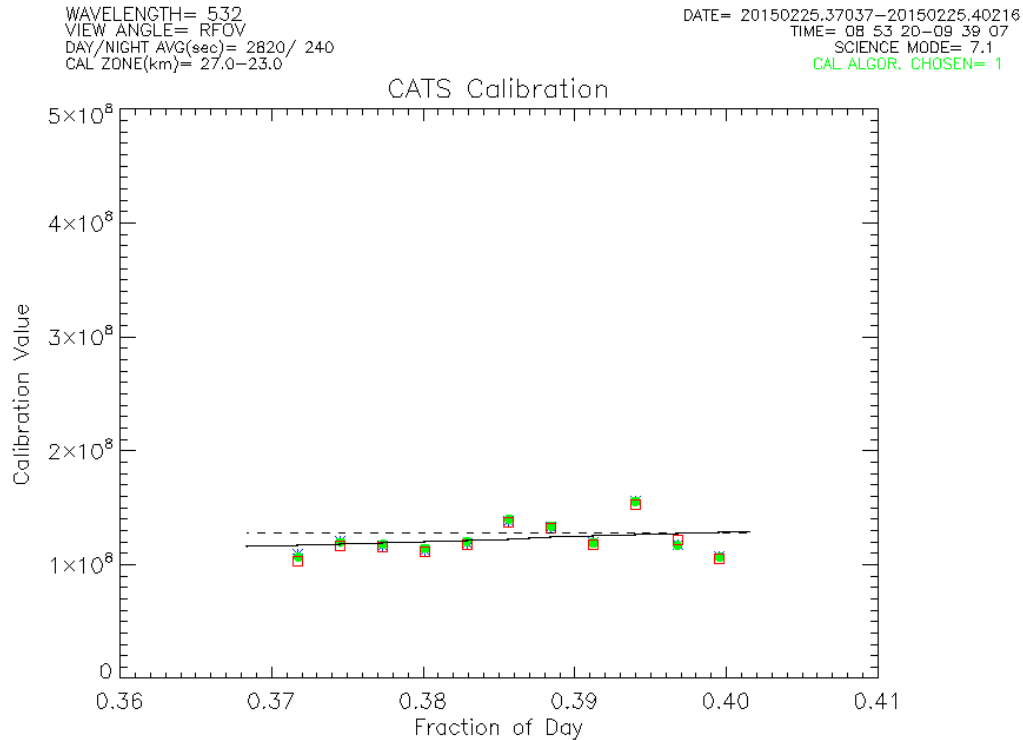
Figure 3.7. The 532 nm RFOV calibration coefficients (green) computed during a nighttime granule on 25 Feb. 2015 around 08 UTC. The linear fit apply to the data (black solid line) is the calibration constant applied to the backscatter data. The dotted black line is the default value that would be used if the calibration coefficients were not within thresholds.

There are two main types of error in the CATS calibration constant: systematic error and random error. The systematic error in the CATS calibration constant is derived similar to

the systematic error in the CALIPSO calibration (Reagan *et al.* 2002, Powell *et al.* 2009). This systematic error has four sources:

1) Error in the stratospheric scattering ratios provided by CALIPSO (ΔR)
2) Error in the molecular backscatter coefficient derived from the GMAO data ($\Delta\beta_M$)
3) Error in the background transmission from molecules and ozone in the atmosphere (ΔT^2)
4) Errors induced by non-ideal optical performance of the CATS lidar system ($\varepsilon+a+c$). This error can be effectively reduced by instrument corrections.

Thus the total relative systematic error in the calibration constant is defined as:

$$\left(\frac{\Delta C}{C}\right)^2_{sys} = \left(\frac{\Delta R}{R}\right)^2 + \left(\frac{\Delta\beta_M}{\beta_M}\right)^2 + \left(\frac{\Delta T^2}{T^2}\right)^2 + (\varepsilon + a + c)^2 \qquad \text{Eq. 3.25}$$

and is estimated to be 5%. The latter 3 terms are constant over time, but the error in the stratospheric scattering ratio is computed for each monthly HDF file described in section 3.2.4 and varies by season and volcanic activity.

The random error in the calibration constant results from normalizing the 532 and 1064 nm signals to the modeled molecular signal and is dominated by noise in the lidar signal. This error can be estimated by determining the variability of the intermediate calculated coefficients (C_i) computed at each averaging segment (4 minutes for nighttime data) that are averaged to produce the final calibration constants. As described above, the intermediate calibration coefficients are computed for each vertical bin inside the calibration zone after averaging horizontally to get a profile for the mean calibration segment with 66 vertical bins (60 m bin size). Thus the equation for computing the random error is:

$$(\Delta C)_{ran} = \frac{stdev(C_i)}{\sqrt{N}} \qquad \text{Eq. 3.26}$$

Typical values of the random error for the CATS calibration constant are 5-7% at 532 nm and 1064 nm. Thus the total error is derived using the equation:

$$\left(\frac{\Delta C}{C}\right)^2_{tot} = \left(\frac{\Delta C}{C}\right)^2_{sys} + \left(\frac{\Delta C}{C}\right)^2_{ran} \qquad \text{Eq. 3.27}$$

The total error in the CATS calibration constants at 532 and 1064 nm is estimated at 5-10%.

3.2.6 Attenuated Backscatter

The primary product in the CATS Level 1B data is the calibrated backscatter, known as the attenuated total backscatter (ATB or γ), which has units of km^{-1} sr^{-1} and is defined as:

$$\gamma(r) = \frac{NRB(r)}{C} = [\beta_P(r) + \beta_M(r)]T_M^2(r)T_P^2(r) \qquad \text{Eq. 3.28}$$

Where C is the calibration constant determined using the algorithm outlined in section 3.2.5. The attenuated backscatter is also computed for the perpendicular and parallel signals using the same calibration constant as the total signal. The primary sources of uncertainty in the CATS attenuated backscatter signal are the calibration constant and signal noise. Thus if the calibration constant is accurate, the CATS ATB profiles should compare favorably with the Rayleigh backscatter model, as shown in Figure 3.8.
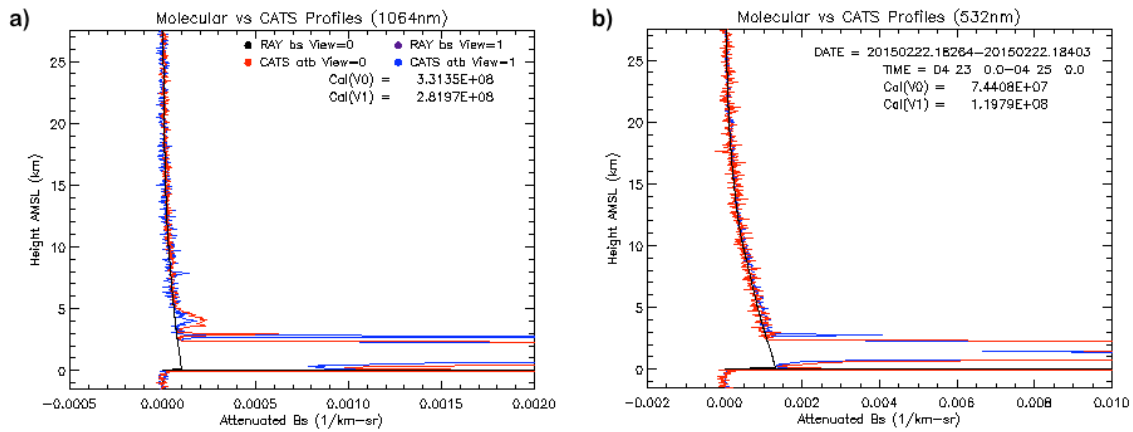

Figure 3.8. The 2-minute mean profiles of the CATS ATB signal at 1064 (a) and 532 (b) for both the RFOV (red) and LFOV (blue) for data on 22 February 2015. These profiles compared favorably with the modeled Rayleigh profiles (black) for each wavelength and field-of-view, demonstrating that the data is well calibrated.

4.0 Overview of Vertical Feature Mask Algorithms

There will be two CATS Level 2 products and therefore two vertical feature mask algorithms. The CATS Heritage Level 2 products (Rodier *et al.* 2015) are created when CATS L1B files are run through the CALIPSO L2 algorithms to provide continuity in the algorithms used for the lidar climate record. These algorithms are outlined in detail in the CALIPSO ATBD (Vaughan *et al.* 2005) and in numerous publications (Hu *et al.* 2009, Liu *et al.* 2009, Omar *et al.* 2009, Vaughan *et al.* 2009, Young *et al.* 2009), so they will not be discussed in this document. The CATS Operational L2 products are produced using new operational CATS L2 algorithms, which will include new capabilities that correspond to new instrument technology such as spectral depolarization and high resolution 1064 nm data. These CATS Operational vertical feature mask algorithms will be outlined in the sections to follow in future releases.

The CATS operational vertical feature mask algorithms were designed based on CPL and CALIPSO algorithms, as well as the performance of the CATS L1B data products. The CATS Mode 7.1 backscatter signal is more robust at 1064 nm than 532 nm. The minimum detectable backscatter for Mode 7.1 for cirrus clouds at 12 km and horizontal resolution of 5 km are shown in Table 4.1. The CATS 1064 nm signal at both night and day has a much lower minimum detectable backscatter than 532 nm. This is attributed to the fact that the laser is outputting more energy at 1064 nm (1.40 mJ compared to 0.88 mJ at 532 nm). Additionally, both CATS wavelengths in daytime conditions have lower signal to noise ratio (SNR) and higher minimum detectable backscatter for Mode 7.1 than nighttime. The poorer performance during daytime is due to solar background noise, typically an issue for daytime operation of any lidar.

Table 4.1. Minimum Detectable Backscatter: CATS Mode 7.1, 5km (hori.) and 60 m (vert.), for cirrus clouds at 15 km

Data	Type	Backscatter (km-1 sr-1)
CATS 1064	Night	1.80E-4 ± 0.49E-4
CATS 532	Night	1.00E-3 ± 0.54E-3
CATS 1064	Day	7.60E-3 ± 0.24E-3
CATS 532	Day	2.20E-2 ± 0.35E-2

Unlike the Mode 7.1 data, where the 532 and 1064 nm signals are comparable, the Mode 7.2 532 and 1064 nm signals are very different. Mode 7.2 data at 532 nm is noisy due to issues with stabilizing the seeded laser (laser 2). Since the frequency stability is poor on laser 2, it is not aligned properly with the CATS etalon causing very weak signal transmission. Unfortunately we do not have the necessary controls to fix the problem, so we recommend averaging the data to *at least* 5 km (roughly 14 raw 20 Hz profiles) when analyzing the 532 nm data.

Table 4.2. Minimum Detectable Backscatter: CATS Mode 7.2, 5km (hori.) and 60 m (vert.), for cirrus clouds at 15 km

Data	Type	Backscatter (km-1 sr-1)
CATS 1064	Night	5.00E-5 ± 0.77E-5
CATS 532	Night	1.60E-2 ± 0.84E-3
CATS 1064	Day	1.30E-3 ± 0.24E-3
CATS 532	Day	3.80E-2 ± 1.05E-2

Due to the signal transmission issues at 532 nm, laser 2 was thermally tuned to increase the laser energy at 1064 nm to 2 mJ per pulse. Thus the 1064 nm signal in mode 7.2 is very robust, with higher SNR and lower minimum detectable backscatter (5.00E-5 km^{-1} sr^{-1}) than the Mode 7.1 data (Table 4.2). Thus, the 1064 nm data is utilized heavily in the vertical feature mask algorithms and for any analysis that is wavelength-independent (i.e.

layer detection, relative backscatter intensity). More assessment of the L1B data product performance, including comparisons with CPL and CALIPSO, will be presented in future publications.

4.1 Atmospheric Layer Detection

The CATS layer detection is performed following the methodology described in the CALIOP Algorithm Theoretical Basis Document (ATBD; Vaughan et al. 2005). It is a threshold-based layer detection method that uses the 1064 nm attenuated scattering ratio, unlike the CALIPSO algorithm that uses 532 nm. A threshold for each profile is calculated from range-variant and range-invariant sources of noise. For an exhaustive description of the details of the layer detection method please see Section 3.2 of the CALIOP ATBD and Vaughan et al. 2009 (Fully automated analysis of space-based lidar data: an overview of the CALIPSO retrieval algorithms and data products).

There are some key differences between the CALIOP layer detection algorithm and CATS layer detection algorithm. The CATS algorithm only performs layer detection at a single horizontal resolution of 5km horizontal (60m vertical). The CALIOP algorithm will successively run the profile scanner at coarser and coarser horizontal resolutions ranging from 5 km to 80 km in order to detect increasingly tenuous layers (Vaughan et al. 2009). The CATS false positive rejection scheme, like CALIOP's, utilizes the feature-integrated backscatter (FIB) of layers as a criterion for rejecting layers. However, unlike the CALIOP scheme, the CATS scheme also takes into account the horizontal persistence of layers (Vaughan et al. 2009). If both a horizontal persistence test and a FIB test are failed then the layer is rejected. The FIB test checks whether the calculated FIB of a layer is less than a certain threshold.

The CATS horizontal persistence test is the more complex part of the false positive rejection algorithm. For a given layer detected within a single profile, a box is created around that layer that includes adjacent profiles. If there are a certain number of layers detected within the box, within adjacent profiles, the layer is considered horizontally persistent and passes the test. If there are less a certain number of layers detected within the box, within adjacent profiles, then the layer fails the horizontal persistence test. The diagram in Figure 4.1 demonstrates how the horizontal persistence test works. Each column represents a lidar profile. Each box represents a lidar bin. A blue box represents that a layer was detected in that bin. In any given profile, consecutive vertical blue bins are considered a single layer. Consider the layer in the center of the red box. The red box represents the region used to test for horizontal persistence. In this scenario, the layer in the red box would fail the horizontal persistence test because there are no other detected layers within the red box. The size of the red box and the number of neighboring layers required are adjustable in a configuration file for the algorithm.
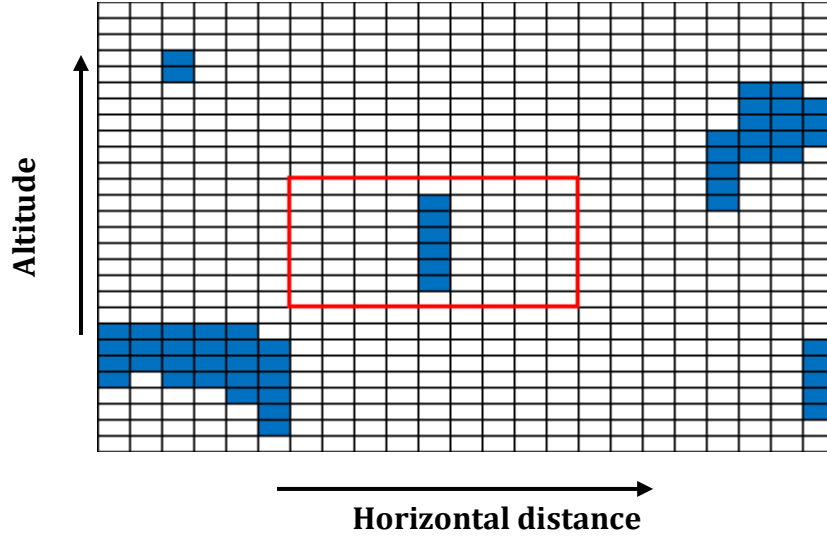
Figure 4.1. Conceptual diagram of horizontal persistence test for layer false positive rejection.

For CATS, the layer detection is performed using the 1064 nm backscatter signal for two main reasons:

1) The CATS 1064 nm minimum detectable backscatter is lower than 532 nm, making it a better option for more accurate layer identification (Tables 4.1 and 4.2).

2) For absorbing aerosols, the absorption optical thickness increases with decreasing wavelength. This effect reduces the backscattered signal at 532 nm with respect to 1064 nm, such that the 532 nm backscatter is not sensitive to entire vertical extent of the aerosol layer [*Torres et al.*, 2013; *Jethva et al.*, 2014; *Liu et al.*, 2014]. Because the 1064 nm wavelength is only minimally affected by aerosol absorption, the vertical extent of the absorbing aerosol layer is more fully captured from 1064 nm backscatter profiles rather than those from 532 nm.

Since CATS can detect the full vertical extent of the aerosol layer above the cloud, it is extremely important that the algorithm also distinguish these as two separate layers.

The CATS layer detection algorithm includes a routine to identify clouds embedded within aerosol layers. The CALIOP algorithm does not include this capability. The CATS cloud-embedded-in-aerosol routine, or CEAL, only considers layers below a certain, configurable altitude. This altitude (operationally set to 6 km) is based on the assumption that the large majority of clouds embedded within aerosol layers will be confined to the lower part of the troposphere. For each layer under consideration a threshold is calculated from the median attenuated backscatter of the layer and a threshold factor. The threshold is equal to the median attenuated backscatter of the layer multiplied by the threshold factor. The threshold is compared to the maximum attenuated backscatter of the layer. If the maximum is greater than the threshold, then a cloud boundary search is conducted as follows:

1. From the bin of the maximum, a search begins upward (altitude-wise) bin by bin until 2 conditions are met: the attenuated backscatter value is below the threshold

and the slope of backscatter (with respect to search direction) reverses a certain number of times.

2. This bin is assigned as the top of the cloud layer.
3. The base of the cloud layer is found analogously by searching downward from the maximum until the same conditions are met.

CEAL can identify up to 2 clouds within a single combined cloud-aerosol layer. Regions of the original candidate layer outside of the identified clouds are considered separate (presumably aerosol, although no typing is performed by the CEAL routine) layers as long as they have a certain minimum vertical thickness, otherwise they are absorbed into the adjacent cloud layers.
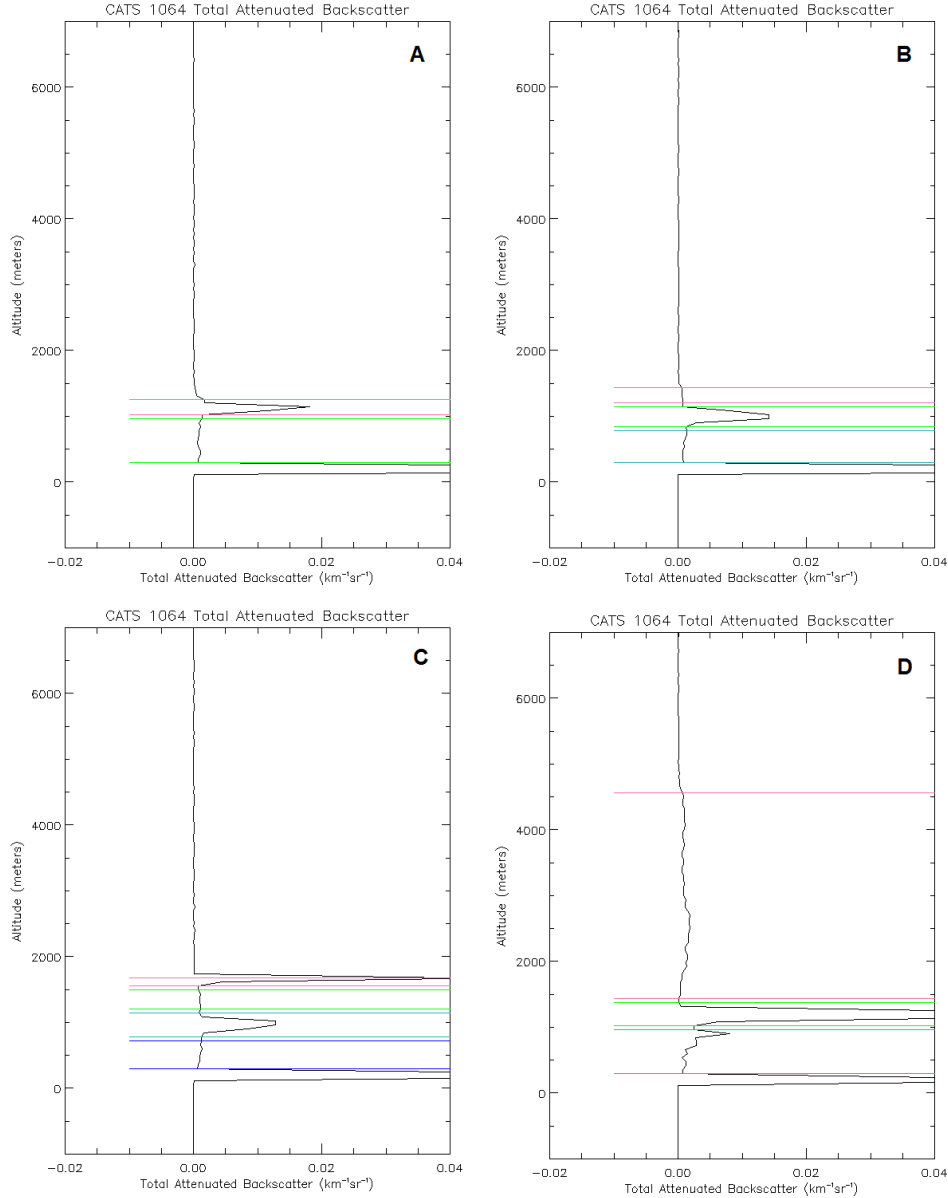
Figure 4.2. Four examples of the CEAL routine applied to profiles containing clouds embedded in aerosol layers. The horizontal, colored lines denote the vertical boundaries of the layers identified by the CEAL routine (first layer is pink, second layer is green, third layer is light blue, fourth layer is dark blue). Each uniquely identified layer gets a unique color. In other words the colors are simply a visual aid for the purpose of making the layers easier to see.

Figure 4.2 shows four examples of the CEAL routine applied to different profiles containing clouds embedded in aerosol layers. Figure 4.2A shows a PBL cloud at the top of an aerosol layer. Figure 4.2B shows a cloud embedded within a PBL aerosol layer. Figure 4.2C shows two clouds embedded in PBL aerosol. Figure 4.2D shows a cloud embedded in a geometrically thick dust layer. In all these examples the sub-layers were all identified as a single layer before applying the CEAL routine.

4.2 Cloud-Aerosol Discrimination

Retrievals of cloud and aerosol optical properties are the staple of the CATS Level 2 data products. The accuracy of these products depends primarily on the accuracy of the extinction-to-backscatter ratio, also known as the lidar ratio, which in turn requires accurate classification of atmospheric layers. Once the layers have been identified, the next step in the CATS L2 processing is cloud– aerosol discrimination (CAD). The CATS CAD algorithm is a multidimensional probability density function (PDF) technique that is based on the CALIPSO algorithm (Liu et al. 2009). The PDFs were developed based on CPL measurements obtained during over 11 field campaigns and 10 years. Table 4.3 shows the dates and locations of all 11 projects. In total, over 1.6 million cloud layers and nearly 1.8 million aerosol layers were included in the dataset, which was limited to the top layer in any given profile. CPL algorithms classify atmospheric layers into four categories: ice clouds, liquid water clouds, PBL aerosols, and elevated aerosols. Frequency plots of layer-integrated attenuated backscatter at 1064 nm, layer-integrated depolarization ratio at 1064 nm, layer-integrated attenuated backscatter color ratio, and mid-layer altitude are useful for discriminating clouds and aerosols. Nearly all layers with a layer-integrated attenuated backscatter at 1064 nm greater than 0.03 sr^{-1} are either liquid water clouds or ice clouds. Also, features with a mid-layer altitude greater than 8.0 km and layer-integrated depolarization ratio at 1064 nm greater than 0.35 are ice clouds. Finally, features with a layer-integrated attenuated backscatter color ratio greater than 1.0 are typically clouds, with the exception of some elevated aerosols (red) that are likely large dust particles.

Differences in cloud and aerosol optical and physical properties are used to classify layers as a cloud or aerosol. The algorithm is driven by a probability function where PDF_C and PDF_A are the multidimensional PDFs, respectively, for clouds and aerosols. For CATS, there are three PDFs used in each mode with attributes i. The performance of the CATS CAD is limited to the amount of overlap between clouds and aerosols in the PDFs. Adding more attributes, or dimensions, to the PDFs will result in a smaller overlap and better classification of clouds and aerosols (Liu et al. 2004). Since the CAD algorithm is employed after the computation of Level 1B data and before retrievals of optical properties (i.e. extinction, optical depth), we are limited in cloud/aerosol properties to use in the PDFs. Measured cloud/aerosol properties available include layer altitudes and thickness, attenuated backscatter, depolarization, and attenuated backscatter color ratio (1064/532-nm). Ancillary data, such as mid-layer temperature can also be utilized. Additionally, data quality, computing time, and ancillary data rates must be considered when selecting attribute dimensions for operational PDFs.

Table 4.3. CPL Data used to produce the CATS Operational CAD PDFs.

Project	Dates	Latitude Range
THORPEX-Atlantic	Nov - Dec 2003	32 to 53
CC-VEX	Jul - Aug 2006	23 to 39
CLASIC	Jun 2007	28 to 40
TC4	Jul - Aug 2007	0 to 39
ATTREX11	Nov 2011	6 to 28
IceAx	Apr 2012	31 to 80
WAVE	Sep 2012	33 to 48
PODEX	Jan - Feb 2013	28 to 38
ATTREX13	Feb - Mar 2013	-10 to 33
HS3	Aug - Sep 2013	10 to 38
SEAC4RS	Aug - Sep 2013	15 to 49

The attributes of the operational CATS PDFs depend on the CATS mode of operations. For Mode 7.1, the dimensions are the layer-integrated attenuated backscatter (γ') at 532 nm, the layer-integrated attenuated backscatter color ratio (χ'), layer-integrated 1064 nm depolarization ratio (δ'), and the midlayer altitude (Z_{top}). The probability score (-100 to 100) for these PDFs are computed using equation 4.1:

$$P^{PDFi} = \frac{PDF_C^i - PDF_A^i}{PDF_C^i + PDF_A^i} \times 100$$

Eq. 4.1

For Mode 7.2, the same variables are used except the layer-integrated attenuated backscatter color ratio is replaced with layer thickness (ΔZ) and the layer-integrated attenuated backscatter at 1064 nm is utilized because the 532 nm signal strength is weak (section 4.0).

The main objective of the CATS operational CAD algorithm is to compute the CAD score, which is an integer value ranging from -10 to 10 for each atmospheric layer. Table 4.4 illustrates that the sign of the CAD score identifies a layer as either cloud (positive) or aerosol (negative), while the magnitude of the CAD score represents the confidence in our classification. A value of 10 indicates complete confidence that the layer is a cloud, while -10 indicates the accurate classification of an aerosol layer. When the CAD score equals 0, the layer is just as likely to be a cloud as it is an aerosol, and thus the classification is undetermined. If the optical and physical properties of the layer are considered invalid for clouds and aerosols, these layers are assigned a CAD score of -99.

Table 4.4. The interpretation of the CATS CAD Score reported in the L2 data products.

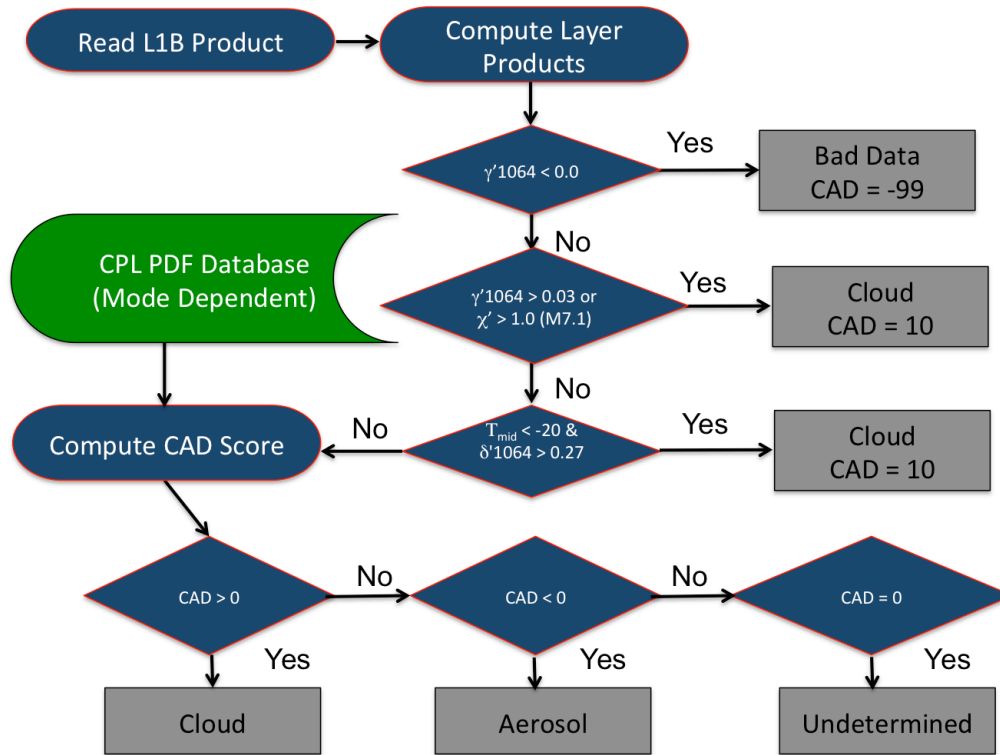| Layer Type	CAD Score
Cloud	1 to 10
Aerosol	-10 to -1
Undetermined	0
Bad Data	-999

CATS CAD Algorithm: Version 1

Figure 4.3. A flowchart of the CATS version 1 operational CAD algorithm.

The flow of the CATS operational CAD algorithm is shown in Fig. 4.3. First, the CAD algorithm computes the layer products, which include γ' at both 532 and 1064 nm, χ', δ' at 1064 nm, Z_{mid}, and T_{mid}. Before classifying an atmospheric layer, the operational CAD algorithm declares the layer invalid if the γ'_{1064} is less than 0. Then, the algorithm identifies high confidence cloud layers (CAD score = 10) as those layers with a γ'_{1064} is greater than 0.03 sr^{-1}. For Mode 7.1 data, high confidence cloud layers are also classified if the χ' is greater than 1.0. No aerosols layers identified in the CPL data were found above these threshold values. Very few aerosol layers are also found below temperatures of -20 C, with the exception of volcanic plumes. In the first year of CATS data, volcanic plumes observed has low depolarization ratios (less than 0.15), typical of plumes that consist mostly of SO_2. If the T_{mid} is less than -20 C and the δ'_{1064} is greater than 0.27, the layer is classified as a high confidence cloud (most likely made of ice particle) and the CAD score is set to 10.

If a layer does not meet any of these criteria, then the PDF technique is used to determine whether the layer is an aerosol or cloud. The PDFs are employed to compute the CAD score that is reported within the CATS L2 products using equation 4.2.

$$CAD = \frac{\left(\dfrac{P^{PDF1}}{100}\right)T^{PDF1} + \left(\dfrac{P^{PDF2}}{100}\right)T^{PDF2} + \left(\dfrac{P^{PDF3}}{100}\right)T^{PDF3}}{T^{PDF1} + T^{PDF2} + T^{PDF3}} \times 10$$

Eq. 4.2

Unlike the CALIPSO CAD algorithm, the CATS version 1 operational CAD method does not test for horizontally oriented ice crystals. In initial assessments, clouds containing high amounts of horizontally oriented ice crystals are not observed as commonly in the CATS data as in the CALIPSO data, possibly do to the orbital environment of the ISS. More testing will be conducted to ensure that these phenomena are infrequent, and the CATS CAD algorithm will be adjusted accordingly in future versions.

Initial testing of the CATS CAD algorithm detected several biases that are corrected once the steps in Figure 4.3 are complete. CATS has observed several volcanic plumes during its first year on orbit. These plumes are typically located in the upper troposphere or lower stratosphere, and have weak backscatter and depolarization signals. Thus, any layer with a base altitude greater than 18 km is classified as an aerosol and assigned a CAD score of -10, since the ISS orbit does not travel to latitudes higher than 51 degrees where CATS would observed Polar Stratospheric Clouds (PSCs). Additionally, any layer with a base altitude greater than 10 km, δ'_{1064} less than 0.25, and γ'_{1064} less than 0.01 sr^{-1} is also classified as an aerosol layer (CAD score = -8). The CAD algorithm also has difficulty distinguishing optically thin dust plumes lofted into the mid to upper troposphere from optically thin ice clouds, since both these layers are depolarizing. To account for this bias, any cloud layer with a γ'_{1064} less than 0.03 sr^{-1}, δ'_{1064} is greater than 0.20 and T_{mid} greater than 0 C is reclassified as an aerosol layer with a CAD score of -5. Also, a CAD score of 3 is assigned to any aerosol layer with a CAD score of greater than -5 and a T_{mid} less than -20 C. These corrections will not fix every misidentified layer in the CAD algorithm, and further testing will likely yield other biases, but they do correct a majority of the known biases in the algorithm. The CATS team will continue performance assessments of the CAD algorithm and update future versions to improve the accuracy of cloud and aerosol identification.

4.3 Cloud Phase

The CATS Cloud Phase (CP) Algorithm, much like the CAD algorithm, employs differences in optical and physical properties of ice and liquid water clouds to classify layers as a ice clouds or water clouds. The CPL measurements used to create the CAD PDFs obtained during over 11 field campaigns and 10 years (Table 4.3) provide a statistical basis for the Cloud Phase Algorithm. According to the CPL data, cloud layers

with a mid-layer altitude greater than 8.0 km and layer-integrated depolarization ratio at 1064 nm greater than about 0.25 are ice clouds. However, mid-layer altitudes of cirrus clouds depend on many variables, including season and geographic location. Hu *et al.* (2009) and Yorks *et al.* (2011a) demonstrate the utility of using layer-integrated depolarization ratio and cloud temperature for distinguishing liquid water clouds from ice clouds. Figure 2 from Yorks *et al.* (2011a), shown below in Figure 4.4, outlines layer-integrated volume depolarization ratio (a) and mid-layer temperature (b) thresholds for ice clouds (red) and liquid water clouds (blue).
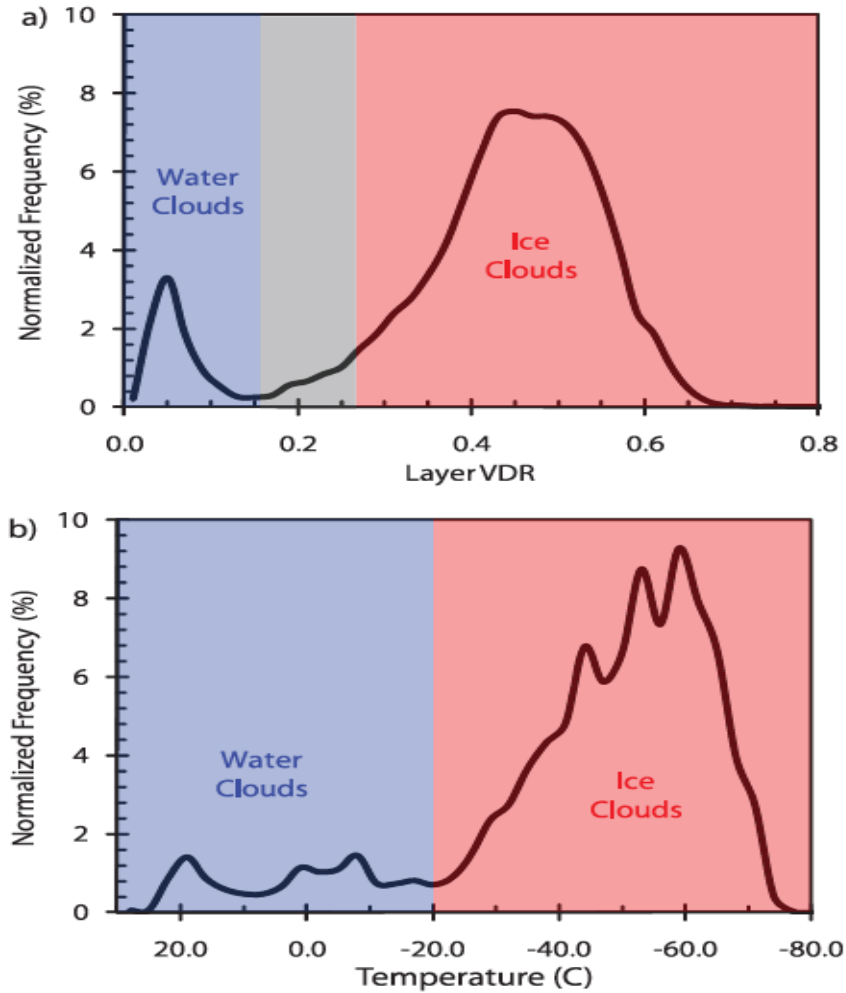

Figure 4.4. CPL data from 2003 to 2007 was used to create frequency distributions of layer-integrated volume depolarization ratio (a) and mid-layer temperature (b) for cloud layers (Yorks *et al.* 2011a, Figure 2).

The flow of the CATS operational CP algorithm is shown in Fig. 4.5. The CPL algorithm uses the layer product and CAD as input, and unlike the CAD algorithm, is the same for both Modes 7.1 and 7.2. First, high confidence liquid water clouds are classified if the cloud layer has a T_{mid} greater than 0 C and high confidence ice clouds are identified as cloud layers with a T_{mid} less than -20 C. These ice clouds and liquid water clouds are assigned a CP score of 10 and -10, respectively. Next, the CP algorithm identifies high confidence ice cloud layers as those layers with a δ'_{1064} is greater than 0.25 or T_{mid} less

36

than -10 C (CP Score = 9). High confidence liquid water clouds are classified if the cloud layer has a δ'_{1064} is less than 0.15 (CP Score = -9). The remaining layers are determined to have lower confidence cloud phase (yellow box) and are assigned a CP Score with an absolute value of 7 or less. For these cloud layers, additional parameters and invoked to help differentiate cloud phase, such as layer thickness (ΔZ) and the layer-integrated attenuated backscatter at 1064 nm. Any remaining cloud layer with a γ'_{1064} greater than 0.08 sr^{-1} is considered a liquid water cloud. Additionally, to identify optically thick but physically thin stratus layers that consist of liquid water clouds, a cloud layer with a γ'_{1064} greater than 0.03 sr^{-1} and ΔZ less than 1.0 km are classified as liquid water clouds. Figure 4.6 shows the frequency plot for layer physical thickness using the CPL data described in section 4.2. Less than 1% of liquid water clouds identified by CPL have a physical thickness greater than 1.5 km. Thus, any remaining cloud layer with a ΔZ greater than 1.5 km and T_{mid} less than 0 C are classified as ice clouds. Any cloud layers that do not meet these criteria are classified as undetermined.
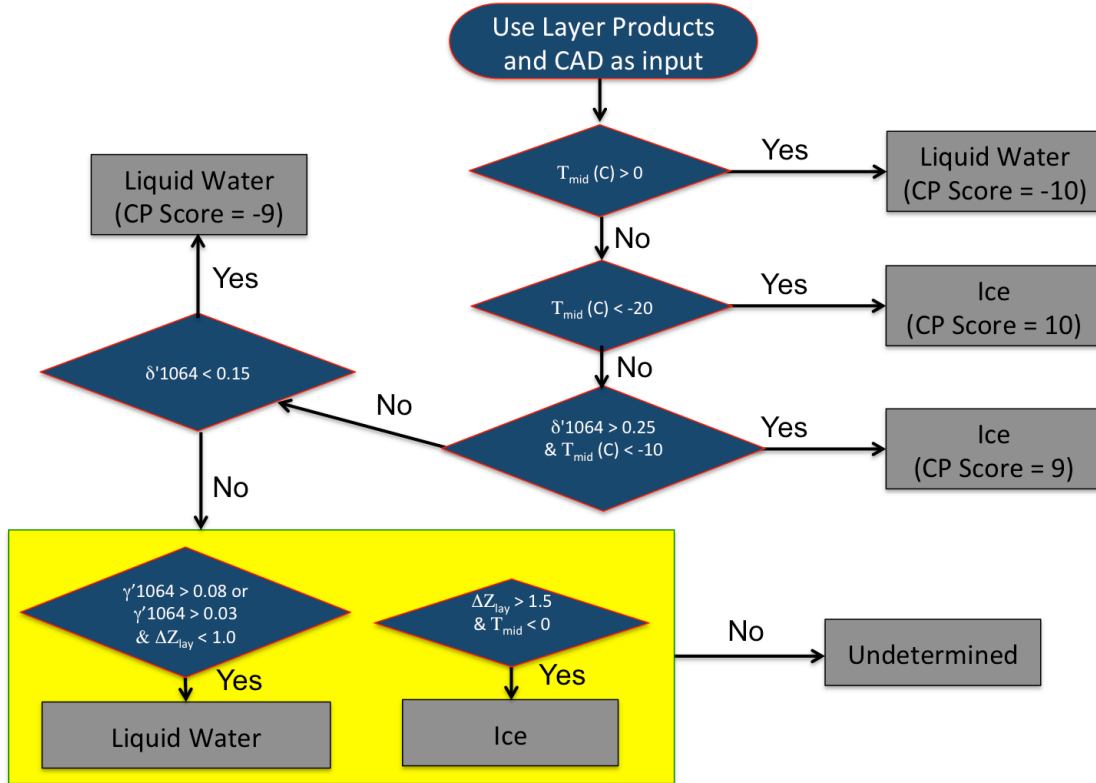

Figure 4.5. A flowchart of the CATS version 1 operational Cloud Phase (CP) algorithm.
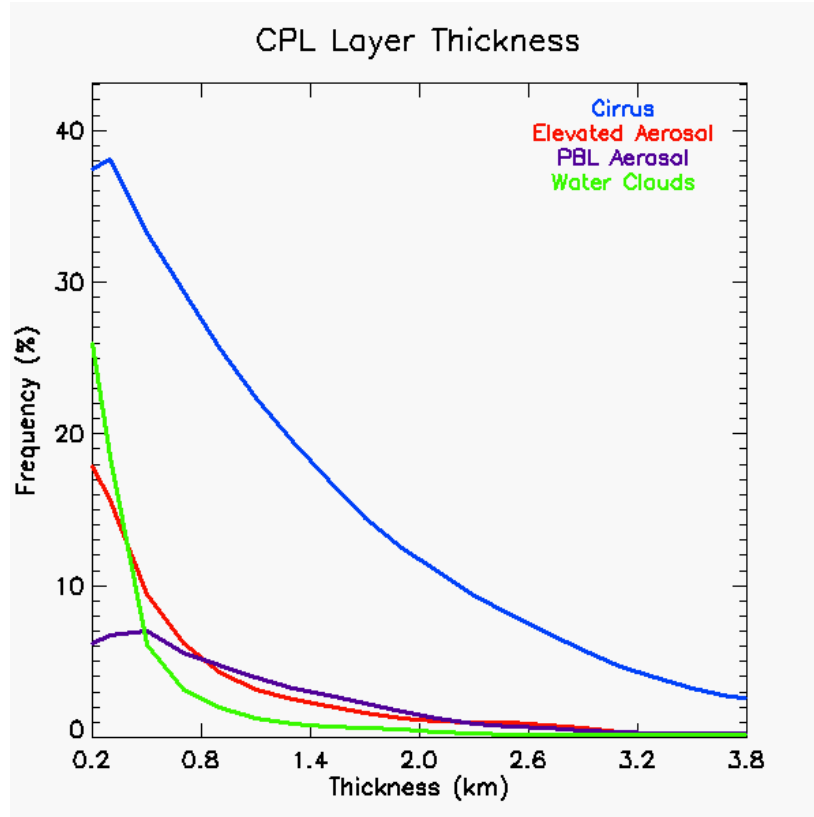
Figure 4.6. A frequency plots from CPL data between the years of 2003 and 2013 for cirrus (blue), water clouds (green), PBL aerosols (purple), and elevated aerosols (red). The frequency is normalized to the total number of observations in each category, and plotted for layer physical thickness.

4.4 Aerosol Typing

The CATS aerosol typing algorithms in both modes have heritage from the CALIOP aerosol typing algorithm (see Figure 2 from Omar et al. 2009) but with several differences:

1. CATS observed inputs from each field of view:
   - Feature Integrated Total Attenuated Backscatter at 1064 nm (γ'_{1064}) (sr^{-1}).
   - Feature Integrated Volume Depolarization Ratio at 1064 nm (δ'_{1064}).
   - Feature Integrated Color Ratio ($\gamma'_{1064}/\gamma'_{532}$) - Mode 7.1 only
   - L1B Spectral Depolarization Ratio Fraction in Feature ($\delta_{1064}/\delta_{532}$) - Mode 7.1 only.
2. Feature integrated total attenuated backscatter and depolarization ratio thresholds have been adjusted for wavelength (1064 nm versus 532 nm) and instrument minimal detectable backscatter.
3. The criteria for an elevated aerosol layer has been modified, yet remain similar to CALIOP. CATS criteria for classifying an aerosol layer as elevated requires the layer base altitude to be above the surface at least 1.0 km and at least 2 km in thickness.
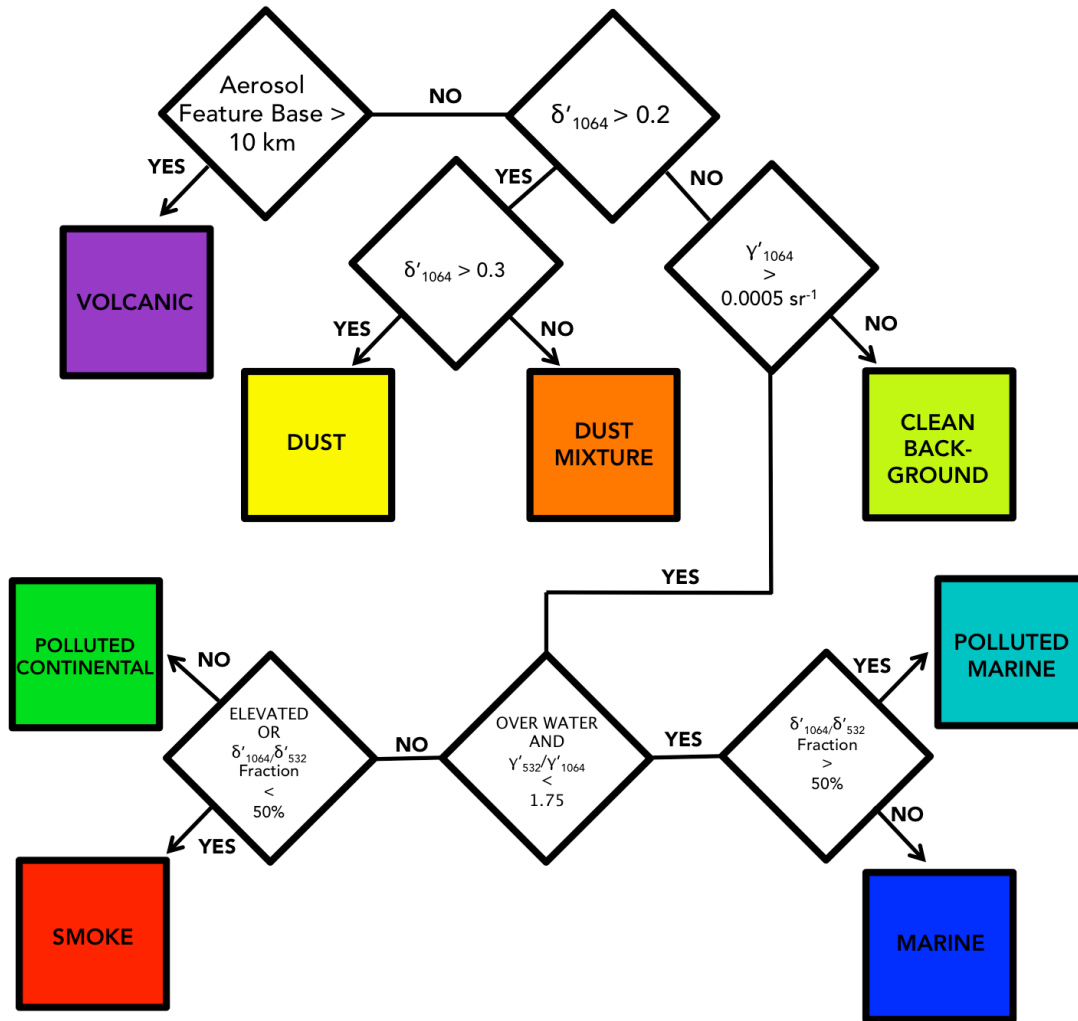
4.4.1 Mode 7.1

Figure 4.7. CATS Mode 7.1 aerosol typing flow chart.

In CATS mode 7.1, eight aerosol types are identified: dust, dust mixture, smoke, marine, marine mixture, polluted continental, clean/background, and volcanic. Mode 7.1 incorporates color ratio and spectral depolarization ratio information that is characteristic of aerosol regimes from Burton et al. (2012). The aerosol typing algorithm in mode 7.1, shown in Figure 4.7, first designates aerosol layers with a base above 10 km as volcanic. The algorithm then identifies strongly depolarizing aerosol layers ($\delta'_{1064} > 0.2$) characteristic of dust at 1064 nm (Burton et al., 2015) for each field of view. Aerosol features with depolarization ratios greater than 0.3 are characterized as dust and those between 0.2 and 0.3 are identified as a dust mixture. For less depolarizing aerosol layers, all features with γ'_{1064} less than 0.0005 sr^{-1} are classified as clean/background. Utilizing the IGBP land surface type database (Loveland et al., 2000), marine and polluted marine aerosol features are defined over water with an inverse color ratio ($\gamma'_{532}/\gamma'_{1064}$) less than 1.75 (Burton et al., 2012). Marine is then distinguished from polluted marine aerosol

features when the feature fraction of L1B spectral depolarization ratios above 1.75 is less than 50% (Burton et al., 2012). Smoke is identified in mode 7.1 when the aerosol layer is elevated, and the feature fraction of L1B spectral depolarization ratios above 1.75 is less than 50% (Burton et al., 2012), the feature is classified as smoke in mode 7.1. Otherwise, the feature is identified as polluted continental.
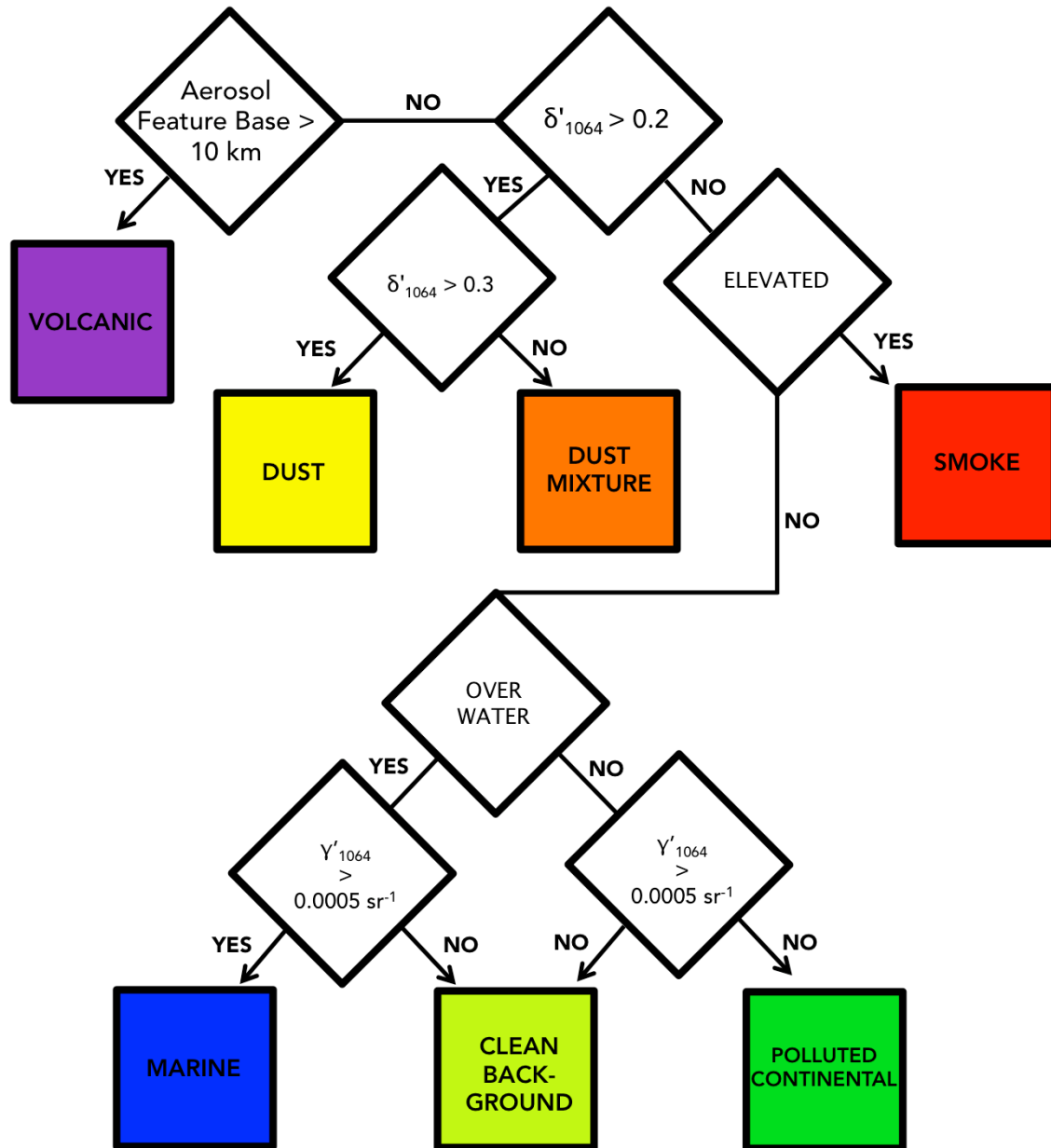
4.4.2 Mode 7.2

Figure 4.8. CATS Mode 7.2 aerosol typing flow chart.

In CATS mode 7.2, seven aerosol types are identified: dust, dust mixture, smoke, marine, polluted continental, clean/background, and volcanic. The CATS mode 7.2

aerosol typing algorithm, shown in Figure 4.8, is more similar to the CALIOP aerosol typing algorithm (Omar et al., 2009). As in mode 7.1, mode 7.2 aerosol typing algorithm first designates aerosol layers with a base above 10 km as volcanic. Then, strongly depolarizing aerosol layers are identified and are classified as dust when the 1064 nm depolarization ratio (δ'_{1064}) is greater than 0.3 and a dust mixture when the ratio is between 0.2 and 0.3 (Burton et al., 2015). For less depolarizing aerosol layers, smoke is identified when the aerosol layer is elevated as defined by the modified CALIOP elevated layer criteria outlined in #3. Again, using the IGBP land surface type database (Loveland et al., 2000), non-elevated aerosol features over water with γ'_{1064} greater than 0.0005 sr^{-1} are classified as marine and those less than 0.0005 sr^{-1} are classified as clean/background. Over land, polluted continental aerosol layers are identified when γ'_{1064} is greater than 0.0005 sr^{-1}, while those that do not meet this threshold are classified as clean/background. It is anticipated that future versions of the CATS mode 7.2 aerosol typing will incorporate simulated aerosol distributions from the NASA GEOS-5 model to help guide aerosol typing in instances where the observed quantities are characteristic of multiple aerosol types. Additionally, utilizing the GEOS-5 aerosol types in this way permits the classification of marine aerosol layers over land in coastal regions which would otherwise create boundaries in aerosol types where land meets ocean using land surface type alone (Nowottnick et al., 2015).

4.5 Initial Performance Assessment

The vertical feature mask algorithms result in numerous parameters that are reported in the L2 profile and layer files. These parameters include physical attributes of the layer, such as the layer top and base altitudes, pressures and temperatures, as well as mid-layer temperature. They also include an attempt to classify the layer using the aforementioned algorithms and create variables such as feature type, feature type score (CAD score), cloud phase, cloud phase score, and aerosol type. The integer values of these parameters are defined in Table 4.5.

The CATS algorithm team has performed rigorous testing of these algorithms using numerous case studies and statistical analysis. Here we demonstrate the performance of the algorithms with just a few examples of the performance assessment. More results will be illustrated in future publications. The first assessment presented here is a case study as the ISS passed along the western coast of Africa on 24 Aug. 2015. The MODIS Terra Corrected Reflectance (True Color) from this day shows Saharan Dust in northern Africa (orange/light brown colors) and extensive cloud cover in the central part of the continental (white) near the equator (Figure 4.9). Fires from biomass burning in southern Africa are shown in red, and smoke from these fires is transported off the coast over the ocean.

Figure 4.10 shows the CATS 1064 nm attenuated total backscatter for the track shown in blue (20:31 to 20:42 UTC) in Figure 4.8. Figure 4.11 shows the CATS vertical feature mask, including cloud phases and aerosol types present, for the same scene. Visible in the backscatter image are layers with high backscatter between 10-17 km that match the cloud cover in the MODIS image. The CATS VFM image (Figure 4.11) shows that the

top and base altitudes of these layers are accurately determined and that these layers are properly classified as ice clouds (light blue). Also visible in the lower troposphere (0-5 km) are layers with weaker backscatter, likely aerosols, and very thin layers with backscatter greater than 0.01 sr^{-1} km^{-1} in the southern latitudes, signatures of liquid water clouds. Figure 4.11 demonstrates very accurate detection of layer boundaries and classification of these cloud and aerosol layers. Dust layers (orange) are properly identified between 0 and 5 km altitude in the more northern latitudes (25.0 to 12.0 degrees), with the exception of a few profiles with weaker depolarization ratios that are classified as smoke (brown). In the southern latitude (0 to 10 degrees), both an elevated smoke plume and liquid water clouds (dark blue) are correctly classified. Just north of the equator (0 to 5 degrees), the CEAL routine appears to have properly identified a marine stratus cloud deck, which consists of liquid water, embedded in a smoke plume (Figure 4.11).

Table 4.5: Definitions of CATS Vertical Feature Mask Parameters

Parameter	Interpretation
Feature_Type	0 = Invalid 1 = Cloud 2 = Undetermined 3 = Aerosol
Feature_Type_Score	| 10 | = high confidence | 1 | = low confidence 0 = zero confidence
Cloud_Phase	0 = invalid 1 = water cloud 2 = unknown cloud phase 3 = ice cloud
Cloud_Phase_Score	| 10 | = high confidence | 1 | = low confidence 0 = zero confidence
Aerosol_Type	0 = Invalid 1 = Marine 2 = Polluted Marine 3 = Dust 4 = Dust mixture 5 = Clean/Background 6 = Polluted Continental 7 = Smoke 8 = Volcanic
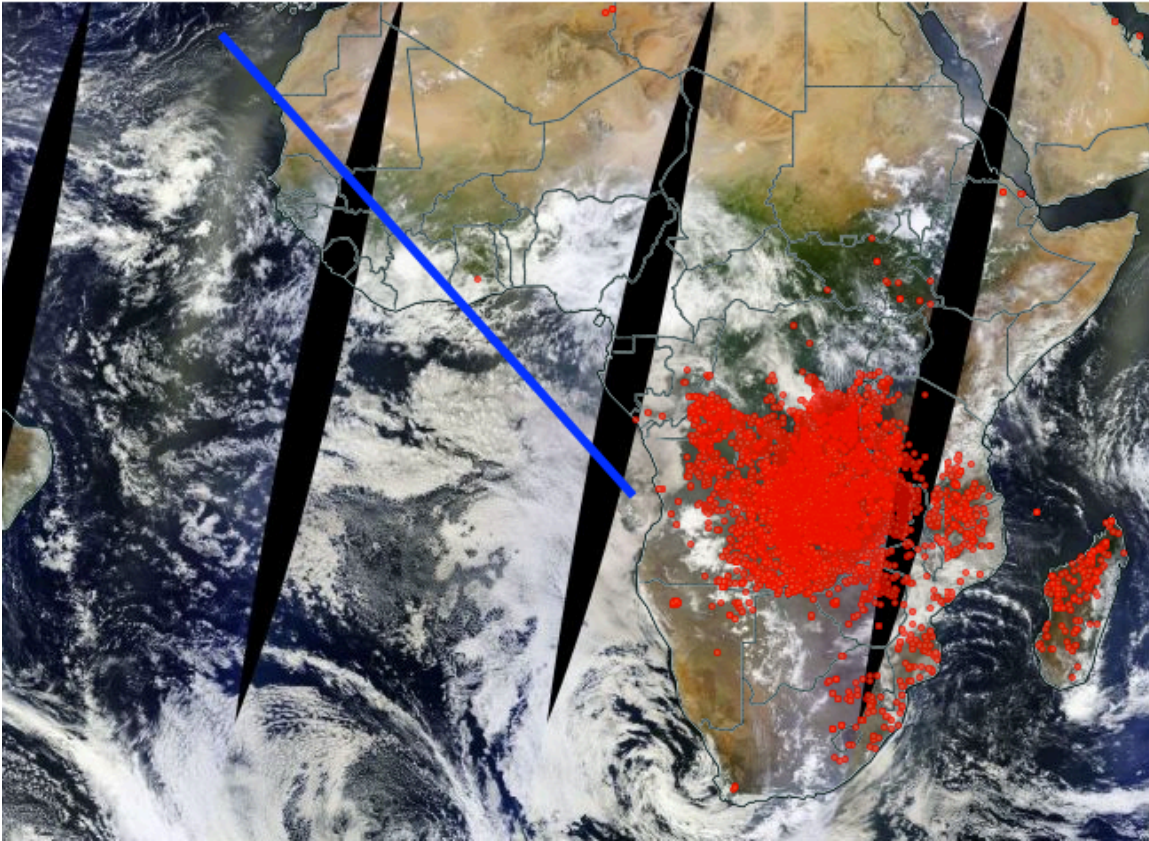
Figure 4.9. The MODIS Terra Corrected Reflectance (True Color) for 24 August 2015 shows Saharan Dust in northern Africa, extensive cloud cover in the central part of the continental, and smoke from biomass burning just off the coast in southern Africa. Biomass burning fires are shown in red, and the CATS track for this day (20:31 to 20:42 UTC) is shown in blue.

Figure 4.10 The 1064 nm attenuated total backscatter data for CATS on 24 August 2015 as the ISS passed over Africa (track shown in blue in Figure 4.8).

43

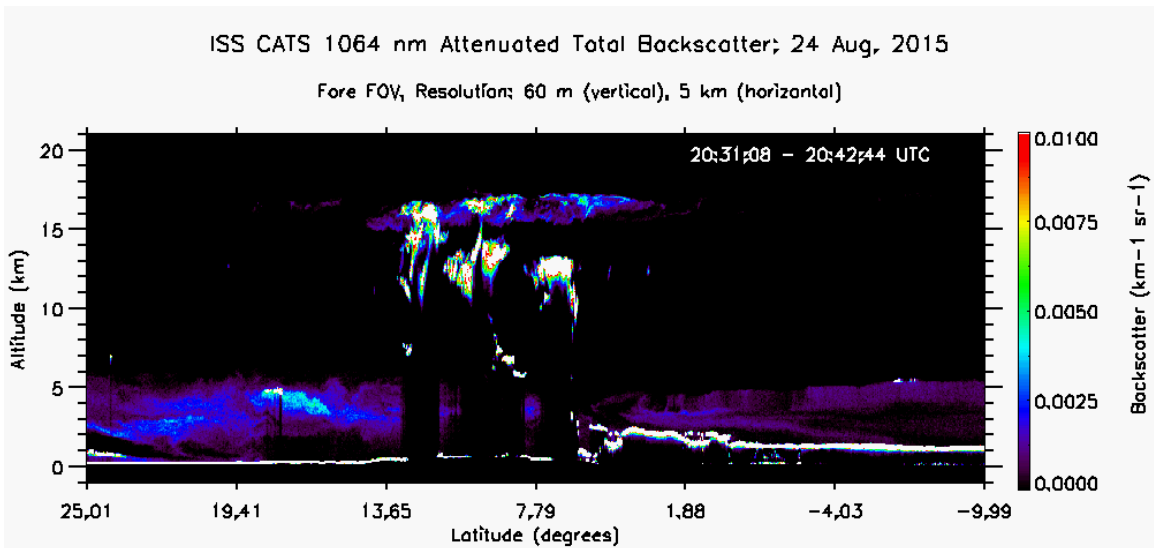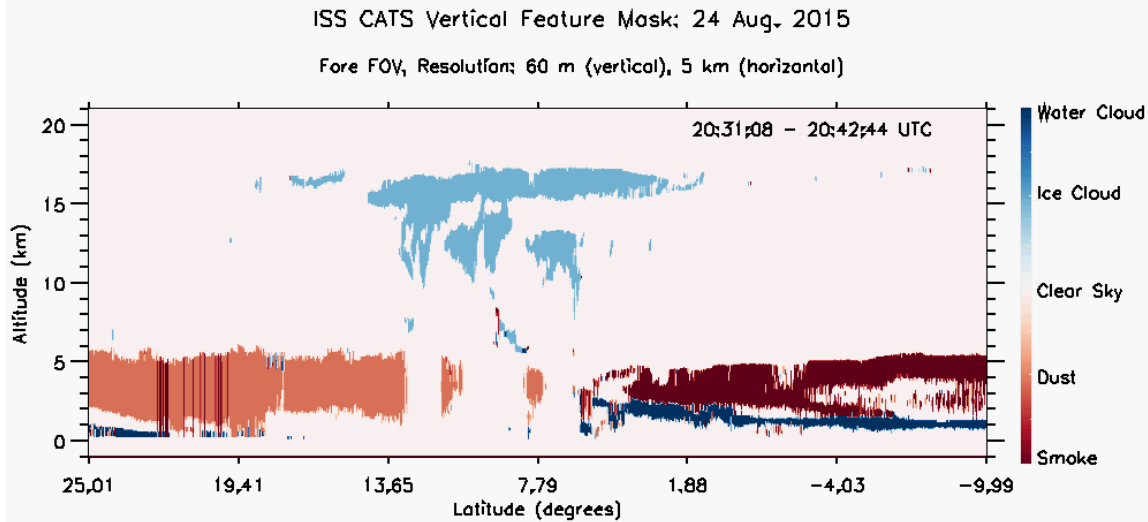**Figure 4.11.** The CATS vertical feature mask for the blue track shown in Figure 4.8 on 24 August 2015 that corresponds to the backscatter in Figure 4.9.
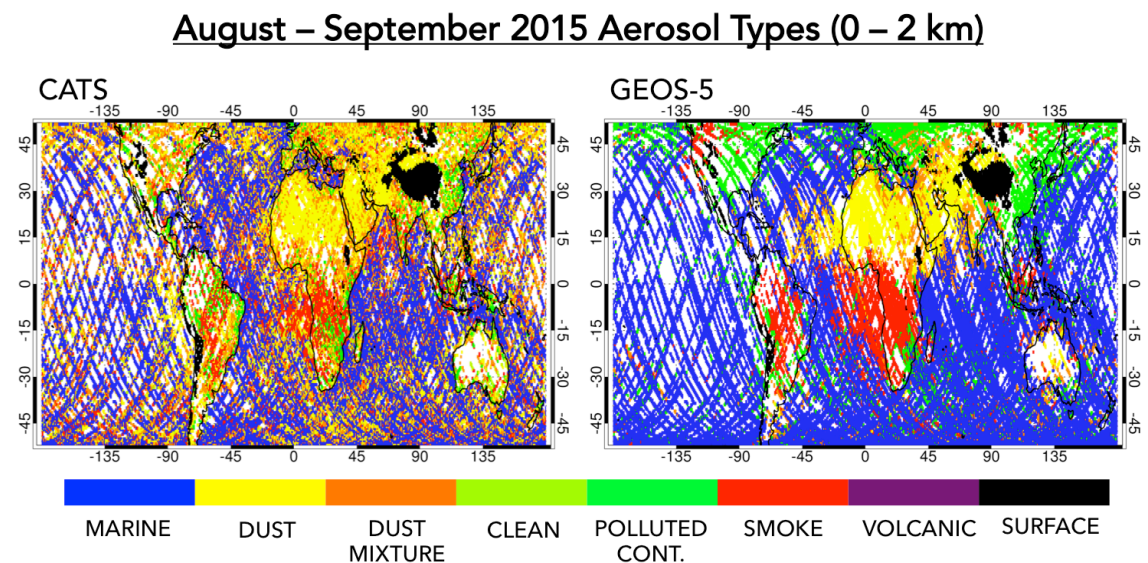
Figure 4.12. The CATS aerosol types for the most commonly observed aerosol layer for the months of August and September 2015. The GEOS-5 aerosol speciation along the CATS track for August and September 2015 is adjusted to fit the CATS aerosol types.

Initial comparisons between CATS aerosol typing and GEOS-5 aerosol speciation demonstrate good agreement. Figure 4.12 shows the CATS aerosol types and GEOS-5 aerosol speciation mapped to these types for the months of August and September 2015. The CATS aerosol typing algorithm accurately identifies dust (yellow and orange) in the Saharan desert and Middle East regions, similar to GEOS-5. The CATS smoke layers (red) due to wildfires in the northern latitudes and biomass burning over southern Africa also agree with the model forecasts. Although the model and CATS aerosol typing agree for dust, smoke, and marine (blue), a low bias in detection of polluted continental

44

aerosols (green) compared to the GEOS-5 forecasts is observed. Further assessment of the CATS aerosol typing is planned, including comparisons with CALIPSO and GEOS-5 using several months of data.

5.0 Overview of Geophysical Parameter Algorithms

When attempting to obtain cloud optical depth from a spacecraft or aircraft elastic backscatter lidar, two assumptions are required regarding the scattering characteristics of the cloud. One of these assumptions is that multiple scattering effects can be reliably quantified. Multiple scattering effects are the modification from the true optical depth caused by the increase in detected signal strength due to the portion of the detected signal, which has experienced more than one scattering interaction. It is primarily the result of photons that are deflected only slightly during the scattering process and stay in the field of view. This is referred to as forward scattering and it serves to decrease the perceived optical depth. Ice particles typically have a very pronounced forward scattering component, which will cause multiple scattering effects to be quite significant, especially from a spacecraft lidar. The other assumption is that the value of the particulate lidar ratio (S_p) is known. For a given scattering layer for CATS processing, S_p is assumed to be constant. In cases of transparent cloud layers and no or very weak aerosol loading this ratio can be estimated from the lidar data itself, but often the process will require externally computed values in order to solve the lidar equation. The values of both of these assumed parameters are determined by the details of the volumetric scattering phase function that quantifies light scattering as a function of scattering angle. The validity of these strongly relies upon former experience with cirrus lidar observations [*Spinhirne et al.*, 1990, 1996].

5.1 Transmittance Solution to the Lidar Equation

CATS data products report particulate layer lidar ratios, particulate backscatter coefficient profiles, particulate extinction coefficient profiles ($\sigma_p(z)$) and particulate layer optical depths ($\tau_p(z)$) for cloud and aerosol layers. These clouds and aerosol optical properties are derived as outlined below and demonstrated in *Spinhirne et al.* [1980, 1996], *Elouragini* [1995] and *Marenco* [1997]. Please note that transmittance, extinction, and optical depths obtained directly from the solution of the lidar equation are actually the apparent or effective values [*Platt*, 1979], which include multiple scattered photons and are denoted with the <u>superscript prime</u>.

The working lidar equation for a high altitude nadir pointing lidar with photon counting detectors has been stated previously in the literature and can be rewritten in the following form:

$$\gamma_n(r) = \beta(r)T'^2(r) = \frac{[n(r)D(n) - n_b]r^2}{O(r)ECT_o^2} . \qquad \text{Eq. 5.1}$$

The raw lidar signal is represented by *n(r)* as a function of range (r) from the lidar. *D* represents the dead time correction as a function of signal strength needed when using photon counting detectors. The solar background signal is n_b. Other instrument corrections are the near field overlap correction (*O*) and outgoing energy normalization

45

factor (E). For CATS, the near field overlap correction factor goes to unity well before the 28 km height where processing begins. The system calibration is C and the ozone transmission factor is T_o^2. The left side of the equation is the calibrated normalized attenuated backscatter coefficient corrected for ozone attenuation. The total (particulate and molecular) volumetric backscatter coefficient at range r from height z to the instrument altitude is denoted by $\beta(r)$ and the two-way particulate and molecular effective transmission factor is expressed as $T'^2(r)$, which is equivalent to $\exp[-2(\tau_m(z) + \tau'_p(z))]$. Optical depth is represented by the symbol τ, while the subscripts m and p designate molecular and particulate contributions, respectively. Furthermore, the influence of the multiple scattering factor (η) on the particulate optical depth is described by:

$$\tau'_p(z) = \int \sigma'_p(z)dz = \int \eta(z)\sigma_p(z)dz \cong \bar{\eta}\int \sigma_p(z)dz ,\qquad \text{Eq 5.2}$$

where σ_p is particulate extinction. CALIPSO corrects retrievals of optical properties using multiple scattering factors (η) of 0.60 (cirrus), 0.40 (liquid water clouds), and 1.00 (aerosols) that does not vary with optical depth or range [*Young and Vaughan*, 2009]. Initial analysis and comparisons with CPL suggest the η for CATS cloud and aerosol layers at 1064 nm is 0.77 for cirrus clouds and 0.60 for liquid water clouds. At 532 nm, initial comparisons with CALIPSO cirrus lidar ratios indicate a η of 0.XX for cirrus clouds. A η of 1.00 is used for all aerosol layers in version 1.02 of the CATS L2 data products, similar to CALIPSO, until more analysis is performed.

Since the molecular contribution to the total backscatter and transmission can be computed from theory, it is advantageous to separate the scattering terms into components, which represent the molecular and particulate contributions independently. Assuming $\beta = \beta_p + \beta_m$ and $T'^2 = T_p'^2 T_m^2$ the lidar equation becomes:

$$\gamma_n = \beta_p T_p'^2 T_m^2 + \beta_m T_p'^2 T_m^2 .\qquad \text{Eq. 5.3}$$

The following relationships must now be defined:

$$T_p'^2 = e^{-2\int \sigma'_p dz} \quad \text{and} \quad S'_p = \frac{\sigma'_p}{\beta_p}\qquad \text{Eq. 5.4}$$

($S'p$ assumed to be a constant for each layer), and

$$T_m^2 = e^{-2\int \sigma_m dz} \quad \text{and} \quad S_m = \frac{\sigma_m}{\beta_m},\qquad \text{Eq. 5.5}$$

where S'_p and S_m are the effective particulate and molecular lidar ratios, respectively. $T_m^2(z)$ and $\beta_m(z)$ can be calculated accurately given the vertical temperature and pressure structure of the atmosphere using the GMAO MET data base by selecting atmospheric profiles closest in space and in time to the ISS satellite track and the fact that S_m is known to be $8\pi/3$ throughout the vertical profile. Once the molecular backscatter coefficient and two-way molecular transmission are computed, the equation can be solved for the vertical profile of β_p. The true particulate optical depth and extinction profiles can then be computed from the values of S_p, β_p, and η. From the above relationships, we see that:

$$\frac{d(T'^2_p)}{dz} = T'^2_p\left(-2S'_p\right)\beta_p.$$
<div align="right">Eq. 5.6</div>

We can use this relationship to substitute for β_p in Equation 5.3 to arrive at:

$$\gamma_n = -\left(\frac{1}{2S'_p}\right)T^2_m\frac{d\left(T'^2_p\right)}{dz} + \beta_m T^2_m T'^2_p \quad \text{or} \quad \frac{d\left(T'^2_p\right)}{dz} - 2S'_p\beta_m T'^2_p = -\frac{2S'_p\gamma_n}{T^2_m} \quad \text{Eq. 5.7}$$

By specifying z as the independent variable and T'^2_p as the dependent variable, this is a first order linear ordinary differential equation; it is a special form of the Bernoulli equation. The solution can be found by using the common integrating factor method where the integrating factor is $F = e^{-2X\int\sigma_m dz}$, and $X \equiv \dfrac{S'_p}{S_m}$. The general solution is:

$$T^{2X}_m T'^2_p = -2S'_p\int T^{2(X-1)}_m\gamma_n dz + K \quad ,$$
<div align="right">Eq. 5.8</div>

where the integrand is defined only where particulates are present and K is a constant of integration.

For convenience, we redefine the coordinate z as the <u>vertical</u> distance from the lidar instrument, increasing downward. If we visualize the situation where the lidar pulse encounters layers of particulates after traveling through the molecular atmosphere below the satellite and allowing for the effect of the lidar pointing off-nadir at a zenith angle of θ, we can define the boundary condition at the top of any particulate layer, $I_B(z_t)$, as:

$$I_B(z_t) = T'^{2\sec\theta}_p(z_t)T^{2X\sec\theta}_m(z_t),$$
<div align="right">Eq. 5.9</div>

where z_t is the vertical distance to the top of the layer. If the layer is the first layer encountered, the $T'^{2\sec\theta}_p(z_t)$ term can be estimated as 1.00. The calculation of $I_B(z_t)$ for subsequent layers in the atmospheric profile is based on the assumption that $T'^{2\sec\theta}_p(z_t)$ will be the same as the $T'^{2\sec\theta}_p$ value at the bottom of the layer above.

So in general, the two-way effective particulate transmission within the particulate layer, whether cloud or aerosol, given a lidar zenith angle of θ is

$$T'^{2\sec\theta}_p\left(z\right) = \frac{I_B(z_t) - 2S'_p\sec\theta\int\limits_{z_t}^{z} T^{2(X-1)\sec\theta}_m\gamma_n(z)dz'}{T^{2X\sec\theta}_m\left(z\right)}.$$
<div align="right">Eq. 5.10</div>

This forward inversion processing continues throughout the series of particulate layers until $T'^{2\sec\theta}_p\left(z\right) < T^2_L$ or the signal from the earth's surface is detected. T^2_L is a limit defined through error consideration because the lidar equation becomes unstable as the transmission approaches 0.0. T^2_L generally is set between 0.003 and 0.004.

5.2 Determination of Lidar Ratio

An important ingredient of the transmission solution of Equation 5.10 is the effective lidar ratio (S'_p). There are four basic categories of lidar ratio input sources: constrained, unconstrained default, modified default, and opaque. A discussion of each follows.

5.2.1 Constrained Lidar Ratio

When the particulate layer being analyzed is determined to meet the "appropriate criteria" for underlying signal loss analysis, an algorithm to calculate an estimate of S'_p is called. If S'_p is found to be within tolerances, currently between 8-100 sr, it will be used in Equation 5.10. Lidar ratios calculated in this manner are categorized as "constrained". "Appropriate criteria" would be a layer that is transparent (with either a lower layer or the earth's surface sensed) and resides just above pristine air (no aerosols) so that it is possible to determine signal loss through the layer. The clear air zone is restricted to a minimum thickness of 0.616 km and maximum of 3 km thickness for CATS processing. Ice clouds above 7 km altitude are the most likely candidates. Under these conditions, an estimate of $T_p'^{2\sec\theta}(z_b)$, where z_b is the vertical distance to the bottom of the layer, can be found using the following equation:

$$T_p'^{2\sec\theta}(z_b) = \frac{\int_{z_b}^{z_c} \gamma_n(z)dz}{T_m^{2\sec\theta}(z_b)\int_{z_b}^{z_c} \beta_m(z)T_m^{2\sec\theta}(z)dz} . \qquad \text{Eq. 5.11}$$

The designation z_c is the distance to the end of the clear air analysis zone. This method is called the signal loss technique. It compares the integrated attenuated total backscatter lidar signal in the presumed clear atmosphere directly below the layer to the estimated integrated molecular signal directly below the layer that incorporates the accumulated molecular transmission loss starting from the instrument height if no cloud or aerosol layer were present. By referring to Equation 5.9, the boundary condition at the bottom of the layer can be defined as $I_B(z_b) = T_p'^{2\sec\theta}(z_b)T_m^{2X\sec\theta}(z_b)$. S'_p can then be calculated through an iterative solution from the following equation:

$$S'_p = \frac{I_B(z_t) - I_B(z_b)}{2\sec\theta \int_{z_t}^{z_b} \gamma_n(z)T_m^{2(X-1)\sec\theta}(z)dz} . \qquad \text{Eq. 5.12}$$

The iterative process is started with an initial guess of S'_p as it relates to the X parameter, with the next iteration using the calculated value until the solution converges to a set tolerance (0.08 sr). A version of this routine initially worked well during automated micro-pulse lidar (MPL) processing of aerosols using the calibrated signal to resolve the layer optical depth similar to the loss of signal in a cloud [*Spinhirne et al.*, 1999]. This routine is also valid for PSCs and enhanced upper tropospheric aerosol layers. This algorithm has been used for over ten years of CPL optical products and the constrained lidar ratios derived from this procedure agree favorably with previous studies of lidar ratio [*Yorks et al.*, 2011a].

5.2.2 Default Lidar Ratio

For atmospheric layers where S'_p cannot be calculated as described above from the transmission loss measurement region due to high aerosol loading or low SNR or proximity to the earth's surface, a value will be assigned for each layer based on the layer type as determined in Section 4.0. CATS ice cloud lidar ratios are assigned based on relationships with layer integrated depolarization ratio geographic location, as explained in Table 5.1, as retrieved from CPL and CALIPSO data [*Yorks et al.* 2011; 2014]. Values of 18.0 sr and 22.0 sr are used for liquid water clouds and clouds of unknown phase, respectively [*Yorks et al.* 2011]. For aerosols, the S'_p is assigned from a look-up table based on the aerosol type and the values are very similar to those used in CALIPSO, CPL and GLAS algorithms, as shown in Table 5.2 [*Palm et al.* 2002; *Omar et al.* 2009]. The default lidar ratios from these relationships or tables are the true S_p. The effective ratio (S'_p) used in equation (10) is determined as:

$$S'_p = \eta \, S_p, \hspace{3cm} \text{Eq. 5.13}$$

where $\eta = 1$ for CPL processing, but not for CATS processing. If this lidar ratio remains unmodified through the solution process, we describe this category as "unconstrained default" or just default.

Table 5.1: CATS Ice Cloud Default Lidar Ratios

Location	Land Type	Sp
Mid-Latitude	Land	24.0
Mid-Latitude	Ocean	30.0
Tropics	Land	28.0
Tropics	Ocean	32.0

Table 5.2: CATS Aerosol Default Lidar Ratios

Aerosol Type	532 nm	1064 nm
Marine	25.0	45.0
Marine Mixture	45.0	40.0
Dust	45.0	55.0
Dust Mixture	35.0	45.0
Clean/Background	55.0	35.0
Polluted Continental	65.0	35.0
Smoke	70.0	40.0
Volcanic	45.0	35.0

5.2.3 Modified Default Lidar Ratio

During CATS processing using Equation 5.10 with the default lidar ratio, if the $T'^{2\sec\theta}_p(z)$ term goes below pre-defined set limits before reaching the bottom of the cloud layer, an

iterative process is invoked where the effective lidar ratio is reduced by 0.5 sr and the layer is reprocessed. This iterative reprocessing continues until the analysis can reach the bottom of the layer or the number of iterations reaches 30. if the $T_p'^{2\sec\theta}(z)$ term goes above pre-defined set limits before reaching the bottom of the cloud layer, an iterative process is invoked where the effective lidar ratio is increased by 0.5 sr and the layer is reprocessed. This iterative reprocessing continues until the analysis can reach the bottom of the layer or the number of iterations reaches 30. We describe these two conditions as the "modified default" category.

5.2.4 Opaque Lidar Ratio

In CATS data processing, opaque layers are treated as special cases of the "constrained" lidar ratio algorithm. For the case of an opaque layer (defined in practical terms as a layer that has no layer sensed below it and no ground signal), the particulate transmission squared term at the "lidar" cloud bottom, where the lidar signal is extinguished, is set to an assumed value consistent with the lidar equation transmission results just before the point of opaqueness (where the lidar equation becomes unstable). Currently, this value of $T_p'^{2\sec\theta}(z_b)$ is set to 0.004. The constrained lidar ratio is then calculated using the $T_p'^{2\sec\theta}(z_b)$ value of this special case for input into Equation 5.12. If the calculated lidar ratio value falls within thresholds, it is used; otherwise a default value is sent to start the full optical processing of the layer. The "opaque" lidar ratio and the resultant transmission and extinction profiles only apply to the segment of the actual cloud above the height where the lidar signal is extinguished.

5.3 Estimates of Geophysical Parameters

In order to obtain the relative density for aerosol and cloud scattering, it is useful to solve for the actual particulate backscatter coefficient with attenuation removed (β_p). An equation for the backscatter coefficient profile can be obtained by using the results from equation (10) as input to equation (3) and rearranging:

$$\beta_p(z) = \frac{\gamma_n(z)}{T_m^{2\sec\theta}(z)T_p'^{2\sec\theta}(z)} - \beta_m(z).$$

Eq. 5.14

Once the particulate effective transmission and backscatter profiles for each layer have been calculated, it is a straightforward procedure to determine the extinction coefficient profiles. Extinction coefficient for particulates (σ_p) is defined as the total scattered energy at range z. The final lidar ratio (S_p') used in Equation 5.10, regardless of the source, now must be corrected for multiple scattering by converting to S_p using Equation 5.13. The extinction profile through the layer is then expressed as a simple product of S_p and $\beta_p(z)$:

$$\sigma_p(z) = S_p\beta_p(z) .$$

Eq. 5.15

Note that multiple scattering has already been accounted for in the calculation of S_p.

The solution to the lidar equation to obtain particulate <u>effective</u> optical depth (τ'_p) at any range z from a nadir pointing high altitude lidar is related to the particulate effective transmission calculated in Equation 5.10:

$$\tau'_p(z) = -\left(\frac{1}{2}\right)\ln\left[T'^2_p(z)\right] \quad \text{or} \quad T'^2_p(z) = e^{-2\tau'_p(z)}. \qquad \text{Eq. 5.16}$$

The specific method CATS uses to calculate the particulate <u>layer</u> optical depth stems from the same transmission solution to the lidar equation, but uses the relationship of the extinction coefficient profile in the layer to optical depth. The final optical depth products from these calculations will be the optical depth (τ_l) for each of the particulate layers meeting the analysis criteria:

$$\tau_l = \int_{z_t}^{z_b} \sigma_p(z)dz \qquad , \qquad \text{Eq. 5.17}$$

where z_b and z_t are the bottom and top locations of the particulate layer, respectively and multiple scattering has already been factored out. The vertical coordinate limits on the integration in Equation 5.10 will be determined by the cloud and aerosol boundary algorithms described in Section 4.1. In practice, the integration will be carried out starting at the top of the first particulate layer. Although the whole molecular transmission vertical profile starting at the instrument altitude and ending at the bottom of the lowest particulate layer sensed will have to be calculated, the particulate transmission vertical profile will be produced only inside cloud and aerosol layers. The boundary condition in Equation 5.9 at the top of any secondary layer will involve the particulate transmission squared at the bottom of the layer above and the molecular transmission squared at the top of the current secondary layer.

The attenuation of the pulse energy due to molecular scattering in the intervening clear air layers is small in the mid to high troposphere where the optically thin clouds are located. The magnitude of the molecular scattering is a significant fraction of lower altitude aerosol scattering since the gaseous atmosphere is relatively dense at the low altitudes of the boundary layer and the optical density of the aerosol particles are typically much lower than that found in cirrus clouds.

An estimate of ice water content is calculated based on its relationship with extinction and temperature [*Heymsfield et al.*, 2014].

$$IWC(z) = 0.303333\sigma(z)D_e \quad , \text{where} \quad D_e = ae^{BT(z)} \qquad \text{Eq. 5.18}$$

a and B change with temperature (T in degrees C) as follows:

for -56 < $T(z)$ < 0, a=308.4, B=0.0152

for -71 < $T(z)$< -56, a=9.1744e04, B=0.117

for -85 < $T(z)$< -71, a=83.3, B=0.0184.

IWC profiles are then developed inside each layer. Ice water path (IWP) for each layer is the integration of the IWC profile in the layer.

$$IWP = \int_{top}^{base} IWC(z)dz \qquad \text{Eq. 5.19}$$

Once the full vertical column of the current profile has been analyzed, the three column parameters are calculated: total column optical depth, total aerosol optical depth, and total cloud optical depth. Each layer has been identified as either aerosol or cloud, so the sums can be calculated accordingly. Layer optical depths are summed vertically in

the column until a known layer has an invalid OD result. Then the column totals are set to invalid (-9.9).   In those columns where a layer is identified as "opaque", the total optical depth values are set to "-1".

5.4 Initial Performance Assessment

The optical properties algorithms result in numerous parameters that are reported in the L2 profile and layer HDF files. These parameters include optical attributes of the layer, such as the layer optical depth, layer 2-way particulate transmission, as well as layer ice water path and layer integrated depolarization ratio. They also include an attempt to classify the layer optical process using the aforementioned algorithms and create variables such as extinction QC flag, lidar ratio selection method flag, and constrained lidar ratio flag. The integer values of these parameters are defined in Table 5.3. Optical attributes of the lidar profile include depolarization ratio, extinction, particulate backscatter, ice water content, and column optical depth.

Table 5.3: Definitions of CATS Optical Properties Flags

Parameter	Interpretation
Extinction QC Flag	-1=calculation not attempted 0 = non-opaque layer extinction analysis nominal 1 = layer hit earth's surface before layer base reached, adjusted base 2 = decrease lidar ratio, iteration process successful 3 = increase lidar ratio, iteration process successful 4 = max # of iterations reached, analysis stopped 5 = saturated before layer base, analysis stopped 6 = opaque, layer OD= -1, initial lidar ratio accepted 7 = opaque, layer OD= -1, iteration process successful 8 = layer OD out of bounds (invalid) OD= -999.99 9 = result invalid, final lidar ratio out of bounds
Lidar Ratio Selection Method	0 = generic default 1 = aerosol lookup table 2 = cloud lookup table 3 = 1064 lidar ratio used 532 OD (for ice clouds only) 4 = constrained result using clear zone below layer 5 = constrained result with opaque layer 6 = lower lidar ratio by max 15sr to reach layer base 7 = raise lidar ratio by max 15sr to reach layer base 8 = open slot (not used) 9 = missing
Constrained Lidar Ratio Flag	0 = value using nominal "constrained' procedure 1 = useful value using opaque "constrained' procedure 2 = constrained lidar ratio outside thresholds 3 = below layer clear zone too small 4 = clear zone signal error > threshold

	5 = Tp_sq < allowed min
	6 = Tp_sq at or below 0.0
	7 = 1064 lidar ratio using 532 OD (for ice clouds only)
	8 = Tp_sq at or below 0.0 in opaque cloud conditions
	9 = missing

The CATS algorithm team has performed rigorous testing of these algorithms using numerous case studies and statistical analysis. Here we demonstrate the performance of the optical properties algorithms with just a few examples of the performance assessment. More results will be illustrated in future publications. The assessment presented here is the same case study as was presented in Section 4.5 for the vertical feature finder when the ISS passed along the western coast of Africa on 24 Aug. 2015. The optical properties correspond to the data displayed in Figures 4.12 and 4.13.

Figure 5.1 shows the CATS 1064 nm particulate backscatter retrieval (with a logarithmic color scale) for the track shown in blue (20:31 to 20:42 UTC) in Figure 4.11. Figure 5.2 shows the CATS extinction retrieval, also with a logarithmic scale, for the same scene. Both the particulate backscatter and extinction values match what would be expected from the lidar signal strength found in Figure 4.12, where the lidar first scans an extended dust layer, then a high cirrus deck with scattered underlying cumulonimbus, followed by an elevated smoke layer over an extended thick stratus cloud deck. Figure 5.3 shows the column optical depth results from integration of the extinction profiles in Figure 5.2. Values shown are the total aerosol optical depth (red), total cloud optical depth (blue), and total column optical depth (green). An opaque layer in the column will cause the total optical depth value to be set to -1, indicating an unknown optical depth over 3.0. The scattered cumulonimbus under the cirrus and much of the stratus deck were analyzed to be opaque to the lidar.

Figure 5.4 demonstrates the results of using Equation 5.18 to calculate the ice water content from relationships with extinction and temperature. The color bar uses a logarithmic scale. The retrieval is only valid in ice clouds and matches well with expected values.
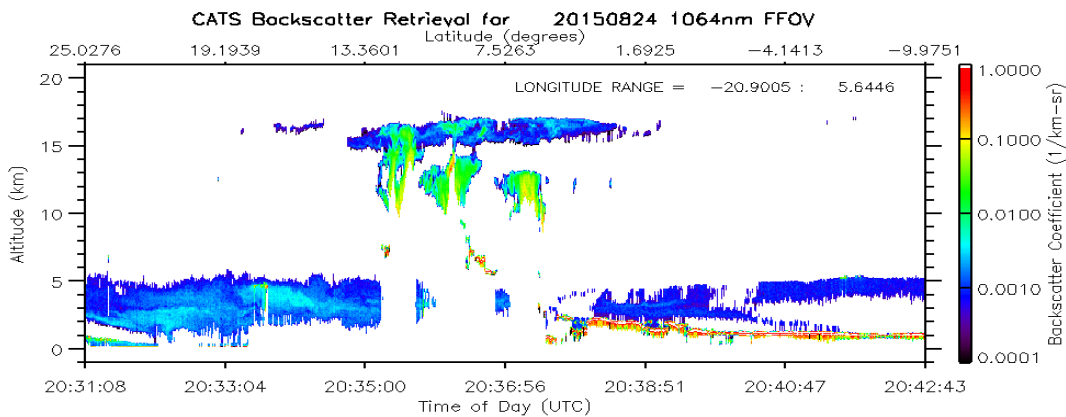

53

Figure 5.1. The 1064 nm particulate backscatter retrieval corrected for attenuation for CATS on 24 August 2015 based on the lidar signals in Figure 4.12 as the ISS passed over first dust, then cirrus, then smoke over stratus (see Fig. 4.13).
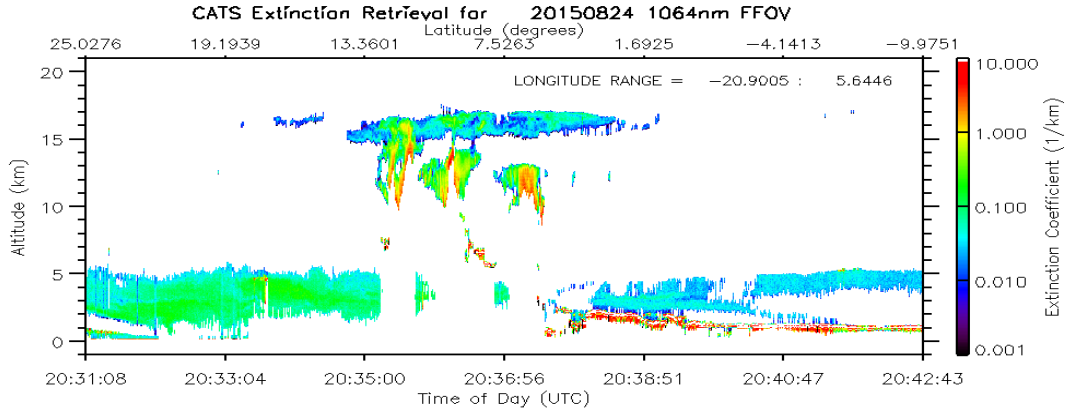

Figure 5.2. The corresponding 1064 nm extinction retrieval for CATS on 24 August 2015 as the ISS passed over first dust, then cirrus, then smoke over stratus during its western Africa pass.
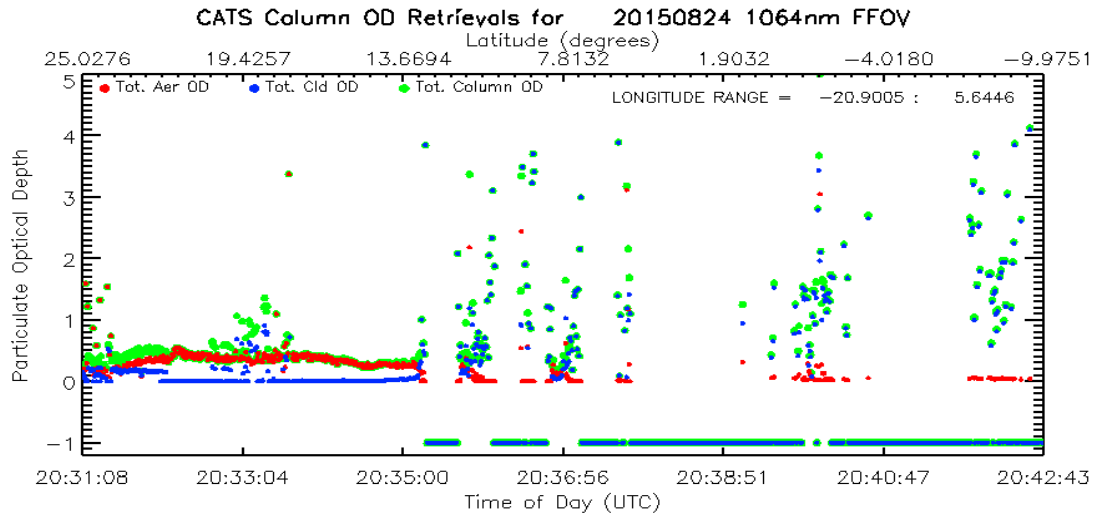

Figure 5.3. The CATS total aerosol optical depth (red), total cloud optical depth (blue) and total column optical depth (green) on 24 August 2015 that corresponds to the extinction in Figure 5.2. Any opaque layer will cause the total optical depths to default to -1, indicating an unknown optical depth over 3.0.
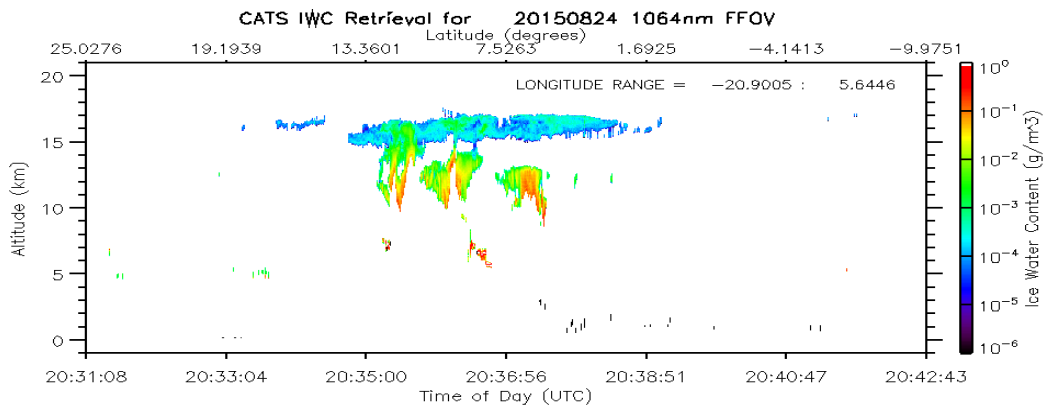

54

Figure 5.4. The 1064 nm ice water content retrieval for CATS on 24 August 2015 as the ISS passed over western Africa. The retrieval is only valid in ice clouds.

References

Behrendt, A.; T. Nakamura, 2002: Calculation of the calibration constant of polarization lidar and its dependency on atmospheric temperature. Optics Express,Vol. 10, No. 16, p. 805-817.

Bodhaine B. A., N. B. Wood, E. G. Dutton, and J. R. Slusser, 1999: On Rayleigh optical depth calculations. J. Atmos. Ocean Technol., 16, 1854-1861.

Bucholtz, A.,1995: Rayleigh-scattering calculations for the terrestrial atmosphere. Appl. Opt., 34, 2765-2773.

Burton, S. P., Ferrare, R. A., Hostetler, C. A., Hair, J. W., Rogers, R. R., Obland, M. D., Butler, C. F., Cook, A. L., Harper, D. B., and Froyd, K. D., 2012: Aerosol classification using airborne High Spectral Resolution Lidar measurements – methodology and examples, Atmos. Meas. Tech., 5, 73-98, doi:10.5194/amt-5-73-2012, 2012.

Burton, S. P., Hair, J. W., Kahnert, M., Ferrare, R. A., Hostetler, C. A., Cook, A. L., Harper, D. B., Berkoff, T. A., Seaman, S. T., Collins, J. E., Fenn, M. A., and Rogers, R. R., 2015: Observations of the spectral dependence of particle depolarization ratio of aerosols using NASA Langley airborne High Spectral Resolution Lidar, Atmos. Chem. Phys. Discuss., 15, 24751-24803, doi:10.5194/acpd-15-24751-2015.

Collis, R. T. H., and P. B. Russell, 1976: Lidar measurement of particles and gases by elastic backscattering and differential absorption", Laser Monitoring of the Atmosphere, E. D. Hinkley (editor), (Springer-Verlag, 1976), Chapter 4.

Del Guasta, M., 1998: Errors in the retrieval of thin-cloud optical parameters obtained with a two-boundary algorithm, Appl. Opt., 37, 5522–5540.

Elouragini, S. (1995), Useful algorithms to derive the optical properties of clouds from a backscatter lidar return, *J. Modern. Opt.*, **42**, 1439-1446.

Fernald, F.G., B. M. Herman and J. A. Reagan, 1972: "Determination of aerosol height distributions with lidar", *Journal of Applied Meteorology*, **11**, 482-489.

Hair, J., C. Hostetler, A. Cook, D. Harper, R. Ferrare, T. Mack, W. Welch, L. Izquierdo, and F. Hovis, 2008: Airborne High Spectral Resolution Lidar for profiling aerosol optical properties, *Appl. Opt.*, **47**, 6734–6752, doi:10.1364/AO.47.006734.

Heymsfield, A. J., et al. (2014), Relationships between ice water content and volume extinction coefficient from in situ observations for temperatures from 0^0 to -86^0C: Implications for spaceborne lidar retrievals , *J. Appl. Meteor. Climatol.*, **53**, 479-505, doi:10.1175/JAMC-D-13-087.1.

Hlavka, D. L., J. E. Yorks, S. A. Young, M. A. Vaughan, R. E. Kuehn, M. J. McGill, and S. D. Rodier, 2012: Airborne validation of cirrus cloud properties derived from CALIPSO lidar measurements: Optical properties. *J. Geophys. Res*., **117**, D09207, doi:10.1029/2011JD017053.

Hovis, F.E. , M. Rhoades, R. L. Burnham, J. D. Force, T. Schum, B. M. Gentry, H. Chen, S. X. Li, J. W. Hair, A. L. Cook, and C. A. Hostetler, 2004: Single-frequency lasers for remote sensing. Proc. SPIE 5332, 263–270.

Hu, Y., Z. Liu, D. Winker, M. Vaughan, V. Noel, L. Bissonnette, G. Roy, and M. McGill, 2006: A simple relation between lidar multiple scattering and depolarization for water clouds. Optics Letters, 31, 1809-1811.

Hu, Y., and coauthors, 2009: CALIPSO/CALIOP cloud phase discrimination algorithm, *J. Atmos. Oceanic Technol.*, *26*, 2293–2309. doi:10.1175/2009JTECHA1280.1.

Iqbal, M., 1983: An Introduction to Solar Radiation, Academic Press, New York, NY.

Klett, J. D., 1981: Stable analytical inversion solution for processing lidar returns. *Appl. Opt*, **20**, 211–220.

Klett, J. D., 1985: Lidar inversion with variable backscatter/extinction ratios. *Appl. Opt*, **24**, 1638–1643.

Liou, K. N., Y. Takano, and P. Yang *et al.*, 2000: Light scattering and radiative transfer in ice crystal clouds: applications to climate research. in Light Scattering by Nonspherical Particles, M. Mishchenko *et al.*, Eds. San Diego, CA: Academic, pp. 417–449.

Liu, Z., M. McGill, Y. Hu, C.A. Hostetler, M. Vaughan, and D. Winker, 2004: Validating lidar depolarization calibration using solar radiation scattered by ice clouds. Geoscience Remote Sensing Letters, 1, doi: 10.1109/LGRS.2004.829613.

Liu, Zhaoyan, and coauthors, 2009: The CALIPSO Lidar Cloud and Aerosol Discrimination: Version 2 Algorithm and Initial Assessment of Performance, *J. Atmos. Oceanic Technol.*, 26, 1198–1213.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L. W. M. J., and Merchant, J. W., 2000: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing, 21(6-7), 1303-1330.

Marenco, F., V. Santacesaria, A. F. Bais, D. Balis, A. di Sarra, A. Papayannis, and C. Zerefas, (1997), Optical properties of tropospheric aerosols determined by lidar and spectrophotometric measurements (Photochemical Activity and Solar Ultraviolet Radiation campaign), *Appl. Opt.*, **36**, 6875-6886.

McGill, M. J., M. Marzouk, V. S. Scott, and J. D. Spinhirne, 1997c: Holographic circle-to-point converter with particular applications for lidar work, *Opt. Eng.*, **36**, 2171–2175.

McGill, M.J., and R.D. Rallison, 2001: Holographic optics convert rings to points for Detection. *Laser Focus World*, **37**, 131-136.

McGill, M. J., D. Hlavka, W. Hart, V. S. Scott, J. Spinhirne, and B. Schmid, 2002: Cloud Physics Lidar: instrument description and initial measurement results. *Applied Optics*, **41**, 3725-3734.

McGill, M.J., D.L. Hlavka, W.D. Hart, E.J. Welton, and J.R. Campbell, 2003: Airborne lidar measurements of aerosol optical properties during SAFARI-2000. *Journal of Geophysical Research*, 108, doi: 10.1029/2002JD002370.

McGill, M. J., M. A. Vaughan, C. R. Trepte, W. D. Hart, D. L. Hlavka, D. M. Winker, and R. Kuehn, 2007: Airborne validation of spatial properties measured by the CALIPSO lidar. *J. Geophys. Res.*, **112**, D20201, doi:10.1029/2007JD008768.

Nowottnick, E. P., Colarco, P. R., Welton, E. J., and da Silva, A., 2015: Use of the CALIOP vertical feature mask for evaluating global aerosol models, Atmos. Meas. Tech., 8, 3647-3669, doi:10.5194/amt-8-3647-2015.

Omar, Ali H., and coauthors, 2009: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *J. Atmos. Oceanic Technol.*, 26, 1994–2014.

Palm, S., W. Hart, D. Hlavka, E. J. Welton, A. Mahesh, and J. Spinhirne, 2002: GLAS atmospheric data products. NASA Goddard Space Flight Center Geoscience Laser Altimeter System Algorithm Theoretical Basis Document Version 4.2, 141 pp.

Platt, C. M. R. (1979), Remote sounding of high clouds: I. Calculation of visible and infrared optical properties from lidar and radiometer measurements, *J. Appl. Meteor.*, **18**, 1130-1143.

Powell, Kathleen A., and coauthors, 2009: CALIPSO Lidar Calibration Algorithms. Part I: Nighttime 532-nm Parallel Channel and 532-nm Perpendicular Channel, *J. Atmos. Oceanic Technol.*, 26, 2015–2033.

Privé, N. C., R. M. Errico, and K-S. Tai, 2013: Validation of the forecast skill of the

Global Modeling and Assimilation Office observing system simulation experiment. Quarterly Journal of the Royal Meteorological Society 139.674, 1354-1363.

Rienecker, M. M. et al., 2008: The GEOS-5 Data Assimilation System–Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0. Technical Report Series on Global Modeling and Data Assimilation, 104606, 92 pp.

Rodier, S., M. Vaughan, S. Palm, J. Yorks, M. McGill, M. Jensen, T. Murray, K.-P. Lee and C. Trepte, 2015: Laser Remote Sensing from ISS: CATS Cloud and Aerosol Data Products, proceedings of the ILRC 2015, New York, New York.

Rogers, R. R., Hair, J. W., Hostetler, C. A., Ferrare, R. A., Obland, M. D., Cook, A. L., Harper, D. B., Burton, S. P., Shinozuka, Y., McNaughton, C. S., Clarke, A. D., Redemann, J., Russell, P. B., Livingston, J. M., and Kleinman, L. I., 2009: NASA LaRC airborne high spectral resolution lidar aerosol measurements during MILAGRO: observations and validation. *Atmos. Chem. Phys.*, 9, 4811-4826, doi:10.5194/acp-9-4811-2009.

Russell, P. B., T. J. Swissler, and M. P. McCormick, 1979: Methodology for error analysis and simulation of lidar aerosol measurements, Appl. Opt., 18, 3783–3797.

Sassen, K., and S. Benson, 2001: A midlatitude cirrus cloud climatology from the Facility for Atmospheric Remote Sensing. Part II: Microphysical properties derived from lidar depolarization. J. Atmos. Sci., 58, 2103– 2112.

Sassen, K., 2002: Cirrus: A modern perspective. *Cirrus*, D. K. Lynch *et al.*, Eds., Oxford University Press, 11–40.

She, C., 2001: Spectral structure of laser light scattering revisited: bandwidths of nonresonant scattering lidars", Appl. Opt., 40, 4875-4884.

Spinhirne J. D., J. A. Reagan, and B. M. Herman, 1980: Vertical distribution of aerosol extinction cross section and inference of aerosol imaginary index in the troposphere by lidar technique. *J. Appl. Meteor.*, **19**, 426–438.

Spinhirne, J. D., M. Z. Hansen, and L. O. Caudill, 1982: Cloud top remote sensing by airborne lidar. Appl. Opt., vol. 21, pp. 1564–1571.

Spinhirne, J. D. and W. D. Hart (1990), Cirrus structure and radiative parameters from airborne lidar and spectral radiometer observations: the 28 October 1986 FIRE study. *Mon. Wea. Rev,* **118**, 2329-2343.

Spinhirne, J. D., W. D. Hart, D. L. Hlavka (1996), Cirrus infrared parameters and shortwave reflectance relations from observations, *J. of Atmos. Sci.*, **53**, 1438-1458.

Spinhirne, J. D., J. R. Campbell, D. L. Hlavka, R. A. Ferrare, D. D. Turner, and C. J. Flynn (1999), Aerosol retrieval comparison during the SGP summer '98 IOP from Multiple Lidar Probing", Poster Abstract, Atmospheric Radiation Measurement (ARM) Science Team Meeting, San Antonio, Texas, March 22-26, 1999.

Tenti, G., C. D. Boley, and R. C. Desai, 1974: On the kinetic model description of Rayleigh–Brillouin scattering from molecular gases. *Can. J. Phys*, **52**, 285–290.

Vaughan, M. A., and coauthors, 2004: Fully automated analysis of space –based lidar data: An overview of the CALIPSO retrieval algorithms and data products, *Proc. SPIE*, 5575, 16-30.

Vaughan, M. A., D. M. Winker, and K. A. Powell, 2005: Part 2: Feature Detection and Layer Properties Algorithms. CALIOP Algorithm Theoretical Basis Document PC-SCI-202.01, 87 pp.

Vaughan, M., and coauthors, 2009: Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements, *J. Atmos. Oceanic Technol.*, 26, 2034–2050.

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, doi:10.1029/2009JD013086.

Vermeille, H., 2002: Direct Transformation from geocentric coordinates to geodetic coordinates. Journal of Geodesy, 76:451-454.

Vigroux, E., 1953: Contribution a l'etude experimentale de l'absorption de l'ozone, Ann. Phys., 8, 709-761.

Winker, D. M., 2003: Accounting for multiple scattering in retrievals from space lidar. *Lidar Scattering Experiments*, C. Werner, U. Oppel, and T. Rother, ed., International Society for Optical Engineering (SPIE Proceedings, Vol. 5059), 128–139.

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**, 1105-1119, doi:10.1175/2009JTECHA1281.1.

Yorks, J. E., D. L. Hlavka, W. D. Hart, M. J. McGill, 2011a: Statistics of Cloud Optical Properties from Airborne Lidar Measurements. *J. Atmos. Oceanic Technol.*, **28**, 869–883. doi: http://dx.doi.org/10.1175/2011JTECHA1507.1

Yorks, J. E., D. L. Hlavka, M. A. Vaughan, M. J. McGill, W. D. Hart, S. Rodier, and R.

Kuehn, 2011b: Airborne validation of cirrus cloud properties derived from CALIPSO lidar measurements: Spatial properties, *J. Geophys. Res*., **116**, D19207, doi:10.1029/2011JD015942.

Yorks, J. E., M. McGill, V. S. Scott, A. Kupchock, S. Wake, D. Hlavka, W. Hart, P. Selmer, 2014: The Airborne Cloud-Aerosol Transport System: Overview and Description of the Instrument and Retrieval Algorithms, *J. Atmos. Oceanic Technol*., 31, 2482–2497, doi:10.1175/JTECH-D-14-00044.1.

Young, A.T., 1981: Rayleigh scattering. *Phys. Today*, 35, 42–48.

Young, S. A., 1995: Analysis of lidar backscatter profiles in optically thin clouds, Appl. Opt., 34, 7019–7031.

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.

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