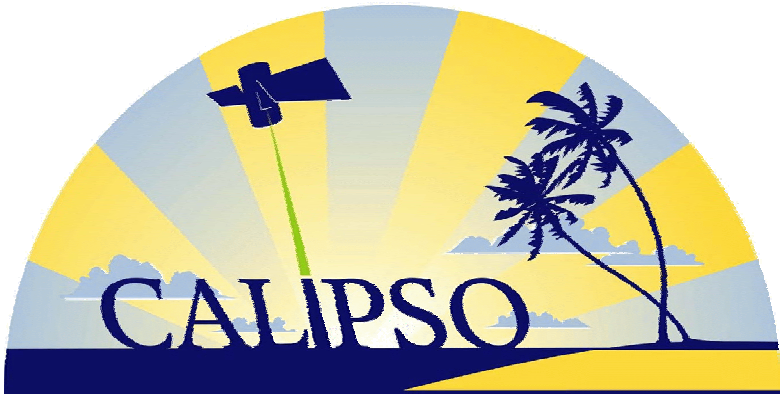
CALIOP Algorithm Theoretical Basis Document

Calibration and Level 1 Data Products



Primary Authors:

Chris A. Hostetler, NASA Langley Research Center, Hampton, Virginia, USA

Zhaoyan Liu, National Institute of Aerospace (NIA), Hampton, Virginia, USA

John Reagan, University of Arizona, Tucson, Arizona, USA

Mark Vaughan, Science Applications International Corp. (SAIC), Hampton, Virginia, USA

David Winker, NASA Langley Research Center, Hampton, Virginia, USA

Mary Osborn, Science Applications International Corp. (retired), Hampton, Virginia, USA

William H. Hunt, Science Applications International Corp. (SAIC), Hampton, Virginia, USA

Kathleen A. Powell, Science Applications International Corp. (SAIC), Hampton, Virginia, USA

Charles Trepte, NASA Langley Research Center, Hampton, Virginia, USA

PC-SCI-201

Release 1.0

7 April 2006

Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observations

CALIOP Algorithm Theoretical Basis Document

Document No: PC-SCI-201

Prepared By:

Chris Hostetler	Date
Zhaoyan Liu	Date
John Reagan	Date
Mark Vaughan	Date
Mary Osborn	Date
William H. Hunt	Date
Kathleen A. Powell	Date
Charles Trepte	Date

Approved By:

David M. Winker Date
CALIPSO Principal Investigator

Table of Contents

1. Introduction

1.1. Purpose

This document describes the algorithms that will be used to calibrate the lidar backscatter profiles acquired by the CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) instrument flown aboard the CALIPSO satellite. The outputs of these algorithms are Level 1 data, consisting of attenuated backscatter coefficient profiles for the three lidar channels along with information on the uncertainties in these products. These data are used by the Level 2 algorithms to produce geophysical parameters such as layer heights and optical depths. 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	Section(s) Affected
2006/03/01	1.0	Initial release	Chris Hostetler	All

1.3. CALIPSO Mission Overview

Current uncertainties in the roles played by clouds and aerosols in the Earth radiation budget limit our understanding of the climate system and the potential for global climate change. Unlike greenhouse gases, tropospheric aerosols are highly variable in space and time, and satellite observations are required to understand the distribution and impacts of aerosols on regional and global scales. Advances in modeling capabilities to predict climate change also require improved representations of cloud processes and decreased uncertainties in cloud-radiation interactions. While models must be used to estimate the impacts of aerosols on the climate in the past and to predict future trends, model-based estimates of aerosol forcing are highly uncertain, largely because current capabilities to observe the global distribution and properties of aerosols are insufficient to constrain several of the key assumptions incorporated into the models. CALIPSO will provide critical observations of the vertical distribution of aerosols, an ability to perform height-resolved discrimination of aerosols into several types, and an improved capability to observe aerosols over bright and heterogeneous surfaces.

The sensitivity of the climate to forcings from aerosols and greenhouse gases is largely controlled by interactions between clouds and radiation. Advances in model capabilities to predict climate change requires improved representations of cloud processes and decreased uncertainties in cloud-radiation interactions. The largest uncertainties involve the use of models to (a) predict cloud properties based on atmospheric state, and (b) use of these cloud properties to calculate radiative energy fluxes. In particular, the largest source of uncertainty in estimating longwave radiative fluxes at the surface and within the atmosphere is connected with current difficulties in determining the vertical distribution and overlap of multilayer clouds.

The CALIPSO mission builds on the experience of LITE, which flew a three-wavelength lidar on the Space Shuttle in 1994 (Winker et al., 1996). The CALIPSO payload consists of the

Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), a two-wavelength polarization-sensitive lidar, the Infrared Imaging Radiometer (IIR), which has three channels in the thermal infrared, and the Wide Field Camera (WFC) with a single channel at 650 nm. Data from these instruments will be used to measure the vertical distributions of aerosols and clouds in the atmosphere, as well as optical and physical properties of aerosols and clouds which influence the Earth's radiation budget. CALIPSO will provide data to address three major objectives:

- to improve observationally-based estimates of direct and indirect aerosol radiative forcing;
- to improve characterization of surface radiative fluxes and atmospheric heating rates; and
- to improve model parameterizations of cloud-climate feedbacks.

CALIPSO will also address a number of secondary objectives, which include observing long-range transport of pollutants, providing coincident measurements to validate and improve retrievals from other instruments within the A-train, and providing aerosol observations useful for atmospheric chemistry applications.

CALIOP will provide global, vertically-resolved measurements of aerosol spatial distributions and aerosol extinction coefficients, and an ability to perform height-resolved discrimination of aerosol into several types. CALIOP can observe aerosol over bright surfaces and beneath thin clouds as well as in clear sky conditions. CALIOP will also provide vertical profiles of single- and multi-layer transmissive clouds. Lidar polarization information will provide profiles of cloud ice/water phase, allowing a determination of the vertical distribution of cloud ice and water. IIR and WFC data will be used to retrieve cloud emissivity and effective particle size. Lidar data is incorporated into this retrieval algorithm to provide constraints to improve the retrieval performance.

1.4. CALIPSO Data Product Levels

The data products generated from the CALIOP measurements are produced according to a protocol which is similar to, but not exactly the same as, that established by NASA's Earth Observing System (EOS). The data product levels for CALIPSO are defined as follows:

- **Level 0**: reconstructed, unprocessed instrument/payload data at full resolution; any and all communications artifacts, e.g. synchronization frames, communications headers, duplicate data removed.
- **Level 1A**: reconstructed, unprocessed instrument data at full resolution which is time-referenced, geo-located, corrected for instrument artifacts, and annotated with ancillary information. CALIPSO Level 1A data is an internal product only and is not archived.
- **Level 1B**: Level 1A data that have been processed to sensor units and archived as Level 1 data.
- **Level 2**: geophysical variables derived from Level 1 data, including those derived using measurements from multiple CALIPSO instruments.
- **Level 3**: Geophysical variables mapped onto uniform space-time grids.
- **Level 4**: Variables derived from measurements on multiple satellites.

In the CALIOP Level 1A data processing, the raw Level 0 data (i.e., the downlinked data, refer to Section 2.3) is converted from engineering format to standard binary format (e.g., IEEE floating point), subtracted for baseline shape, linearized, geolocated, and normalized to laser energy and amplifier variable gain. However, the Level 1A data processing is not addressed further in this document. All the algorithms described in this document concern the Level 1B data processing, which consists primarily of calibration of the three CALIOP channels.

2. Instrument Description

CALIOP consists of a laser transmitter subsystem and a receiver subsystem. Figure 2.1 shows an exploded view of the "lidar core" contained inside the payload housing. The instrument is built around a T-shaped optical bench which assures stability of the transmitter-to-receiver alignment. The lidar receiver telescope is attached to one side of the bench with the receiver optics and detector assemblies on the other side. The laser transmitter assembly is attached to the top of the "T" by a precision linear drive mechanism and gimbal assembly, allowing precise and accurate pointing adjustments.

2.1. Transmitter subsystem

The laser transmitter subsystem includes two identical, redundant laser transmitters, each with a beam expander, and a beam steering system that ensures alignment between the transmitter and receiver. Only one laser is operated at a time. The lasers produce simultaneous pulses at 1064 nm and 532 nm at a pulse repetition rate of 20.16 Hz. The lasers are Q-switched to provide a pulse length of about 20 ns. Each laser generates nominally 220 mJ per pulse at 1064 nm, which is frequency-doubled to produce about 110 mJ of pulse energy at each of the two wavelengths. The output pulse energy at each wavelength is measured using energy monitors located within each canister. Beam expanders reduce the angular divergence of the transmitted laser beam to produce a beam diameter of 70 meters at the Earth's surface (corresponding to a nominal laser beam divergence of 100 μrad). Transmitter specifications are summarized in Table 2.1.

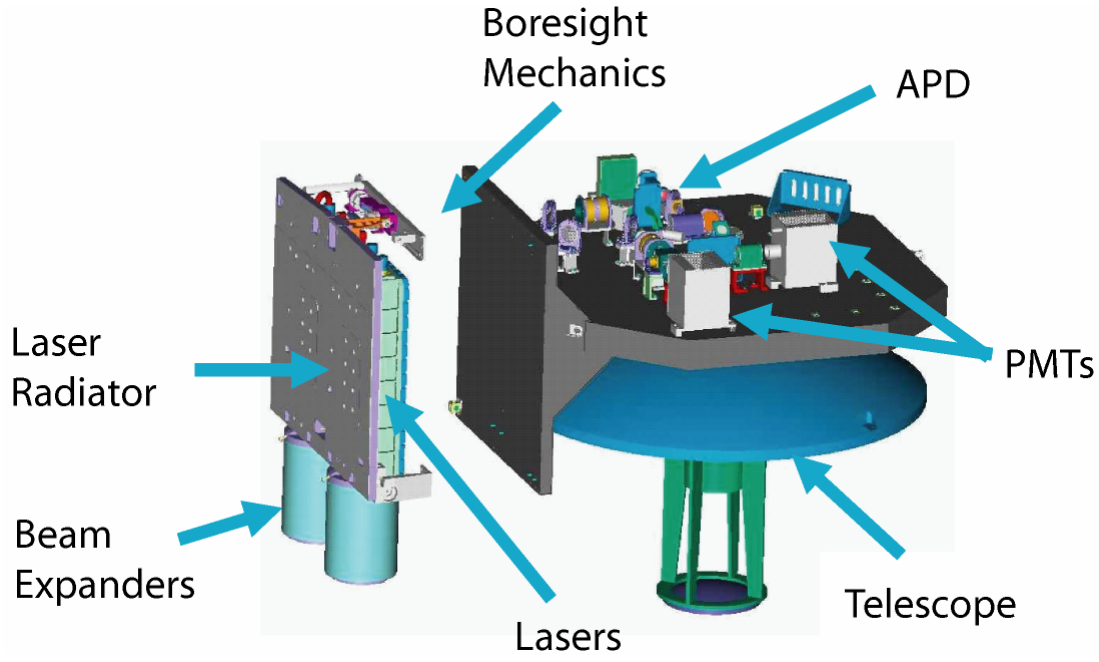


Figure 2.1 CALIOP transmitter and receiver subsystems.

Table 2.1 CALIOP transmitter characteristics

Laser	Diode-pumped Nd:YAG
Pulse Energy	110 mJ: 532 nm 110 mJ: 1064 nm
Rep Rate	20.16 Hz
Pulse Length	20 nsec
Linewidth	30 pm at 532 nm 100 pm at 1064 nm
Polarization Purity	> 1000:1 (532 nm)
Nominal Beam Divergence	100 μrad (after beam expander) [a]
Boresight Range	±1 degree, 1.6 μrad steps
Laser Environment	18 psia, dry air

[a] Measurements on the beam expanding optics indicate that the actual divergences are somewhat larger than the nominal value, ranging from 110 to 115 μrad at 532 nm and 130 to 140 μrad at 1064 nm.

2.2. Receiver subsystem

Shown schematically in Figure 2.2, the receiver sub-system consists of the telescope, relay optics, detectors, preamps, and line drivers, all mounted on a stable optical bench. The completed payload is shown in Figure 2.3. Signal processing and control electronics are contained in boxes mounted on the payload housing. The receiver telescope is an all-beryllium 1-meter diameter design similar to the telescope built for the GLAS instrument on the ICESat satellite. The telescope primary mirror, secondary mirror, metering structure, and inner baffle are all made of beryllium, for lightness and to minimize the effect of thermal gradients. A light shade prevents direct solar illumination of the mirrors. The telescope is thermally isolated from the optical bench. A field stop at the focus of the telescope defines the receiver field of view of 130 μrad (full angle) and also rejects stray light. A movable shutter placed downstream from the focus can be used to block incoming light to allow measurements of detector dark current. The shutter mechanism also allows a pseudo-depolarizer (McGuire and Chapman, 1990) to be moved into the 532 nm beam for depolarization calibration. A polarization beamsplitter is used to separate the 532 nm parallel and perpendicular returns. A narrowband etalon is used in combination with a dielectric interference filter in the 532 nm channel to reduce the solar background illumination, while an interference filter alone provides sufficient solar rejection for the 1064 nm channel. Photomultiplier tubes (PMTs) are used for the 532 nm detectors as they provide large linear dynamic range, very low dark noise, and reasonable quantum efficiency. An avalanche photodiode (APD) is used at 1064 nm. The APD has good dynamic range and quantum efficiency, but the dark noise is much larger than for the PMTs. Thus, the 532 nm channels have lower noise-equivalent power.

The magnitude of lidar return signal spans a very large range. CALIOP is required to accurately measure signal returns from the aerosol-free region between 30 km and 35 km as well as the

strongest cloud returns. For this reason, all detectors are used in analog mode followed by dual 14-bit digitizers which provide an effective linear dynamic range of ~22 bits. The electronic gains of the 532 nm channels are set high enough to allow detection of single photoelectron events. Table 2.2 lists specifications of the receiver subsystem.

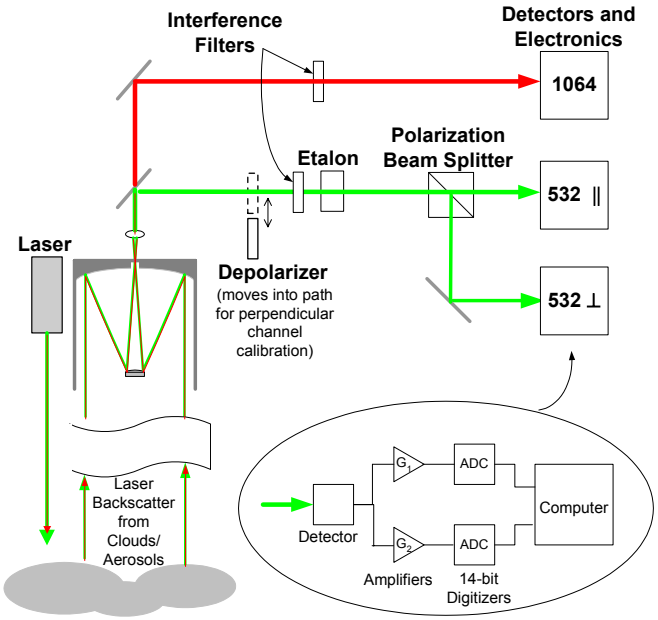


Figure 2.2 Functional block diagram of CALIOP.

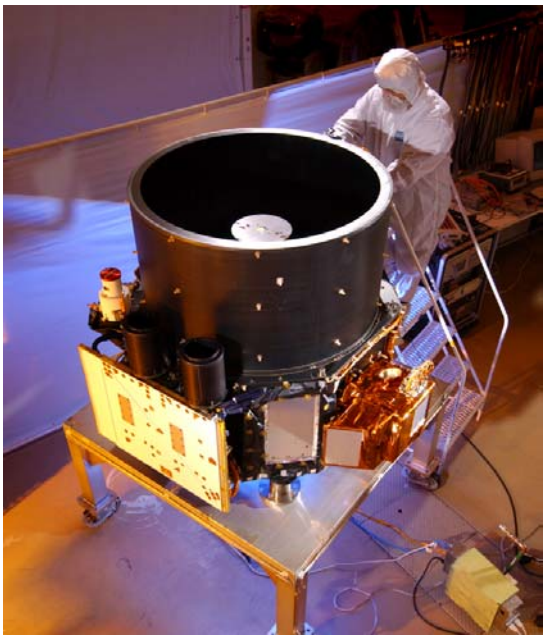


Figure 2.3 Completed CALIOP payload under test in BATC facilities.

Table 2.2 CALIPSO receiver parameters

Telescope diameter	1 meter
Field of View	130 μrad
Digitization Rate	10 MHz
Linear Dynamic Range	4 E+6 : 1
532 nm Channel:	
Detector	PMT
Etalon Passband	37 pm
Etalon Peak Transmission	85%
Blocking Filter	770 pm
1064 nm Channel:	
Detector	APD
Optical Passband	450 pm
Peak Transmission	84%

2.3. Data Acquisition and Signal Processing

A number of functions are performed by the instrument to convert the analog detector signals into the profiles which are downlinked. These include range determination, background subtraction, digitization, merging, and averaging. There are two opposing drivers on data processing: to maximize spatial resolution and dynamic range of the signal while simultaneously minimizing the telemetry data volume. Several features were implemented to reduce the required telemetry bandwidth by more than an order of magnitude relative to the raw data, while impacting the information content of the data as little as possible.

Due to the oblateness of the Earth, the range from the circular orbit of the CALIPSO satellite to mean sea level (MSL), defined by the height of the geoid, varies by about 21 km through the orbit. The geoid is an equipotential surface which (approximately) coincides with mean ocean surface and the geoid height (undulation) is a height relative to the surface of an ellipsoid Earth model (see, for example, Li and Götze, 2001), which is based on the NASA EGM96 360th order gravity model. The rate of change of the range to MSL is as much as 22 meters/sec. To allow on-orbit averaging of profile data while maintaining the vertical resolution of 30 meters the data acquisition timing must be adjusted to account for the changing range to the Earth's surface. The Payload Controller makes a real-time determination of the range to MSL for each laser shot using an on-board geoid model and an orbit propagator, which is updated using data from the spacecraft GPS. The data acquisition timing is then adjusted so that each profile will have the same altitude registration with respect to the geoid.

Data acquisition timing is illustrated in Figure 2.4 in terms of height above MSL. A timer is started when the laser fires. When the laser pulse reaches an altitude of 115 km above MSL, the PMT detectors are gated on and the profile signals from all three channels are acquired (the APD detector has no gate and is always on). The analog profiles are sampled at 10 MHz (corresponding to 15-meter range interval) until the elapsed time corresponds to a range of 18.5 km below sea level, at which point the PMTs are gated off and the digitizers stop sampling. The samples acquired between 40 km (30 km for the 1064 nm channel) and –2 km are used to create the profile data which are downlinked. The portions of the profile above 60 km and below –11 km are used to measure DC signal levels.

The solar background signal can be significant – as large as the clear-sky atmospheric signal. The instrument measures the DC background of each profile from the signal acquired between 112 km and 97 km, where the laser backscatter signals are negligible. This DC signal is electronically subtracted from the analog profile before digitization to allow the dynamic range of the digitizer to be used most effectively. This subtraction will result in negative-going noise excursions if the laser backscatter signal is small. These negative spikes will be clipped by the digitizers producing a bias of the mean signal. To avoid this, a fixed electrical offset is added, prior to digitization, to the portion of the profile below 97 km. The magnitude of this offset is accurately measured using the 1000 15-meter samples acquired between 75.3 km and 60.3 km, which are averaged and downlinked as a single value. This average of the upper background region is numerically subtracted from the profile during later processing. The 500 15-meter samples between –11 km and –18.5 km are also averaged and downlinked, providing another measure of the offset.

Two 10 MHz 14-bit analog-to-digital converters (ADCs), set for different gains, are used in each channel to provide the required 22-bit effective dynamic range. On each channel, the high gain

ADC measures weak signals and the low gain ADC acquires signals which saturate the high gain digitizer. The profile samples are taken from the high-gain ADC if they are on-scale. If a sample is saturated on the high-gain ADC, the corresponding sample from the low-gain ADC is used. The outputs of each pair of digitizers are re-scaled and merged into a single profile before being downlinked. The fundamental sampling resolution of the lidar is 30 meters vertical and 333 meters horizontal, determined by the receiver electrical bandwidth and the laser pulse repetition rate. Therefore, each pair of adjacent 15-meter samples is averaged to produce a profile of 1400 30-meter samples extending from 40 km to -2 km. 1064 nm profiles extend only from 30 km to -2 km, as 1064 nm returns from the purely molecular atmosphere above 30 km are negligibly small.

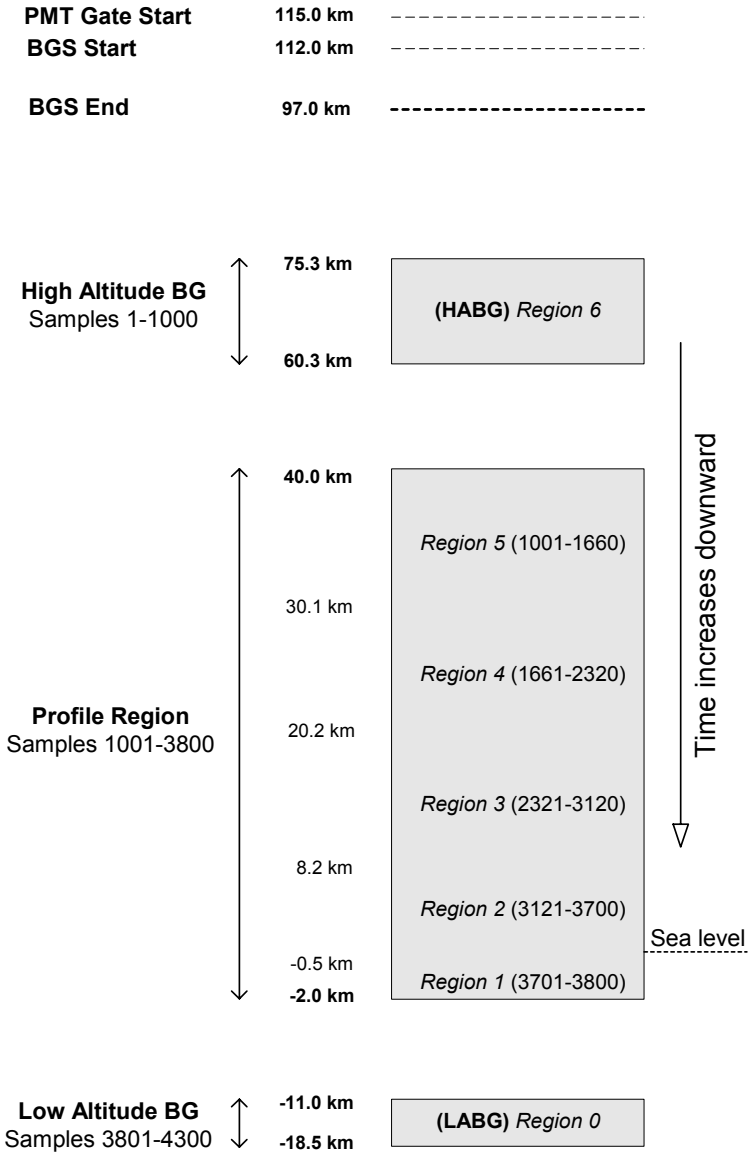


Figure 2.4 CALIPSO on-orbit data acquisition timing. Sample numbers refer to the number of 15-meter range bins.

The atmosphere becomes more spatially uniform with increasing altitude. Further, signals from higher in the atmosphere tend to be weaker and require more averaging. Therefore, an altitude-dependent on-board averaging scheme was developed which provides full resolution in the lower troposphere – where the spatial variability of cloud and aerosol is greatest – and lower resolution in the higher atmosphere. The degree of averaging varies with altitude, as detailed in Table 2.3 (altitudes are with respect to the geoid).

The averaging scheme provides "full resolution" (30 meters vertical, 0.33 kilometers horizontal) through the depth of the planetary boundary layer and as much of the lower troposphere as possible. The lowest land elevation is the shore of the Dead Sea, with a depression of 392 meters below sea level. Therefore, the full-resolution region extends to -0.5 km. The region from -0.5 km to -2.0 km is provided at a vertical resolution of 300 meters to allow monitoring of potential delayed recovery from the surface return. Data between 8.2 km and 20.2 km has a vertical resolution of 60 meters, which allows several vertical samples within even the thinnest cirrus layers. The stratosphere is much more spatially homogeneous than the troposphere. Therefore, between 20.2 km and 30.1 km the vertical resolution is degraded to 180 meters, as stratospheric aerosols and polar stratospheric clouds found in this region tend to have much greater vertical extent than the thin cirrus layers found in the upper troposphere. Only coarse spatial resolution is required above 30 km, as the aerosol concentration is near zero. Molecular returns from this region are used for calibration of the 532 nm parallel channel.

Table 2.3 Spatial resolution of downlinked data

Altitude Range (km)	Horizontal Resolution (km)	532 nm Vertical Resolution (m)	1064 nm Vertical Resolution (m)
30.1 to 40.0	5.0	300	---
20.2 to 30.1	1.67	180	180
8.2 to 20.2	1.0	60	60
-0.5 to 8.2	0.33	30	60
-2.0 to –0.5	0.33	300	300

3. Overview of Level 1 Lidar Calibration Algorithms

The calibration algorithms are designed to accomplish two major functions: (1) the determination of calibration coefficients for the three lidar channels and (2) the application of those calibration coefficients to produce attenuated backscatter profiles used in Level 2 processing. An overview of the calibration algorithms is shown in Figure 3.1 and described below.

Determination of the calibration coefficients is basically a three-step process applied to the Level 1A data products:

(a) The calibration coefficient is determined for the 532 nm parallel channel. For the baseline approach, this is done by comparing the measured 532 nm parallel channel signal from the 30-34 km region to an estimate of the parallel backscatter coefficient computed from a modeled atmospheric density profile. The 30-34 km altitude range is chosen because there is little aerosol at that height range, especially in mid and high latitudes. At low latitudes, the aerosol contribution can be reasonably well represented and extracted so that virtually all the backscatter is from molecules. The molecular backscatter coefficients can be well estimated using knowledge of the molecular number density and theoretically derived estimates of the molecular backscatter cross section (Reagan, et al., 2002).

(b) The 532 nm perpendicular channel is then calibrated relative to the calibration obtained for the parallel channel. There is not enough signal to calibrate the perpendicular channel using stratospheric molecular returns, because the depolarization of clear-air 180-backscatter is much less than 1%. The calibration is therefore transferred from the parallel to the perpendicular channel using data collected during the Polarization Gain Ratio (PGR) operation.

(c) Calibration of the 532 nm parallel and perpendicular channels is then transferred to the 1064 channel. As with the 532 nm perpendicular channel, the signal from the 1064 channel in the mid-stratosphere is too low to provide a reliable calibration measurement. Transfer of calibration from the 532 nm channels to the 1064 nm channels is accomplished using the backscatter from properly chosen cirrus clouds. Because cirrus cloud particles are large, the ratio of the 532 nm and 1064 nm backscatter coefficients and the ratio of the 532 nm and 1064 nm extinction coefficients are both nearly equal to 1. The 1064 nm calibration coefficient is determined by comparing the 1064 nm backscatter signal with the calibrated 532 nm cirrus backscatter measurements (Reagan, et al., 2002).

These three steps are defined in detail in Sections 4 through 7. The calibration coefficients produced are stored in the Calibration Data Product File. The complete contents of this file are described in the CALIPSO Data Products Catalog (PC-SCI-503).

In addition to accomplishing the two functions described above, the Level 1 processing also computes a noise scale factor (NSF) that is used to estimate random errors due to shot noise in the lidar measurements. The random uncertainties are then estimated for the computed calibration coefficients. The NSF computation is described in Section 8. The details of the application of the NSF to the estimation of random uncertainties contained in the calibration coefficients can found in Sections 4 through 7.

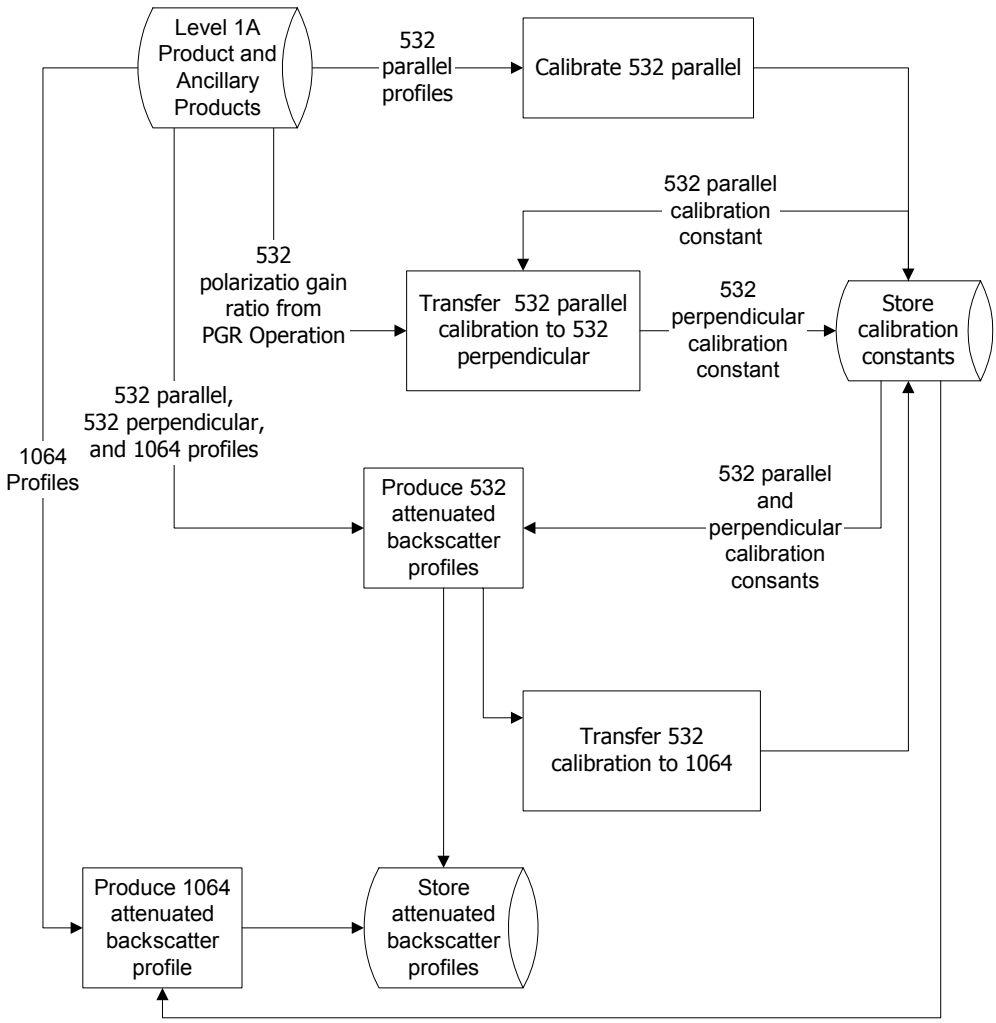


Figure 3.1 Block diagram representing the basic flow of the Level 1B processes.

3.1. Basic Equations and Nomenclature

In this section we present the basic equations that represent the lidar profile data and define the nomenclature common to the sections that follow.

3.1.1. General Form of Lidar Equation

Prior to calibration processing the lidar signal has been corrected for offset voltages, background signals, and instrument artifacts. The resulting signal to be calibrated can be written as

$$P(r) = \frac{1}{r^2} E_0 \, \xi \, \beta(r) \, T^2(r), \tag{3.1}$$

where

$$T^2(r) = \exp\left[-2\int_0^r \sigma(r')\,dr'\right] \tag{3.2}$$

and

r = range from the satellite to the sampled volume

P = measured signal after background subtraction and artifact removal

E_0 = average laser energy for the single-shot or composite profile

ξ = lidar system parameter

β = volume backscatter coefficient at range r

T = one-way transmittance from the lidar to the scattering volume at range r

σ = volume extinction coefficient at range r.

In analyzing the data, we work with profiles composed of returns from a single laser shot or returns composed from the average of several laser shots. The parameters in the lidar equation can change from profile to profile, which we can indicate explicitly with a profile index, k, as follows:

$$P(r,k) = \frac{1}{r^2} E_0(k)\, \xi(k)\, \beta(r,k)\, T^2(r,k) \tag{3.3}$$

The signal is written in term of the range from the satellite, r; however, this notation can lead to confusion. The atmospheric parameters, β and σ, are more properly considered functions of altitude, z, above mean sea level. The one-way transmittance is a function of optical path length through the attenuating atmospheric volume above the scattering volume and, hence, can be considered a function of either altitude or range. For off-nadir measurements, a recasting of the above equation in more explicit terms is appropriate:

$$P(z,k,\theta) = \left(\frac{\cos(\theta)}{z_{sat}(k) - z} \right)^2 E_0(k)\ \xi(k)\ \beta(z,k) \exp\left[-2 \int_0^r \sigma\left(z(r',k),k\right) dr' \right]. \tag{3.4}$$

Here

$$z(r,k) = z_{sat}(k) - r\cos\left(\theta(k)\right) \tag{3.5}$$

and

$\theta(k)$ = average off-nadir angle for the k^{th} laser shot or composite profile

$z_{sat}(k)$ = average altitude of the satellite for the k^{th} laser shot or composite profile.

The nominal off-nadir angle for the CALIPSO mission is 0.3˚, and for this small value the correction for the cosine of the angle is insignificant. However, for larger angles, correcting the optical path lengths by the cosine of the angle can be important.

The off-nadir pointing direction is aligned with the satellite flight vector such that the laser footprint at ground level is along the satellite ground track ahead of the satellite subnadir location.

3.1.2. Lidar System Parameter and Calibration Coefficient

The lidar system parameter is the product of many instrument parameters that govern the sensitivity and efficiency of the lidar. Factors making up the lidar system parameter include the receiver area, the receiver optical efficiency, the detector quantum efficiency, the electronic gain

in the detection electronics, and the geometric overlap between the laser footprint and the receiver field of view (boresight overlap). Many of these factors vary with time. Some of this temporal variability is in response to commanded and autonomous instrument operations, while the remainder is due to non-ideal instrument effects and environmental changes. Examples of commanded changes include changes in amplifier gain and boresight position. The amplifier gain will be changed twice per orbit in response to changing lighting conditions as the satellite crosses the terminator. The boresight overlap will change abruptly upon a commanded change to the boresight operating position. Examples of non-ideal instrument effects include changes in detector sensitivity and boresight drift. The detector sensitivity is expected to change gradually over the course of the mission due to degradation of the detectors. The boresight overlap may also change slowly due to subtle thermal/mechanical effects. Amplifier gains were well characterized prior to launch and changes in the gain are accounted for in post processing. Other components of the lidar system parameter must be measured during the on-orbit calibration process.

In processing the data, we separate the lidar system parameter into two factors that are accounted for separately:

$$\xi(k) = G_A(k)\ \ C(k)\,, \tag{3.6}$$

where

$G_A(k)$ = amplifier gain in effect for profile k

$C(k)$ = lidar calibration coefficient (or calibration constant).

The amplifiers are commanded to one of a set of discrete values, which are all determined prior to launch; hence, the amplifier gain is known and is downlinked with the data.

3.1.3. Range-scaled, Energy-Normalized, Gain-Normalized Profile

Much of the algorithm description throughout this document is posed in terms of the range-scaled, energy normalized, and gain-normalized versions of the signal, P, which can be written as

$$X(z,k) = \frac{r^2\ P(z,k)}{E(k)\ G_A(k)} = C(k)\ \beta(z,k)\ T^2(z,\theta,k)\,. \tag{3.7}$$

In the sections that follow, we analyze the algorithms in terms of the range-scaled, energy normalized, gain-normalized signal, X; however, to make the equations less cumbersome, we simplify the notation by

$$X(z) = C\ \beta(z)\ T^2(z)\,, \tag{3.8}$$

while maintaining an understanding of the implicit dependencies on profile index and off-nadir angle. All of the calibration operations that are discussed in the following sections use the range-scaled, energy-normalized, gain-normalized signal X.

The raw lidar data consists of profiles for the three lidar channels. The range scaled, energy and gain normalized profiles for the 532 parallel, 532 perpendicular, and 1064 channels are defined, respectively, as follows:

$$X_{532,\parallel}(z) = C_{532,\parallel}\ \beta_{532,\parallel}(z)\, T_{532}{}^{2}(z) \tag{3.9}$$

$$X_{532,\perp}(z) = C_{532,\perp}\ \beta_{532,\perp}(z)\, T_{532}{}^{2}(z) \tag{3.10}$$

$$X_{1064}(z) = C_{1064}\ \beta_{1064}(z)\, T_{1064}{}^{2}(z) \tag{3.11}$$

In the sections that follow, some of the subscripts may be dropped when there is little chance of confusion. In particular, for the 532 perpendicular and parallel parameters, the "532" is frequently omitted from the subscript.

3.1.4. Output Data Products

The data products archived upon the completion of the Level 1 processing include profile products and calibration products.

The profile products are

- 532 nm total attenuated backscatter coefficients
- 532 nm perpendicular attenuated backscatter coefficients
- 1064 nm total attenuated backscatter coefficient

as defined by the following equations:

$$\beta'_{532,\,Total}(z) = \left[\beta_{\parallel}(z) + \beta_{\perp}(z)\right] T_{532}{}^{2}(z) \tag{3.12}$$

$$\beta'_{532,\perp}(z) = \beta_{\perp}(z)\, T_{532}{}^{2}(z) \tag{3.13}$$

$$\beta'_{1064}(z) = \beta_{1064}(z)\ T_{1064}{}^{2}(z) \tag{3.14}$$

The calibration products are composed of several files which archive final and intermediate calibration coefficient estimates, statistics of the calibration coefficients, and various parameters relevant to the calibration process (see PC-SCI-503, CALIPSO Data Management System Data Products Catalog).

4. 532 nm Parallel Channel Calibration

4.1. Algorithm Description

4.1.1. General Description

The 532 nm parallel channel is calibrated by comparing the magnitude of the measured signal to the predicted backscatter from a region in the atmosphere for which an accurate independent estimate of the backscatter coefficient is available. The altitude range extending from 30-km to 34-km will be initially used to calculate the CALIOP 532 nm parallel channel calibration coefficient. In addition, calculations will also be done in four other diagnostic ranges (see Table 4.1) below and above this default height range. Two factors were considered in selecting this range: choosing a lower region leads to increased uncertainties due to additional aerosol backscatter; choosing a higher region leads to increased uncertainties in the resulting calibration coefficient due to the decreasing molecular backscatter, resulting in lower SNR. The selection of the calibration altitude range was optimized by balancing these two factors. The location of the calibration altitude range will be reassessed after launch.

An estimate of the relative contribution of aerosol backscatter present at altitudes between 25 and 40 km prior to the launch of CALIPSO can be predicted from the 20-year SAGE II aerosol record (SPARC, 2006). Figure 4.1 presents the total (Rayleigh + aerosol) to Rayleigh backscatter ratio, R, as a function of altitude, latitude, and season. A lidar ratio or extinction-to-backscatter ratio of $S_a = 50$ sr^{-1} (Jäger et al., 1995; Jäger and Deshler, 2003) was used in the conversion of SAGE II aerosol extinction to backscatter. This 2-year period of observations (2001-2002) is characteristic of "background" levels, the last major eruption to have impacted the stratospheric aerosol layer being Mt. Pinatubo in 1991 (SPARC, 2006). Background aerosol conditions are still present, just prior to the launch of CALIPSO in spring 2006 since the stratosphere aerosol layer has remained unperturbed since 2002.

The overall structure of the aerosol distributions shown in Figure 4.1 reveals an aerosol reservoir in the tropical stratosphere and downward sloping contour surfaces with increasing latitude. Above 35 km, little aerosol is usually present (R-1 ~ 2%). However, over the tropics and in the default calibration height range (30-34 km), backscatter levels may be significant and vary in magnitude as the tropical aerosol reservoir ascends or descends. The formation of the tropical reservoir and its altitude variation is a product of aerosol microphyics and the interplay of the quasi-biennial oscillation and the annual circulation (Trepte and Hitchman, 1992).

The presence of aerosol within the calibration range requires the application of an aerosol backscatter correction to the calibration procedure. For the at-launch calibration procedure, an aerosol model is used based on a composite of the 2 years of observations displayed in Fig 4.1. During the mission, CALIPSO observations will be analyzed to determine if this aerosol model remains appropriate. In any case, the aerosol correction terms used in the calibration procedure will be routinely monitored to identify the impact of new eruptions and the effects of transport.

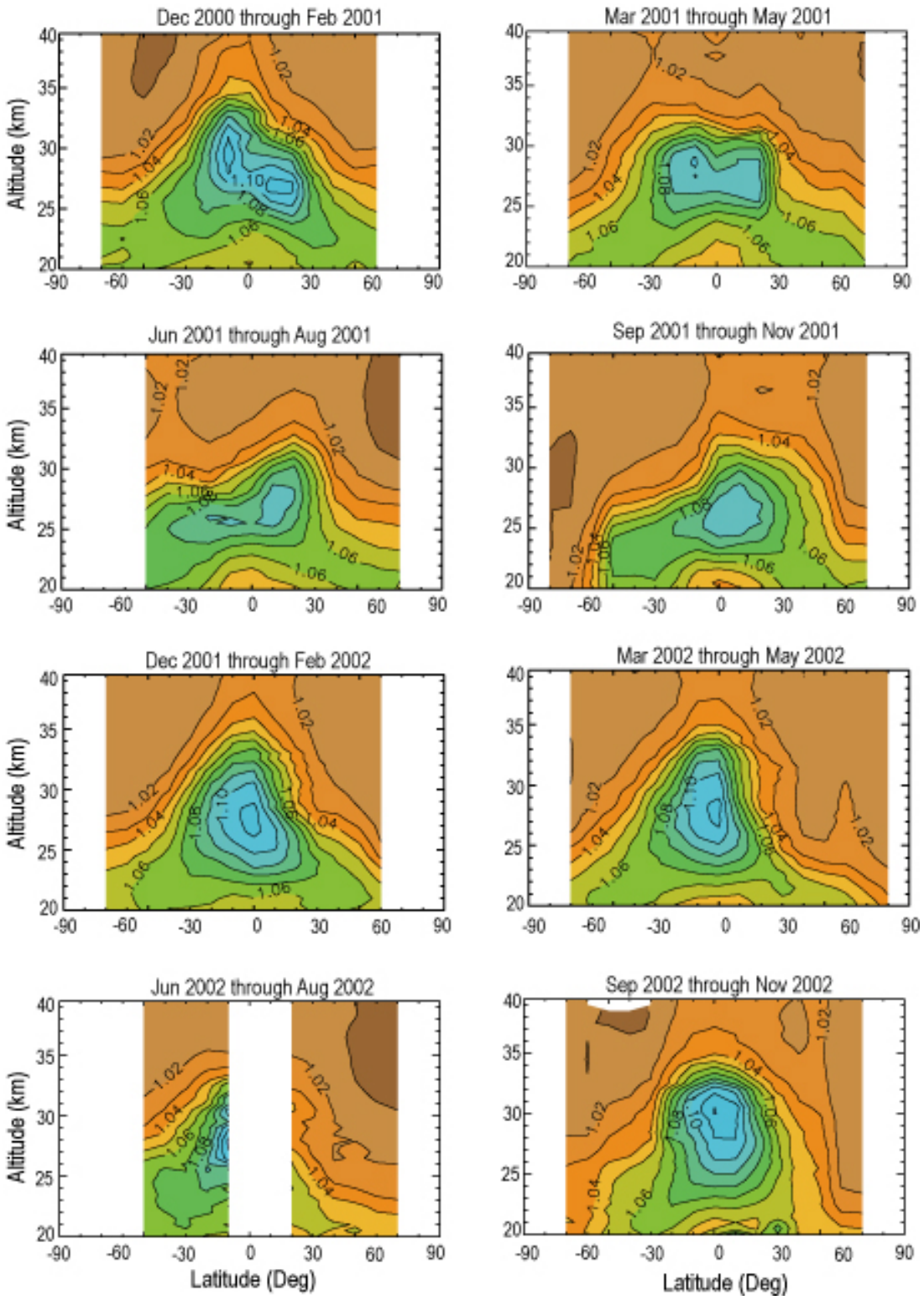


Figure 4.1 Aerosol total-to-Rayleigh backscatter ratio as a function of latitude, altitude and season converted from the SAGE II aerosol extinction observations at a wavelength of 525 nm using a lidar ratio of 50 sr^{-1} for a time period of 2001–2002 which is characteristic of "background" levels.

Due to background light and hence lower SNR under daytime lighting conditions, the 532 nm parallel channel calibration coefficient can only be estimated on the night side of the orbit. Estimates of the calibration coefficient computed on the night side of the orbit are stored in a calibration data product file, and the stored values are used to interpolate and/or extrapolate instantaneous values to be used at any point in the orbit. The calibration coefficient will change abruptly with commanded changes to detector gains or boresight operations. When such a commanded change occurs, only estimates of the calibration coefficient generated from data collected after the change are used to estimate the calibration coefficient for any time following that change.

4.1.2. Mathematical Basis

4.1.2.1. Basic Algorithm

The calibration coefficient is determined from the 532 nm parallel channel range-scaled, energy-normalized, gain-normalized profile,

$$X_{\|}(z) = C_{\|}\ \beta_{\|}(z)\, T^2(z)\,. \tag{4.1}$$

Rearranging to solve for $C_{\|}$ yields

$$C_{\|} = \frac{X_{\|}(z_C)}{\hat{\beta}_{\|}(z_C)\ \hat{T}^2(z_C)}\,. \tag{4.2}$$

In Eq.(4.2), $X_{\|}$ is measured by CALIOP and the other two parameters, $\hat{\beta}_{\|}$ and $\hat{T}^2$, are estimated from external data sources. As described earlier, these external data sources include the output from global meteorological analyses and measurements from other instruments (e.g., SAGE II observations). We have chosen to explicitly denote those parameters that are estimated from external data sources with the "hat" notation (ˆ).

We can break down the total parallel backscatter coefficient, $\beta_{\|}$, into the components due to molecular and aerosol backscatter separately

$$\beta_{\|} = \beta_{m,\|} + \beta_{a,\|} \tag{4.3}$$

where

$\beta_{m,\|}$ = parallel component of the molecular backscatter coefficient

$\beta_{a,\|}$ = parallel component of the aerosol backscatter coefficient.

The parallel component of molecular backscatter is calculated from an estimate of the total molecular backscatter and the expected depolarization ratio for molecular backscatter. The total molecular backscatter is the product of the number density and backscatter cross-section for air. The appropriate value for the backscatter cross section and depolarization ratio are, in part, functions of the bandwidth of the optical filter in the lidar receiver. Because the full-width-half-maximum bandwidth of the filter is of the order of 40 pm, only the central Cabannes line of the backscatter will be detected (She, 2001). The ratio of perpendicular to parallel backscatter for the Cabannes line is ~0.00366 (refer to Section 4.2; also see Cairo *et al.*, 1999), hence the

parallel component of the molecular backscatter is very nearly identical to the total molecular backscatter for CALIOP:

$$\beta_{m,\parallel}(z_C) = \frac{1}{1+\delta_m}\,\beta_m(z_C) = 0.996\,\beta_m(z_C)\,, \tag{4.4a}$$

where

$$\delta_m = \frac{\beta_{m,\perp}}{\beta_{m,\parallel}} = 0.00366 \tag{4.4b}$$

is the depolarization ratio for Cabannes scattering.

The computation of β_m is described in Section 4.2. To account for the aerosol backscattering in the calibration algorithm, the calibration coefficient is recast in terms of the total parallel scattering ratio, $R_\parallel$,

$$C_\parallel = \frac{X_\parallel(z_C)}{\hat{\beta}_{m,\parallel}(z_C)\;\hat{R}_\parallel(z_C)\;\hat{T}^2(z_C)} \tag{4.5}$$

where

$$R_\parallel(z_C) = \frac{\beta_{m,\parallel}(z_C) + \beta_{a,\parallel}(z_C)}{\beta_{m,\parallel}(z_C)} \tag{4.6}$$

As mentioned earlier, a global model of R for the background aerosol is used, currently based on SAGE II aerosol extinction data. Because aerosols in the calibration region are predominantly spherical, the perpendicular backscatter from the aerosols is insignificant, and the parallel aerosol backscatter is equivalent to the total aerosol backscatter.

Because of the relatively low SNR of the data samples in the calibration region, significant averaging is required to produce an accurate estimate of the calibration coefficient. Increased averaging is achieved in three steps: (1) horizontal averaging, (2) vertical averaging, and (3) application of smoothing filters. These steps occur as follows:

1. The profile data, $X_\parallel$, are averaged horizontally over 11 full-resolution profiles. At the calibration altitude, the full-resolution profiles correspond to a 5-km along track resolution (due to averaging done over 15 shots on board the satellite for data compression purposes), hence the 11-profile average corresponds to a 55-km along track average. Intermediate values of the calibration coefficient are then computed for each vertical sample in this horizontally averaged composite profile
2. The calibration coefficients computed at each altitude are averaged vertically over the entire calibration region, resulting in a single estimate for each 55 km orbit segment along the night side of the orbit.
3. These calibration coefficients are further averaged by smoothing the sequence of values via a 13-point running average filter, resulting in an effective 715 km average between independent samples.

Both the smoothed and unsmoothed calibration coefficients are stored in the Calibration Data Product. The smoothed calibration coefficients are used to interpolate/extrapolate calibration coefficients throughout the orbit and applied to produce attenuated backscatter coefficient profiles.

The algorithm used to compute the 532 nm parallel calibration coefficient can be summarized in the following two equations:

$$C_{\|}(y_k) = \underbrace{\frac{1}{j_{34} - j_{30} + 1} \sum_{j=j_{30}}^{j=j_{34}} \frac{\overbrace{\frac{1}{11} \sum_{i=11k-5}^{11k+5} X_{\|}(y_i, z_j)}^{\text{Step 1: average horizontally over 11 5-km profiles (one 55-km cell)}}}{\hat{\beta}_{m,\|}(y_k, z_j)\ \hat{R}_{\|}(y_k, z_j)\, \hat{T}^2(y_k, z_j)}}_{\text{Step 2: average intermediate calibration constants vertically from 30 to 34 km}} \tag{4.7}$$

$$\tilde{C}_{\|}(y_k) = \underbrace{\frac{1}{13} \sum_{k-6}^{k+6} C_{\|}(y_k)}_{\substack{\text{Step 3: compute running} \\ \text{average over 13 55-km cells} \\ \text{(715 km)}}} , \tag{4.8}$$

where

i = index for the i^{th} full-resolution profile (horizontal resolution = 5 km)

j = index for the j^{th} vertical sample in a high resolution profile (vertical resolution = 300 m)

j_{30} = vertical index corresponding to 30 km

j_{34} = vertical index corresponding to 34 km

k = index for the k^{th} calibration coefficient (computed every 55 km)

y = horizontal distance along the ground track

z = vertical distance

$C_{\|}(y_k)$ = unsmoothed calibration coefficient computed every 55 km along track

$\tilde{C}_{\|}(y_k)$ = smoothed calibration coefficient computed every 55 km along track.

Figure 4.2 schematically demonstrates the intermediate calibration coefficients $C_{\|}(y_k)$ that are computed every 55 km along the night side of the orbit in the 30-34 km region. We note that the indexing of i from $11k$ –5 to $11k$+5 in the horizontal average computed in step 1, as

$$\text{55-km horizontally averaged profile} = \frac{1}{11} \sum_{i=11k-5}^{11k+5} X_{\|}(y_i, z_j) , \tag{4.9}$$

accomplishes a decimation in horizontal index, from the 5-km index, i, to the 55-km index, k.

We also note that $\hat{\beta}_{m,\|}(y_k, z_j)$, $\hat{R}_{\|}(y_k, z_j)$, and $\hat{T}^2(y_k, z_j)$ are estimates of the corresponding parameters computed for the k^{th} 55-km along-track cell and j^{th} altitude bin. As mentioned earlier, these parameters are estimated from external data sources, which include the output from global meteorological analyses and measurements from other instruments. The external data include profiles of pressure, temperature, aerosol extinction and backscatter, and ozone concentration. The external data are interpolated along the ground track such that the sampling is identical in space and time to the CALIPSO full-resolution profiles. Horizontal averages of the estimated quantities are produced on the 55-km grid established for the CALIPSO calibration data product.

The estimate for the parallel component of molecular backscatter, $\hat{\beta}_{m,\|}(y_k, z_j)$, is calculated according to Eq.(4.4a), where the molecular volume scattering coefficient is computed from temperature and pressure profiles extracted from global meteorological analyses using Eq. (4.17) described in Section 4.2. The estimate for the parallel scattering ratio, $\hat{R}_{\|}(y_k, z_j)$, is calculated using Eq.(4.6), where the parallel component of aerosol backscatter, $\beta_{a,\|}(z)$, can be estimated using the global aerosol model derived from the SAGE II data and/or other data sources. The calculation of the two-way transmission involves integration over altitude, and merits a slightly more detailed description, which is provided in Section 4.1.2.2.

For diagnostic purposes, calibration coefficients are computed over four other altitude ranges in addition to the nominal 30-34 km range of the baseline calibration coefficient. These additional ranges are listed in Table 4.1. The vertical range varies with altitude to provide approximately similar SNRs for each average.

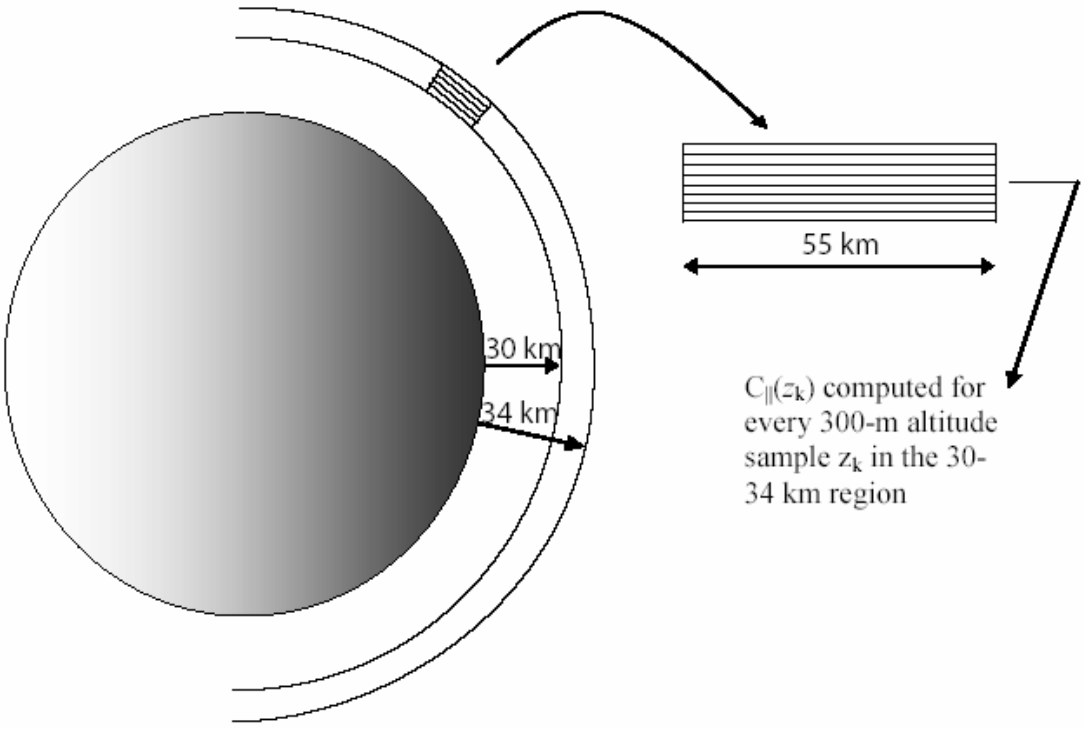


Figure 4.2 Intermediate calibration coefficients are computed every 55 km along the night side of the orbit in the 30-34 km region.

Table 4.1. Ranges for calculation of diagnostic and baseline calibration coefficients for the 532 nm parallel channel.

Range		Use
Lower Boundary	**Upper boundary**	
25.0 km	27.1 km	Diagnostic
26.0 km	28.5 km	Diagnostic
28.0 km	30.2 km	Diagnostic
30.3 km	34.2 km	Baseline 532 nm Parallel Calibration Range
32.0 km	39.0 km	Diagnostic

4.1.2.2. Calculation of the Two-Way Transmission, T^2

The two-way transmission to the calibration altitude, z_C, can be written as

$$T^2(z_C) = \exp\left\{-2\int_{z_C}^{z_{sat}} \left[\sigma_m(z') + \sigma_a(z') + \sigma_{O_3}(z')\right] dz'\right\}, \tag{4.10}$$

where

σ_m = extinction coefficient due to molecular scattering

σ_a = extinction coefficient due to aerosol scattering

σ_{O_3} = extinction coefficient due to ozone absorption

z_{sat} = altitude of the satellite

z_C = altitude at which the intermediate calibration coefficient is to be computed.

In practice, the integration extends from z_C to 40 km, as the atmospheric attenuation above 40 km is negligible. The algorithm used to compute the two-way transmission can be written as

$$\hat{T}^2(y_k, z_C) = \exp\left\{-2\sum_{j=j_C}^{j_{40}} \left[\hat{\sigma}_m(y_k, z_j) + \hat{\sigma}_a(y_k, z_j) + \hat{\sigma}_{O_3}(y_k, z_j)\right] \Delta z\right\}, \tag{4.11}$$

where we have again used the "hat" notation to identify parameters estimated using data from external sources and

Δz = vertical sample spacing of the extinction parameters determined from meteorological analyses

j = index for the j^{th} vertical sample of the estimated extinction parameters

j_C = vertical index corresponding to the altitude of the intermediate calibration coefficient computed in step 1

j_{40} = vertical index corresponding to 40 km

k = index for the k^{th} calibration coefficient (computed every 55 km horizontally along the orbit track)

4.2. Rayleigh Scattering

This section provides a brief introduction to atmospheric Rayleigh scattering and describes the calculations of Rayleigh scattering parameters for "standard" air for use in CALIOP data processing. Standard air is defined as dry air at a pressure of 760 mm Hg (1013.25 mb), temperature of 15°C, and containing 300 ppm CO_2 by volume. Atmospheric molecular scattering consists of Rayleigh scattering and vibrational Raman scattering. The scattering cross section for the latter is negligibly small compared with the former. Therefore, only Rayleigh scattering is considered when computing the molecular scattering (Bucholtz, 1995; Bodhaine *et al*., 1999; She, 2001). Rayleigh scattering consists of a central Cabannes line which has a Doppler-broadened full width of approximately 5 GHz (or 5 pm) and wavelength-shifted rotational Raman side lines with a much wider bandwidth of approximately ±200 cm^{-1} (±5 nm) at 0.5 μm wavelength (She, 2001). The total Rayleigh scattering cross section per molecule for the standard air is given by (see, for example, Bucholtz, 1995; Bodhaine *et al*., 1999)

$$
\begin{aligned}
Q_s(\lambda) &= \frac{24\pi^3\left(n_s^2(\lambda)-1\right)^2}{\lambda^4 N_s^2\left(n_s^2(\lambda)+2\right)^2}\left(\frac{3+6\delta_m(\lambda)}{3-4\delta_m(\lambda)}\right) = \frac{24\pi^3\left(n_s^2(\lambda)-1\right)^2}{\lambda^4 N_s^2\left(n_s^2(\lambda)+2\right)^2}F_k(\lambda), \\
F_k(\lambda) &= \frac{3+6\delta_m(\lambda)}{3-4\delta_m(\lambda)} \quad \text{or} \quad \delta_m(\lambda) = \frac{3F_k(\lambda)-3}{4F_k(\lambda)+6},
\end{aligned}
\tag{4.12}
$$

where δ_m is the depolarization ratio defined as the ratio of perpendicular to parallel components for linearly polarized light (She, 2001). n_s and N_s are the refractive index and molecular number density (2.54743×10^{19} cm^{-3}) of standard air, respectively. n_s is a function of wavelength and can be well estimated by a five-parameter formula (see, for example, Bucholtz, 1995; Bodhaine *et al*., 1999) developed based on laboratory experiments for a wavelength range of 0.185-1.69 μm. F_k is the King factor which accounts for the anisotropy of molecules. A spectral formula has also been developed, based on both theoretical and experimental works, that computes a weighted-average of the partial King factors for the main constituents of standard air, N_2 (78.084%), O_2 (20.946%), Ar (0.934%) and CO_2 (300 ppmv) (Bates, 1984; Bodhaine, *et al*., 1999). In the formula, the partial King factor is 1.15 for CO_2 and 1.00 for Ar, while the factor for N_2 and O_2 is wavelength dependent. We note that the amount of CO_2 in air is so low that the King factor is not sensitive to its variation. For example, adding 300 ppmv of CO_2 increases the King factor no more than 0.04% throughout the spectral region considered here. We also note that the definition of depolarization ratio used in Eq. (4.12) is different from that used by Bucholtz (1995) and Bodhaine *et al*. (1999). Those authors used a depolarization ratio defined as the ratio of horizontal to vertical components relative to the plane of incident light and scattering light at 90° for nonpolarized natural light (i.e., $\rho = \beta_{n,\parallel}(\pi/2)/\beta_{n,\perp}(\pi/2) = 2\delta/(1+\delta)$, usually called depolarization factor), whereas the depolarization ratio used in this work is defined

by Eq. (4.4b). This difference in depolarization ratio definition is reflected in corresponding differences in the formulation of the King factors.

The volume molecular scattering coefficient at a given altitude z can then be estimated using

$$\sigma_m(z,\lambda) = N(z)Q_s(\lambda), \text{ or} \tag{4.13a}$$

$$\sigma_m(z,\lambda) = \frac{N_A P(z)}{R_a T(z)} Q_s(\lambda) = \frac{C_s(\lambda)P(z)}{T(z)}, \tag{4.13b}$$

where P(z) and T(z) are the atmospheric pressure (hPa) and temperature (K) at altitude z, respectively, N(z) is the number density of molecules at z, N_A = 6.02214×10^{23} (1/mol) is Avogadro's number and R_a = 8.314472 (J/K/mol) is the gas constant. $C_s(\lambda) = N_A Q_s(\lambda)/R_a$ (K/hPa/m) can be computed using Eq. (4.12) with $n_s(\lambda)$ and $F_k(\lambda)$ data (Bodhaine, *et al.*, 1999) for the standard air. The volume molecular backscattering coefficient is estimated using,

$$\beta_m(z,\lambda) = \frac{\sigma_m(z,\lambda)}{S_m(\lambda)} = \frac{\sigma_m(z,\lambda)}{(8\pi/3)k_{bw}(\lambda)}, \tag{4.14}$$

where S_m=(8π/3)k_{bw} is the extinction to backscatter ratio or the lidar ratio for the molecular scattering and is usually approximated as 8π/3 in the lidar community (Collins and Russell, 1976) by neglecting the dispersion of the refractive index and King factor of air which is quantified by k_{bw} here. As mentioned earlier, Rayleigh scattering has a wide bandwidth. When a narrow band (< ~ 10 nm at 532 nm) optical filter is used, only a part of the Rayleigh scattering signal is detected. Two typical cases are given by

$$\begin{aligned} k_{bw}^{\mathrm{T}}(\lambda) &= \frac{1+2\delta_m^{\mathrm{T}}(\lambda)}{1+\delta_m^{\mathrm{T}}(\lambda)} && \text{fot total Rayleigh, and} \\ k_{bw}^{\mathrm{C}}(\lambda) &= \frac{1+2\delta_m^{\mathrm{T}}(\lambda)}{1-\frac{1}{6}\delta_m^{\mathrm{T}}(\lambda)} = \frac{1+6\delta_m^{\mathrm{C}}(\lambda)}{1+\frac{10}{3}\delta_m^{\mathrm{C}}(\lambda)} = \frac{6k_{bw}^{T}(\lambda)}{13-7k_{bw}^{T}(\lambda)} && \text{for Cabannes only,} \end{aligned} \tag{4.15}$$

with superscripts "C" and "T" representing the Cabannes and total Rayleigh scattering, and

$$\delta_m^T(\lambda) = \frac{4\delta_m^{\mathrm{C}}(\lambda)}{1+4\delta_m^{\mathrm{C}}(\lambda)} \text{ or } \delta_m^{\mathrm{C}} = \frac{\delta_m^T(\lambda)}{4-4\delta_m^T(\lambda)}. \tag{4.16}$$

Computed F_k, δ_m^T, δ_m^C, k_{bw}^T, and C_s as a function of wavelength are plotted in Figure 4.3. In general, large dispersion is shown for short wavelengths. Parameters computed for Nd:YAG laser's fundamental, second, third, and fourth harmonic oscillations, and for 550 nm wavelength which is commonly used as a reference (Collis and Russell, 1976), are listed in Table 4.2. It is seen that neglecting the atmospheric dispersion (i.e., setting k_{bw}=1) could causes an overestimate in the molecular volume backscatter coefficient of ~1.4% and ~3%, respectively, for the Rayleigh and Cabannes scattering at both CALIOP 532 nm and 1064 nm wavelengths.

We note that the computed Q_s at 550 nm is 4.51×10^{-27} (cm^2) which is ~1% smaller than that (4.56×10^{-27} cm^{-1}) in Collis and Russell (1976), which is used to compute the Rayleigh scattering cross section for other wavelengths using the approximated wavelength dependence of $Q_s(\lambda)$= 4.56×[λ(nm) / 550 nm]$^{-4}$ (×10^{-27} cm^2). As pointed by Young (1980), a major error source in

computing Rayleigh scattering cross section is that how the King factor, in particular its dependence on the wavelength, is taken into account. In Collis and Russell's computation which was based on Elterman (1968), the wavelength dependence of the King factor was neglected due to the limited available experiment data of the depolarization factor; i.e., a constant value of 1.061, corresponding to a King factor value near 260 nm of the computation result here, was used. Therefore, our computed values for the Rayleigh scattering parameters are used in the CALIOP data processing.

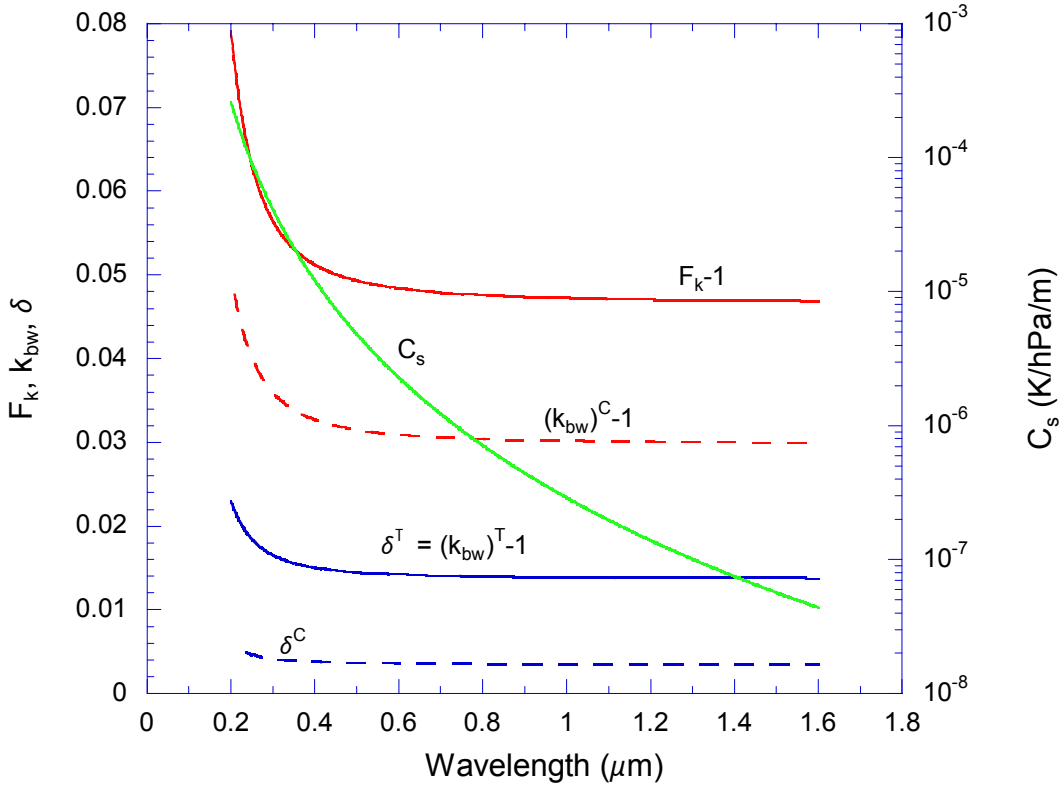


Figure 4.3 Rayleigh scattering parameters computed for standard air.

Table 4.2 Parameters computed for total Rayleigh scattering and central Cabannes line for Nd:YAG laser's fundamental, second, third, and fourth harmonic oscillation, and for 550 nm

λ (nm)	$(n_s - 1)$ ($\times 10^{-3}$)	F_k	δ_m (%)	δ_m^C (%)	k_{bw}	k_{bw}^C	C_s (K/hPa/m)	Q_s ($\times 10^{-27}$ cm^2)
266	0.2975	1.0604	1.768	0.4500	1.0174	1.0384	6.924e-05	95.59
355	0.2857	1.0529	1.554	0.3945	1.0153	1.0337	1.998e-05	27.59
532	0.2782	1.0490	1.441	0.3656	1.0142	1.0313	3.742e-06	5.167
550	0.2778	1.0488	1.436	0.3643	1.0142	1.0312	3.267e-06	4.510
1064	0.2740	1.0472	1.400	0.3523	1.0137	1.0302	2.265e-07	0.3127

The use of the λ^{-4} approximation, i.e., $Q_s(\lambda) = 4.56 \times [\lambda(\text{nm})/550]^{-4}$ ($\times 10^{-27}$ cm^2), could cause an error in Q_s of about -13%, -5%, -1% and 4%, respectively, at 266 nm, 355 nm, 532 nm and 1064 nm. Collis and Russell (1976) suggested the use of the approximation of $\lambda^{-4.09}$. We use

$$Q_s(\lambda) = 4.5102 \times 10^{-27} \left(\frac{\lambda(\text{nm})}{550} \right)^{-4.025 - 0.05627 \times (\lambda/550)^{-1.647}} \quad (\text{cm}^2) \tag{4.17}$$

and

$$C_s(\lambda) = \frac{N_A Q_s(\lambda)}{R_a} = 3.732 \times 10^{-6} \left(\frac{\lambda(\text{nm})}{550} \right)^{-4.025 - 0.05627 \times (\lambda/550)^{-1.647}} \quad (\text{K/hPa/m}) . \tag{4.18}$$

The accuracy of Q_s and C_s computed using Eqs. (4.17) and (4.18) is better than 1% through the spectral region (200-1600 nm) considered here and better than 0.2% for a spectral region of 350-1600 nm.

For CALIOP, because the bandwidth of the optical filter is very small (~ 40 pm or 1.4 cm^{-1} at 532 nm and ~ 0.5 nm or 4.4 cm^{-1} at 1064 nm), only the central Cabannes line is received. Therefore, for CALIOP the molecular volume scattering coefficient at 532 nm and 1064 nm is computed using Eq. (4.13a) with $Q_s(532) = 5.167 \times 10^{-27}$ (cm^2) and $Q_s(1064) = 3.127 \times 10^{-28}$ (cm^2), or using Eq. (4.13b) with $C_s(532) = 3.742 \times 10^{-6}$ (K/hPa/m) and $C_s(1064) = 2.265 \times 10^{-7}$ (K/hPa/m) as well as the ancillary pressure and temperature data. The volume backscatter coefficient is computed using Eq. (4.14) with $k_{bw}(532) = k_{bw}^C(532) = 1.0313$ and $k_{bw}(1064) = k_{bw}^C(1064) = 1.0302$, respectively, at 532 nm and 1064 nm.

4.3. Error Analysis

Uncertainties in the calibration coefficient are the result of both random and systematic errors:

$$\left(\frac{\Delta C_\parallel}{C_\parallel} \right)^2 = \left(\frac{\Delta C_\parallel}{C_\parallel} \right)^2_{RAN} + \left(\frac{\Delta C_\parallel}{C_\parallel} \right)^2_{SYS} \tag{4.19}$$

4.3.1. Systematic Error

The systematic component of error is given by

$$\left(\frac{\Delta C_\parallel}{C_\parallel} \right)^2_{SYS} \approx \left(\frac{\Delta \hat{R}_\parallel(z_c)}{\hat{R}_\parallel(z_c)} \right)^2 + \left(\frac{\Delta \hat{\beta}_{m,\parallel}(z_c)}{\hat{\beta}_{m,\parallel}(z_c)} \right)^2 + \left(\frac{\Delta \hat{T}_{532}(z_c)}{\hat{T}_{532}(z_c)} \right)^2 + (\varepsilon + a + c)^2 . \tag{4.20}$$

The systematic uncertainties are comprised by the error in the parameters estimated largely from the GMAO meteorological analysis product and global estimates of stratospheric aerosol loading. The last term, $(\varepsilon + a + c)^2$, concerns errors induced by non-ideal optical performance of the lidar system (see Section 6.2.1). This error term can be effectively reduced along with other corrections to the profile products described in Section 6.2.1, and hence will be removed from the calculation of systematic error for the calibration coefficient. For reporting systematic error, Eq.(4.21) will be used.

$$(\Delta C_\parallel)^2_{SYS} \approx \left(\left(\frac{\Delta \hat{R}_\parallel(z_c)}{\hat{R}_\parallel(z_c)} \right)^2 + \left(\frac{\Delta \hat{\beta}_{m,\parallel}(z_c)}{\hat{\beta}_{m,\parallel}(z_c)} \right)^2 + \left(\frac{\Delta \hat{T}_{532}(z_c)}{\hat{T}_{532}(z_c)} \right)^2 \right) C_\parallel^2 \tag{4.21}$$

Assuming uncertainties that should reasonably apply in the stratosphere for 532 nm in the calibration altitude between 30 and 34 km, the preliminary estimate for the systematic error is broken out below.

Table 4.3. Preliminary estimates of systematic error in the calculation of the 532 nm parallel calibration coefficient.

$\left(\frac{\Delta C_{\parallel}}{C_{\parallel}}\right)_{SYS}$	$\frac{\Delta \hat{R}_{\parallel}(z_c)}{\hat{R}_{\parallel}(z_c)}$	$\frac{\Delta\hat{\beta}_{m,\parallel}(z_c)}{\hat{\beta}_{m,\parallel}(z_c)}$	$\frac{\Delta \hat{T}^2_{532}(z_c)}{\hat{T}^2_{532}(z_c)}$	$(\varepsilon + a + c)$
0.05	0.04	0.03	0.005	0.0

(Reagan *et al.*, 2002)

Estimates of these error terms will be improved as more knowledge is gained on the accuracy of the products used to compute them. We note that events may occur that will drastically affect the estimate of the systematic uncertainties. For instance, a volcanic eruption may drastically affect the error in both $\hat{R}_{\parallel}(z_c)$ and $\hat{T}^2_{532}(z_c)$. In such cases, estimates for these error terms will have to be updated based on expert analysis of all available data.

4.3.2. Random Error

The random error is dominated by the noise in the lidar data itself. The random error is estimated using two different approaches: (1) from the variability of the intermediate calibration coefficients that are averaged to produce the final calibration coefficient estimate and (2) using a parameter we call the Noise Scale Factor (NSF; see section 8). Two quantities are therefore computed to represent the random error. The first quantity is computed by Eq. (4.22)

$$\left(\Delta C_{\parallel}(y_k)\right)_{RAN} = \frac{\mathrm{Stdev}\left(C_{\parallel}(y_i)\right)}{\sqrt{11}} = \frac{1}{11}\left(\sum_{i=11k-5}^{11k+5}\left[C_{\parallel}(y_i) - C_{\parallel}(y_k)\right]^2\right)^{1/2} \tag{4.22}$$

where

$$C_{\parallel}(y_i) = \frac{1}{(j_{34} - j_{30} + 1)}\sum_{j=j_{30}}^{j_{34}} C_{\parallel}(y_i, z_j) \tag{4.23}$$

and $C_{\parallel}(y_k)$ is the average value of calibration coefficients computed from all data points in the calibration altitude range over a 55 km horizontal distance as defined by Eq. (4.7). $\mathrm{Stdev}\left(C_{\parallel}(y_i)\right)$ is the standard deviation of $C_{\parallel}(y_i)$. The quantity defined by Eq. (4.22) is therefore a standard deviation equivalent to that for the averaged $C_{\parallel}(y_k)$ over 11 intermediate $C_{\parallel}(y_i)$ estimates. Each sample is assumed to be statistically independent so the standard deviation is reduced by a factor of $\sqrt{11}$ due to averaging. We refer to this quantity as equivalent standard deviation.

The second quantity representing the random uncertainty in the averaged $C_{532,\parallel}(y_k)$ is estimated using NSF and is given by

$$\Delta C_{\|}(y_k) = \frac{f_{\|}(N_{range})}{11\times(j_{upper}-j_{lower}+1)}\left\{\sum_{j=j_{lower}}^{j_{upper}}\frac{\sum_{i=11k-5}^{11k+5}\left[\frac{r_j^2}{E_{532}}NSF_{\|}^2(y_i)X_{\|}(y_i,z_j)+\left(\frac{r_j^2\Delta P_{bd,\|}(y_i)}{E_{532}G_{A,\|}}\right)^2\right]}{\left[\beta_{m,\|}(y_k,z_j)R_{\|}(y_k,z_j)T^2(y_k,z_j)\right]^2}\right\}^{1/2} \quad (4.24)$$

Where $\Delta P_{\mathrm{bd},\|}$ is the standard deviation of background signal, in science digitizer counts, i.e., the baseline RMS noise (see Eq.(8.9) for its measurement). $f_{\|}(N_{\mathrm{range}})$ is a factor used to correct the correlation between neighbour samples (see Eq. (8.15) or (8.16)) which is due to the limited bandwidth of the preamplifier. The value of $f_{\|}(N_{\mathrm{range}})$ at any point is a function of the number of range bins averaged on-board the satellite (i.e., prior to downlinking the data). A look-up table was constructed for this function, with values derived from autocorrelation coefficients. The details are discussed in Section 8.

Note that the NSF used in Eq. (4.24) is dependent on both range-resolution and data averaging. The fundamental value for the NSF is computed at a range resolution of 15 meters and for single shots (i.e., some quantities used for estimating the NSF are derived as part of the on-board processing and hence are computed at a 15-m resolution, even though the highest vertical resolution of the downlinked profile information is 30 m). Thus if the downlinked range resolution is ΔR and the data has been averaged for N_{shot} shots, the NSF must be converted from the 15-meter, single shot NSF (refer to Section 8.3), using

$$NSF(N_{bin},N_{shot}) = \frac{NSF}{\sqrt{N_{bin}N_{shot}}} = \frac{NSF}{\sqrt{\frac{\Delta R}{15}N_{shot}}} \quad (4.25)$$

The background noise ΔN_{bd} must also be converted using

$$\Delta P_{bd}(N_{bin},N_{shot}) = \frac{\overline{\Delta P}_{bd}}{\sqrt{N_{bin}N_{shot}}} = \frac{\overline{\Delta P}_{bd}}{\sqrt{\frac{\Delta R}{15}N_{shot}}} \quad (4.26)$$

where $\overline{\Delta P}_{bd}$ is the background RMS noise averaged over N_{shot}. Values computed using Eqs.(4.22) and (4.24) should be approximately equal since the change of calibration coefficient over the 55-km horizontal average range is expected to be very small (i.e., the natural variation of the atmosphere is insignificant).

4.4. Simulation Results

The calibration procedure was tested using synthetic data generated by the CALIOP simulator (Powell *et al.*, 2002; Powell, 2005). Figure 4.4 presents an example. Shown in the figure are the 532 nm parallel calibration coefficients for the nighttime portion of an orbit computed for each 55-km profile using Eq.(4.7) and further smoothed over 13 55-km profiles using Eq.(4.8). The 55-km calibration coefficients are saved as intermediate values and the smoothed values are used in the lidar data calibration. The relative RMS error of simulated smoothed values is 3.5%.

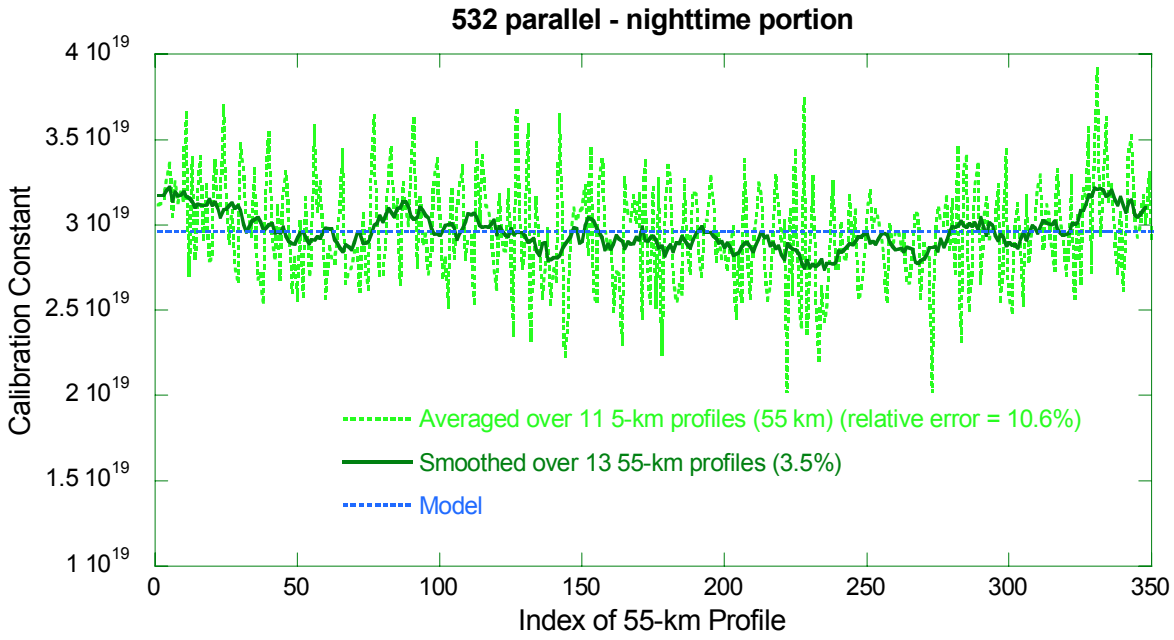


Figure 4.4 Simulation results of 532 nm parallel channel calibration using simulated CALIOP signals.

4.5. Anomalous Conditions and On-Orbit Checkout

Anomalous conditions that may affect the way we interpolate/extrapolate the calibration coefficient include the following:

- Volcanic eruption
- High temporal variability of calibration coefficient
- Low SNR (e.g., laser energy or etalon throughput low)

Data from the on-orbit checkout period will be used to evaluate the stability of the calibration coefficient. This information may affect the way we interpolate/extrapolate the calibration coefficient to the day side of the orbit. Issues to be investigated include the following:

- Stability of boresight
- Stability of laser-etalon tuning

5. 532 Perpendicular Channel Calibration

5.1. Algorithm Description

5.1.1. General Description

Calibration of the 532 perpendicular channel is transferred from the 532 parallel channel using data acquired during the Polarization Gain Ratio (PGR) operation. It is anticipated that Polarization Gain Ratio Operations will be performed approximately once per week throughout the mission. The required frequency of PGR operations will be determined during the early phases of the on-orbit mission. During the calibration, a spatial pseudo-depolarizer is inserted into the 532 nm optical path upstream of the polarization beam splitter (Figure 5.1). Insertion of this device results in a randomly polarized backscatter signal, and thus nominally equal optical power is directed into the detectors for the two orthogonal polarization orientations, regardless of the target being measured. Inserting the pseudo-depolarizer allows the relative sensitivity of the two 532 nm receiver channels to be determined. The ratio of the two detection channel signals is called the Polarization Gain Ratio (PGR) and is denoted in the equations that follow by the symbol K_P. The PGR accounts for differences in the responsivity and gain of the two detection channels and the relative transmission of the optics downstream of the polarizing beam splitter (PBS). To compute the 532 nm perpendicular attenuated backscatter profiles, other effects must also be taken into account, including differences in optical transmission of the receiver channels upstream of the PBS and optical cross talk between the receiver channels. These effects are treated separately in Section 6.2.1.

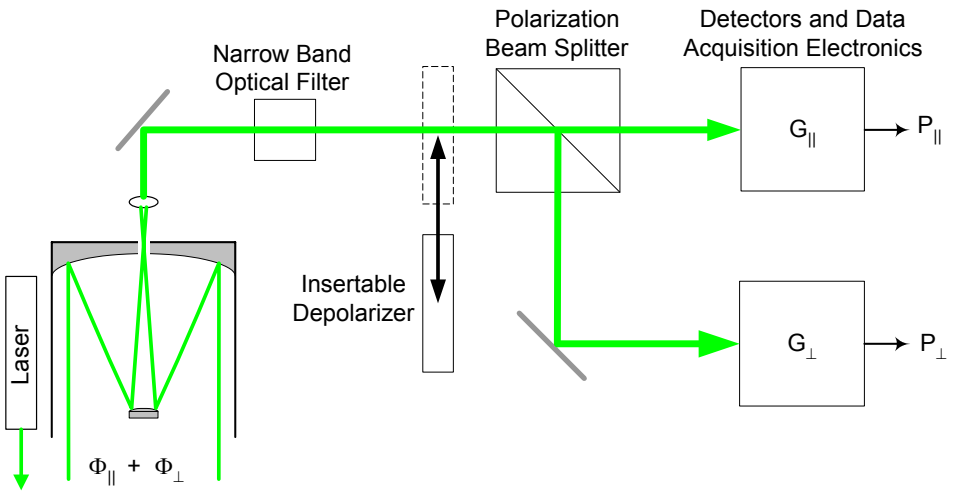


Figure 5.1 Conceptual block diagram of the 532 receiver channels. During the PGR Operation, a pseudo-depolarizer is inserted into the optical path upstream of the PBS, putting nominally equal optical fluxes into the receiver optics downstream of the PBS.

After the on-orbit check out phase of the mission, PGR Operations will nominally be carried out only on the night side of the orbit. Data will be acquired in PGR mode for approximately 5 minutes, or as required, depending upon instrument performance. The acquired data will be averaged horizontally over the entire PGR acquisition segment (~2100 km) and vertically from 18 km to 25 km (the Baseline PGR Computation Region). Averaging horizontally and vertically

will insure that a relative random error of <1% can be achieved in the relative calibration of the two channels. Using this altitude range insures that strong cloud signals will be excluded from the calculation to avoid errors in the calculation of the polarization gain ratio due to differences in the transient response of the two channels. Also, this altitude region ensures against using data for which the backscatter signal has been completely extinguished, as would be the case for deep convective cloud systems. The vertical range over which the averaging is done is programmable and can be changed after results from the on-orbit checkout become available.

As a diagnostic, the PGR calculation will also be done over other vertical ranges as defined in Table 5.1.

Table 5.1 Altitude ranges for the PGR calculation diagnose

Range		Use
Lower Boundary	**Upper boundary**	
Surface + 1 km	6 km	Diagnostic
6 km	12 km	Diagnostic
12 km	18 km	Diagnostic
18 km	25 km	Baseline PGR Computation Range

5.1.2. Mathematical Basis

For an ideal instrument, the measured 532 parallel and perpendicular channel signals can be written as

$$X_{\parallel}(z) = C_{\parallel} \ \beta_{\parallel}(z)\, T^2(z) \tag{5.1}$$

$$X_{\perp}(z) = C_{\perp} \ \beta_{\perp}(z)\, T^2(z) \tag{5.2}$$

We can write the calibration coefficients in terms of the optical throughputs upstream and downstream of the depolarizer and the residual component containing all other factors. The expanded expressions are

$$C_{\parallel} = T_{U,\parallel}\, T_{D,\parallel}\, G_{\parallel}\, C_0 \tag{5.3}$$

$$C_{\perp} = T_{U,\perp}\, T_{D,\perp}\, G_{\perp}\, C_0 , \tag{5.4}$$

where

$T_{U,\parallel}$ = parallel channel optical throughput upstream of the depolarizer

$T_{D,\parallel}$ = parallel channel optical throughput downstream of the depolarizer

$T_{U,\perp}$ = perpendicular channel optical throughput upstream of the depolarizer

$T_{D,\perp}$ = perpendicular channel optical throughput downstream of the depolarizer

$G_{\parallel}$ = overall responsivity and gain of the parallel channel

$G_{\perp}$ = overall responsivity and gain of the perpendicular channel

C_0 = all remaining factors in the calibration coefficients that are common to both channels (includes telescope area, laser-receiver optical overlap, etc.)

We define the PGR, K_P, as follows.

$$K_P \equiv \frac{C_{\perp}}{C_{\parallel}} = \frac{T_{U,\perp}\, T_{D,\perp}\, G_{\perp}\, C_0}{T_{U,\parallel}\, T_{D,\parallel}\, G_{\parallel}\, C_0} = \left(1-\alpha_U\right)\left(1-\alpha_D\right)\frac{G_{\perp}}{G_{\parallel}}, \tag{5.5}$$

where

$1-\alpha_i$ = ratio of the optical throughput of the perpendicular channel to that of the parallel channel, where the subscript i is either "U" or "D", corresponding to the portion of the receiver upstream or downstream of the depolarizer, respectively.

We note that α_U and α_D will be numbers very much smaller than one, hence

$$K_P \equiv \frac{C_{\perp}}{C_{\parallel}} \approx \frac{G_{\perp}}{G_{\parallel}}. \tag{5.6}$$

As mentioned earlier, K_P is estimated from signals acquired during the PGR operation, when the depolarizer is inserted into the optical path of the system so that nearly equal optical fluxes are incident on the two detectors. K_P can be estimated from the ratio of the signals measured during the PGR Operation as follows:

$$\begin{aligned}
\hat{K}_P &= \frac{X_{\perp}^{PGR}}{X_{\square}^{PGR}} \\
&= \frac{T_{D,\perp}\, G_{\perp}\, C_0 \left[\dfrac{T_{U,\parallel}\beta_{\parallel}(z)+T_{U,\perp}\beta_{\perp}(z)}{2}\right] T^2(z)}{T_{D,\parallel}\, G_{\parallel}\, C_0 C_{\parallel} \left[\dfrac{T_{U,\parallel}\beta_{\parallel}(z)+T_{U,\perp}\beta_{\perp}(z)}{2}\right] T^2(z)} \\
&= \left(1-\alpha_D\right)\frac{G_{\perp}}{G_{\parallel}} \\
&= \frac{1}{1-\alpha_U} K_P \,.
\end{aligned} \tag{5.7}$$

We note that $\hat{K}_P$ is not a perfect estimate of K_P, and should be scaled by $(1-\alpha_U)$ to correct for the upstream throughput ratio. This is not the only correction that is required, however. Several non-ideal polarization effects in the transmitter and receiver create deviations in the measured signals from those that would be measured by an ideal instrument. In addition to the relative

throughput differences already noted, non-ideal effects include the polarization impurity of the transmitter, angular offset between the polarization reference axes of the transmitter and receiver, optical cross talk between polarizations in the receiver, and non-ideal depolarizer performance (unequal split of the optical signals). Corrections for all of these effects are carried out when the calibration coefficients are applied to the measured profiles to produce the attenuated backscatter profile data products. These corrections are described in Section 6.2.1.

In practice K_P is estimated using signals averaged horizontally and vertically. The actual algorithm for estimating K_P implements the following equation:

$$\hat{K}_P \equiv \frac{\overline{X}_{\perp}^{PGR}}{\overline{X}_{||}^{PGR}} = \text{ estimate for } \hat{K}_P \text{ produced by algorithm} \tag{5.8}$$

and

$$\overline{X}_{\perp}^{PGR} = \frac{1}{N\left(j_{high} - j_{low} + 1\right)} \sum_{k=1}^{N} \sum_{j=j_{low}}^{j=j_{high}} X_{\perp}^{PGR}(y_k, z_j) \tag{5.9}$$

$$\overline{X}_{||}^{PGR} = \frac{1}{N\left(j_{high} - j_{low} + 1\right)} \sum_{k=1}^{N} \sum_{j=j_{low}}^{j=j_{high}} X_{||}^{PGR}(y_k, z_j) \ , \tag{5.10}$$

where N = the total number of profiles acquired under the PGR Calibration Mode

5.2. Error Analysis

Estimates of the uncertainties in $C_{\perp}$ and $\hat{K}_P$ are calculated and stored in the Calibration Data Product File. Uncertainties in the Polarization Gain Ratio, $\hat{K}_P$, are due to both random and systematic errors

$$\left(\frac{\Delta K_p}{K_p}\right)^2 = \left(\frac{\Delta K_p}{K_p}\right)_{RAN}^2 + \left(\frac{\Delta K_p}{K_p}\right)_{SYS}^2 \tag{5.11}$$

The systematic error is due primarily to non-ideal polarization effects in the transmitter and receiver (e.g., non-ideal performance of the depolarizer, relative differences in throughput upstream of the depolarizer, etc.). These effects also create errors in the 532 nm parallel calibration coefficient as discussed in Section 4.3.1. All the known non-ideal instrumental polarization errors are corrected when the final attenuated backscatter coefficients are computed at a later stage of the processing and will not be further described here. However, the systematic error is initially set to zero as default and will be updated after launch when sufficient information is available.

As for the 532 nm calibration coefficient, two quantities, the equivalent standard deviation and an uncertainty based on the NSF, are computed to estimate the random error of the PGR. The equivalent standard deviation in $\hat{K}_P$ is given by

$$\left(\Delta K_p\right)_{RAN} = \frac{1}{N}\left\{\sum_{i=1}^{N}\left[K_p(y_i) - \overline{K}_p(y_i)\right]^2\right\}^{0.5} \ , \tag{5.12}$$

where

$$K_p(y_i) = \frac{\sum_{j=j_{low}}^{j_{high}} X_{\perp}^{PGR}(y_i, z_j)}{\sum_{j=j_{low}}^{j_{high}} X_{\parallel}^{PGR}(y_i, z_j)} \tag{5.13}$$

$$\bar{K}_p = \frac{1}{N}\sum_{i=1}^{N} K_p(y_i), \tag{5.14}$$

and N is the number of profiles acquired during the PGR operation.

The uncertainty estimate using the NSF is given by

$$\left(\Delta K_p\right)_{RAN} = \bar{K}_p \left\{ \frac{f_{\parallel}^2(N_{range}) \sum_{k=1}^{N} \sum_{j=j_{low}}^{j=j_{high}} \left[\frac{r_j^2}{E_{532}} NSF_{\parallel}^2(y_k) X_{\parallel}^{PGR}(y_k, z_j) + \left(\frac{r_j^2 \Delta P_{bd,\parallel}(y_i)}{E_{532} G_{A,\parallel}} \right)^2 \right]}{\left[\sum_{k=1}^{N} \sum_{j=j_{low}}^{j=j_{high}} X_{\parallel}^{PGR}(y_k, z_j) \right]^2} \right.$$
$$\left. + \frac{f_{\perp}^2(N_{range}) \sum_{k=1}^{N} \sum_{j=j_{low}}^{j=j_{high}} \left[\frac{r_j^2}{E_{532}} NSF_{\perp}^2(y_k) X_{\perp}^{PGR}(y_k, z_j) + \left(\frac{r_j^2 \Delta P_{bd,\perp}(y_i)}{E_{532} G_{A,\perp}} \right)^2 \right]}{\left[\sum_{k=1}^{N} \sum_{j=j_{low}}^{j=j_{high}} X_{\perp}^{PGR}(y_k, z_j) \right]^2} \right\}^{0.5} \tag{5.15}$$

Note again, the NSF provided by the Level 1 processing must be converted using Eq. (4.17), based on the range resolution of the data and number of shots averaged. The background noise must also be converted using Eq.(4.18). The estimated relative random error for the perpendicular calibration coefficient, $C_{\perp}$, is

$$\left(\frac{\Delta C_{\perp}}{C_{\perp}}\right)^2 = \left(\frac{\Delta K_p}{K_p}\right)^2 + \left(\frac{\Delta C_{\parallel}}{C_{\parallel}}\right)^2 . \tag{5.16}$$

Hence, the estimated random error in $C_{\perp}$ archived in the calibration file is

$$\Delta C_{\perp} = C_{\perp}\left[\left(\frac{\Delta K_p}{K_p}\right)^2 + \left(\frac{\Delta C_{\parallel}}{C_{\parallel}}\right)^2\right]^{\frac{1}{2}} . \tag{5.17}$$

However, neither the 532 nm perpendicular calibration coefficient nor the uncertainties will be included explicitly in the lidar data products.

5.2.1. Anomalous Conditions and On-orbit Checkout

Conditions that may lead to corrections in the PGR algorithms and/or operational procedures include:

- High temporal variability of the PGR calibration coefficient, $\hat{K}_P$
- Systematic variation of the $\hat{K}_P$ over an orbit that can be estimated and corrected.

During the on-orbit checkout period, any systematic error in the PGR due to nonlinearities discovered in either of the detection channels will be analyzed and appropriate algorithms created to compute and archive the systematic uncertainty associated with the nonlinearities. As an example, systematic variation in the PGR with respect to calibration target signal level may be an indication of nonlinearity in one or both of the polarization channels.

Also, during on-orbit checkout, the PGR operation should be conducted at several points in the day and night side of the orbit to determine the stability of the PGR and any effects that should be considered in the interpolation of PGR values between calibration operations. These tests will also enable estimation of optimal averaging schemes for daytime PGR operations, should these be considered in the future, and whether there are systematic variations due to differences in detector gain ratio due to lighting conditions or instrument temperature. The PGR validation, as described in Section 5.2.2, will also provide means for these onboard checkouts.

5.2.2. Validation of the On-board PGR Procedure

This section discusses the use of solar background radiation signals scattered from ice clouds to validate the onboard PGR determination. The background light signal measured by lidar during daytime (i.e., the Background Monitor reading for CALIOP) is the sunlight signal scattered by the land and ocean surfaces as well as by clouds, aerosols, and molecules in the atmosphere. Because the background signal scattered by ice clouds is largely unpolarized, the difference in backscatter intensity between the parallel and perpendicular channels ought therefore to be minimal. In addition, multiple scattering can further depolarize the background light scattered from optically dense ice clouds.

Figure 5.2 presents an example of ice cloud measurements made by the Cloud Physics Lidar (CPL) (McGill et al., 2002). CPL is a three wavelength, polarization-sensitive airborne lidar that provides down-looking measurement of the atmosphere from an altitude of ~20 km.

The upper panel of Figure 5.2 shows a time history of 532 nm attenuated backscatter profiles acquired on February 22, 2003 during a flight over the Pacific Ocean. For much of the flight, a totally attenuating cirrus cloud is observed between 10-14 km. The aircraft made four passes over this cirrus layer. This layer consists mostly of ice crystals as identified from the depolarization measurement (mostly larger than 20% as shown in the second panel from the top). The third panel shows the ratio of perpendicular to parallel components of the solar background signal at 1064 nm. These data are derived by averaging the subsurface measurements (100 samples) of each profile where there is no laser backscattering signal. The ratio of these signals remains almost constant throughout the entire extent of the cirrus layer, except at the edges. The mean value of the ratio is consistent with the PGR value (solid line in the third panel) determined for this flight via the CPL's half-wave plate calibration technique (McGill et al., 2002), which is similar in some respects to the CALIOP onboard PGR procedure. This demonstrates that the scattered solar radiation from the densest parts of the cirrus layer is unpolarized, regardless of

solar elevation angle, and can therefore be used to validate the polarization gain ratio of the two polarization channels. Multiple scattering may have also contributed an additional measure of depolarization. Deviations are seen at the edges of the cirrus cloud where the cloud layer is transmissive and the polarization is affected by the lower water clouds and/or ocean surface. This is as expected since the scattering of solar radiation from both water clouds and the ocean surface results in significant polarization, and the measured ratios depend on the solar elevation angle as predicted by theory (Liu *et al.*, 2004).

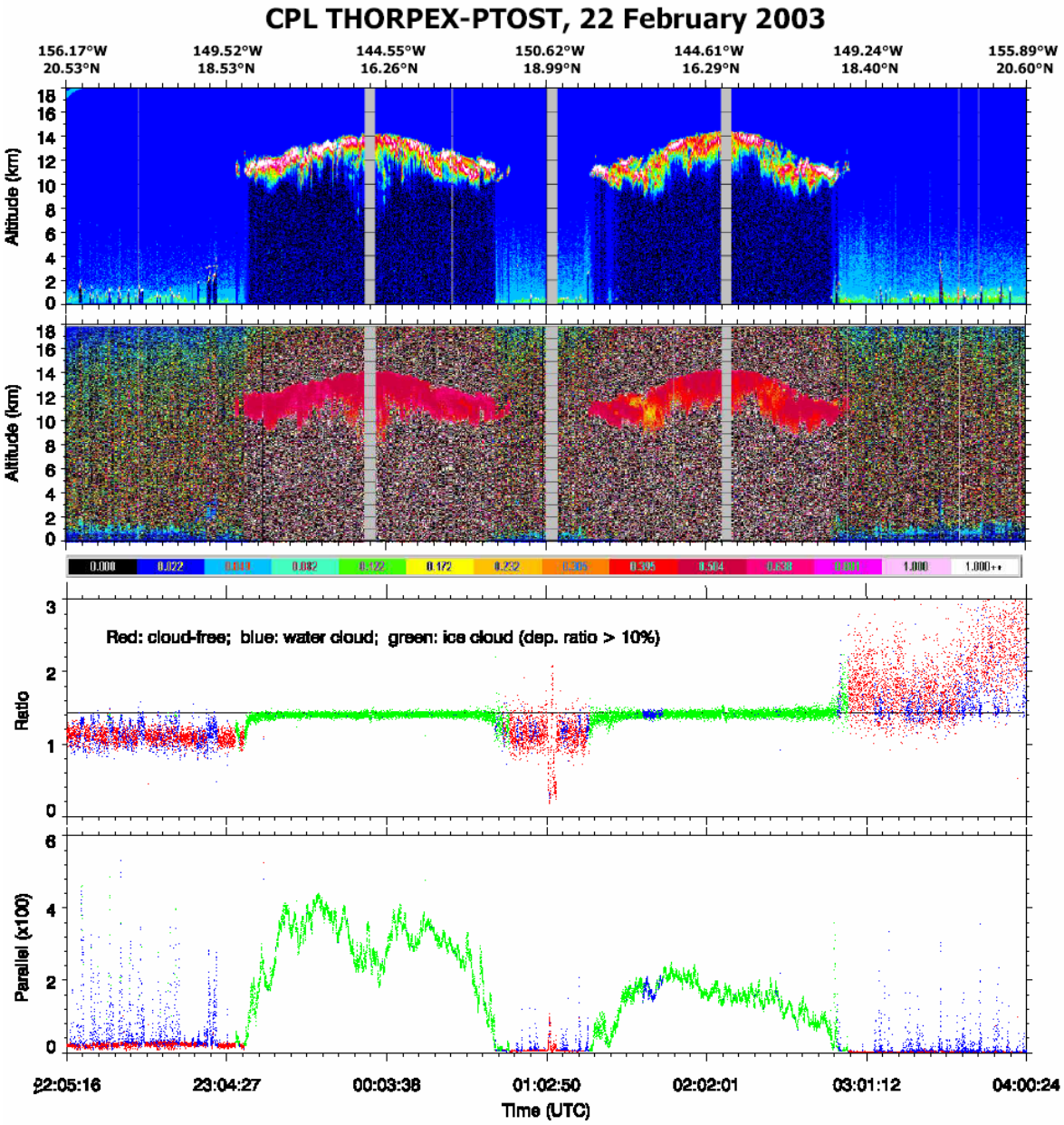


Figure 5.2 Attenuated backscatter at 532 nm (upper panel), depolarization ratio at 1064 nm along with color bar (second panel from top), perpendicular-to-parallel component ratio (third panel from top), and parallel component (lower panel) of background signal (mostly scattered solar radiation), observed on 22 February 2003 by CPL.

Using CALIOP measurements similar to those described above, we will validate the CALIOP depolarization calibration using two different approaches. Approach 1 is a linear-fit method. Figure 5.3(a) presents scatter plots of perpendicular versus parallel components for all background solar signals of Figure 5.2. Figure 5.3(b) is for cirrus only. The background signal is identified as being scattered from an ice cloud by checking the depolarization ratio of any feature found in the corresponding lidar profile. If the layer-integrated depolarization ratio of the feature is larger than 20%, then the feature is classified as an ice cloud and the background signal is consequently identified as being scattered from ice cloud. When a linear fit is applied to these points, the slope of the fitted line yields the PGR of the two channels. A large spread of data points is seen in Figure 5.3(a), while a substantially better correlation is seen when fitting the ice-cloud-only data points as shown in Figure 5.3(b). The PGR value determined by the ice-cloud-only fitting (1.41) is very close to the value determined by the half wave-plate method (1.44). Results suggest that the ice-cloud-only calibration technique can then be used as a diagnostic method to check the stability of the CALIPSO gain ratio calibration. The onboard method will be applied only periodically, and it requires the insertion of additional onboard optics. The depolarization ratio is used to select the appropriate (ice cloud) data for the gain ratio calibration.

An alternative approach is to select a high, dense cirrus anvil similar to the one shown in Figure 5.2. The ratio of perpendicular to parallel components of the solar background signal is plotted as in the third panel of Figure 5.2. The PGR can then be derived from the mean value of the flattest part of the curve. The flatness of the curve can be used as a metric to select an appropriate ice cloud for analysis.

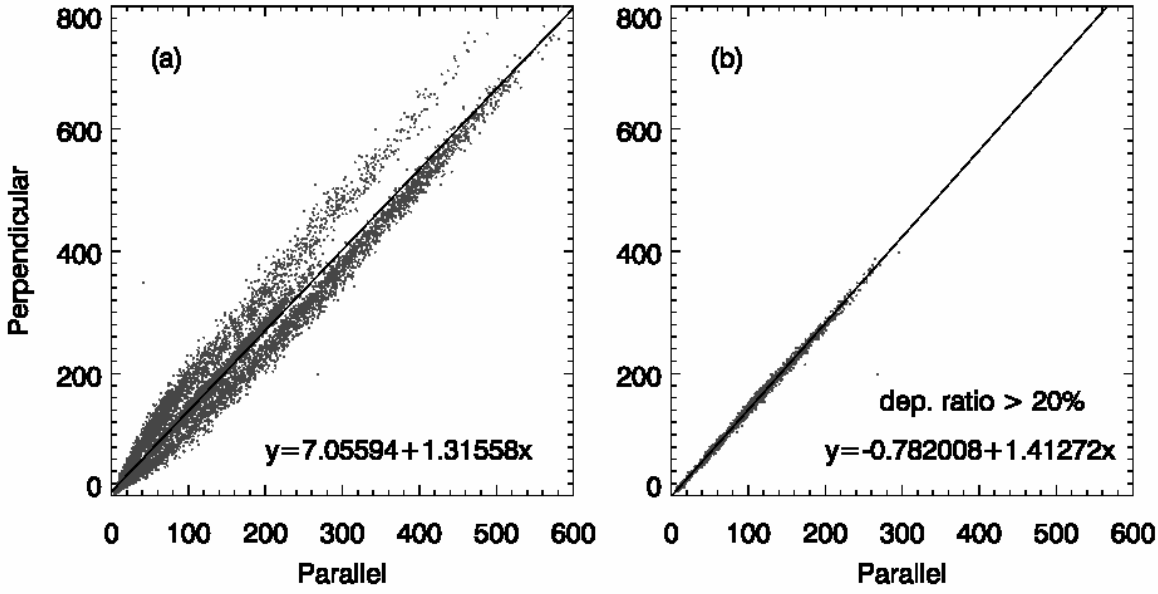


Figure 5.3 Scatter plots of background signals of perpendicular versus parallel components for (a) all cases and (b) for the ice-cloud only cases with linearly fitted line, for the flight of 19 February 2003.

6. Constructing Composite Profiles of 532 nm Total Backscatter Coefficients

6.1. General Description

The main products from the Level 1 algorithms are 532 nm and 1064 nm attenuated backscatter profiles constructed from the three instrument channels. The attenuated backscatter coefficient is the product of the volume backscatter coefficient and the two-way optical transmission from the lidar to the sample volume in question:

$$\beta'(z) = \beta(z)\ Exp\left[-2\int_0^{r(z)} \sigma(r')\,dr'\right] = \beta(z)\ T^2(z)\,. \tag{6.1}$$

Retrieval of the backscatter and extinction coefficients from the attenuated backscatter coefficient is one of the objectives of the Level 2 algorithms.

In this section we discuss how the calibration coefficients are applied to the data to produce the attenuated backscatter profiles. The application to the 1064 nm data is straightforward. However, the 532 nm data requires additional corrections for non-ideal instrument effects. As these corrections, and the error analysis involved, are fairly complicated, only the results are presented here. More details on the derivations can be found in the reference cited below.

6.2. Algorithm Description: 532 Channels

For an ideal instrument, attenuated backscatter profiles would be calculated exactly as the 1064 nm profiles are calculated; that is

$$\beta'_{\parallel,532}(k,z) = \frac{1}{C_{\parallel,532}(k)}\,X_{\parallel,532}(k,z) \tag{6.2}$$

$$\beta'_{\perp,532}(k,z) = \frac{1}{K_P(k)\ C_{\parallel,532}(k)}\,X_{\perp,532}(k,z)\,. \tag{6.3}$$

However, several non-ideal instrument effects associated with the polarization-sensitive measurement at 532 nm corrupt the measured signals. These effects include:

- polarization impurity in the transmitter,
- misalignment between the transmitter and receiver polarization reference axes
- cross talk of optical power in one receiver polarization reference frame to the other
- differences in the throughput between polarizations in the receiver

Section 6.2.1 defines the minimum set of parameters required to quantify these instrument effects. Section 6.2.2 shows how these parameters and the calibration coefficients described earlier in the document are combined to compute the attenuated backscatter profiles.

6.2.1. Definition of Non-ideal Instrument Optical Parameters

The instrument parameters used in the computation of 532 nm attenuated backscatter coefficient profiles are defined below. See the document, "Errors in Backscatter and Depolarization Due to Non-Ideal Optical Effects in the CALIPSO Lidar", for a more complete definition of each.

These are parameters thought to be the minimum set necessary to quantify optical instrument effects. The effect of any one parameter can be further allocated to individual optical elements in the transmitter and/or receiver. Initial values for these parameters were measured during the integration and test phase of the instrument development. For some of the parameters, estimates will be updated using data collected during on-orbit operations. All of the parameters are defined so as to represent a perturbation term such that their values should be much less than 1.

ε = ratio of the energy output from the transmitter that is polarized perpendicular to the transmitter's nominal polarization axis to the total energy output from the transmitter. This parameter concerns the light output from the beam expander. The perpendicularly polarized component is assumed to be incoherent with respect to the parallel component.

a (or c) = parallel-to-perpendicular polarization cross talk parameter for the portion of the optical train that is upstream (or downstream) of the depolarizer. This parameter represents the fraction of the optical power polarized parallel to the receiver polarization reference plane that is transferred to the perpendicular channel. The model separates cross talk that occurs upstream of the insertion point of the depolarizer (parameter a) from that that occurs downstream of the depolarizer (parameter b).

b (or d) = perpendicular-to-parallel polarization cross talk parameter for the portion of the optical train that is upstream (or downstream) of the depolarizer. (See comments on parallel-to-perpendicular cross talk parameter above.)

$1-\alpha_U$ = ratio of the optical throughput of the perpendicular channel to that of the parallel channel for the portion of the optical train upstream of the depolarizer. The value of this parameter may be positive or negative.

$\frac{\Delta\phi}{\phi}$ = parameterization for non-ideal depolarizer performance.

$= \frac{\phi_{D,\parallel} - \phi_{D,\perp}}{\phi}$, where

ϕ = optical power incident on the depolarizer

$\phi_{D,\parallel}$ = optical power in the parallel channel downstream of the depolarizer

$\phi_{D,\perp}$ = optical power in the perpendicular channel downstream of the depolarizer

Note 1: the upstream cross talk parameters, a and b, quantify the effects of both (1) misalignment of the transmitter and receiver polarization reference planes and (2) polarization cross talk in the receiver upstream of the depolarizer. These effects are easily separated as outlined in the document describing the model used to generate the specifications: "Errors in Backscatter and Depolarization Due to Non-Ideal Optical Effects in the CALIPSO Lidar."

Note 2: The parameterization for non-ideal depolarizer performance does not provide all the information that might be needed. This parameter is introduced to capture the error in the calibration of the perpendicular channel relative to the parallel channel.

6.2.2. Calculation of 532 Attenuated Backscatter Coefficients

Estimates of the attenuated backscatter coefficients for the 532 nm channels are calculated according to the following equations:

$$\hat{\beta}'_{\parallel}(k,z) = \frac{1}{\hat{C}_{\parallel}(k)} \frac{\left(1-a-b-2c-\alpha_U+\frac{\Delta\phi}{\phi}\right) X_{\parallel}(k,z) - \frac{(b+d+\varepsilon)}{\hat{K}_P} X_{\perp}(k,z)}{1-a-b-2c-\alpha_U+\frac{\Delta\phi}{\phi}} \tag{6.4}$$

$$\hat{\beta}'_{\perp}(k,z) = \frac{1}{\hat{C}_{\parallel}(k)} \frac{\left(1-2a-c-d-\frac{\Delta\phi}{\phi}\right) \frac{1}{\hat{K}_P} X_{\perp}(k,z) - (a+c+\varepsilon) X_{\parallel}(k,z)}{1-a-b-2c-\alpha_U+\frac{\Delta\phi}{\phi}} \tag{6.5}$$

$$\hat{\beta}'_{total}(k,z) = \hat{\beta}'_{\parallel}(k,z) + \hat{\beta}'_{\perp}(k,z). \tag{6.6}$$

Of these, only $\hat{\beta}'_{total}(k,z)$ and $\hat{\beta}'_{\perp}(k,z)$ are reported in the CALIPSO data products. ($\hat{\beta}'_{\parallel}(k,z)$ can, of course, be derived as necessary by simple subtraction of the reported quantities.)

7. 1064 nm Channel Calibration

7.1. Algorithm Description

7.1.1. General Description

As described in Section 4, calibration of the 532 nm parallel channel will be accomplished via normalization to a high altitude region where the backscatter is nearly entirely from molecular returns. The molecular backscatter is too weak to permit use of this calibration technique for the longer wavelength 1064 nm channel. Instead, the calibration at 532 nm will be transferred to the 1064 nm channel via comparison of the returns from cirrus clouds, where the spectral dependence of backscatter at the two wavelengths is expected to be fairly stable. Cirrus clouds are good targets for this purpose because they occur with sufficient frequency and provide strong and nearly spectrally flat backscatter and extinction. Another consideration favorable to the use of cirrus clouds for calibration transfer is that, because cirrus clouds occur at high altitudes, corrections for the spectral transmission differences between satellite and the cloud top are relatively small and fairly predictable for the two wavelengths. In fact, given the very low aerosol loadings currently typical of the stratosphere and upper troposphere, spectral transmission differences due to non-molecular constituents can be neglected entirely for the CALIOP wavelengths.

Theoretical and experimental studies have indicated that extinction is essentially independent of wavelength for large particles such as ice crystals in cirrus; the extinction cross-section approaches two times the geometric area as particle size increases to sizes typical of cirrus cloud particles (see, for example, van de Hulst, 1981, p107). Theoretical considerations suggest that lidar backscatter varies weakly with wavelength even in the geometric limit, due to variation in the refractive index of ice (Bohren and Huffman, 1983). A limited number of measurements have indicated that cirrus cloud backscatter has little wavelength dependence (Ansmann, *et al.*, 1993; Beyerle, *et al.*, 2001) although the uncertainty in some of these measurements is significant, and significant wavelength dependence appears to exist for some types of cirrus. The wavelength dependence of cirrus backscatter will be further investigated after the launch of CALIPSO, and the calibration algorithm will be refined to take any new information into account.

The 532 nm and 1064 nm cloud signals measured by CALIOP contain a molecular/aerosol component in addition to the stronger cloud component. This molecular/aerosol backscatter is rarely spectrally flat. One algorithm approach is to remove the non-cloud component. This is difficult due to the low SNR of the non-cloud backscatter and the requirement for the lidar signal inversion, and thus does not lend itself to an easily-implemented, operational algorithm. The algorithm approach selected will use only very strong cloud returns where the influence of non-cloud scattering is small and the error introduced by it is insignificant. This algorithm has the added advantage that it can be used under both daytime and nighttime lighting conditions.

7.1.2. Mathematical Basis

7.1.2.1. Basic Algorithm

The normalized lidar signal in clouds is given by

$$X_\lambda(r) = C_\lambda \left[\beta_{\lambda,c}(r) + \beta_{\lambda,m}(r) \right] T_{\lambda,m}{}^2(0,r)\, T_{\lambda,O_3}{}^2(0,r)\, T_{\lambda,a}{}^2(0,r)\, T^2_{\lambda,c}\left(0, r_{top}\right) T_{\lambda,c}{}^2\left(r_{top}, r\right) \quad (7.1)$$

where

C_λ = lidar calibration coefficient at wavelength λ

$T^2_{\lambda,n}(0,r)$ = two-way transmittance for constituent n from the lidar (at $r = 0$) to range r; atmospheric constituents considered include molecules (m), ozone (O_3), background aerosol (a), and clouds (c)

$T^2_{\lambda,c}\left(r_{top}, r\right)$ = two-way transmittance of cloud particles from cloud top (r_{top}) to range r

$\beta_{\lambda,n}(r)$ = volume backscatter coefficient for constituent n at range r

We note that in the most general case, an aerosol backscatter would also be included. However, in the cirrus cloud calibration regions, the aerosol backscatter is usually quite small when compared with the molecular backscatter (typically, $R_a < 1.04$), and will be essentially negligible when compared to the much larger cloud backscatter term. Correcting the signal for the known (from meteorological data and models) attenuations due to molecules, ozone, and background aerosols yields

$$\begin{aligned} \tilde{X}_\lambda(r) &= \frac{X_\lambda(r)}{T_{\lambda,m}{}^2(0,r)\, T_{\lambda,a}{}^2(0,r)\, T_{\lambda,O_3}{}^2(0,r)} \\ &= C_\lambda \left(\beta_{\lambda,m}(r) + \beta_{\lambda,c}(r) \right) T^2_{\lambda,c}\left(0, r_{top}\right) T_{\lambda,c}{}^2\left(r_{top}, r\right) \end{aligned} \quad . \quad (7.2)$$

Consistent with the backscatter assumptions stated in Section 7.1.1, we now define a cloud backscatter color ratio,

$$\chi_c = \beta_{1064,c}(r) \Big/ \beta_{532,c}(r), \quad (7.3)$$

and a molecular backscatter color ratio

$$\chi_m = \beta_{1064,m}(r) \Big/ \beta_{532,m}(r). \quad (7.4)$$

The λ^{-4} approximation of wavelength dependence of Rayleigh scattering allows us to immediately evaluate the molecular backscatter color ratio, such that $\chi_m \approx 1/16$. The appropriate value for the cloud backscatter color ratio is less well known, but is generally assumed to be in the near neighbor-hood of 1.0.

Deriving the 1064 nm calibration constant involves taking the ratio of the signals acquired at each of the two wavelengths. To do so requires that we first obtain the total backscatter signal for the 532 nm channel. This calculation is straightforward:

$$X_{532}(r) = X_{532,\parallel}(r) + \frac{X_{532,\perp}(r)}{K_p}. \quad (7.5)$$

The ratio of the signals at 1064 nm and 532 nm can now be written as

$$\frac{\tilde{X}_{1064}(r)}{\tilde{X}_{532}(r)} = \frac{C_{1064}\left(\beta_{1064,c}(r)+\beta_{1064,m}(r)\right)T^2_{1064,c}\left(0,r_{top}\right)T_{1064,c}{}^2\left(r_{top},r\right)}{C_{532}\left(\beta_{532,c}(r)+\beta_{532,m}(r)\right)T^2_{532,c}\left(0,r_{top}\right)T_{532,c}{}^2\left(r_{top},r\right)}$$
$$= \frac{C_{1064}}{C_{532}}\left(\frac{\chi_c\beta_{532,c}(r)+\chi_m\beta_{532,m}(r)}{\beta_{532,c}(r)+\beta_{532,m}(r)}\right)\left(\frac{T^2_{1064,c}\left(0,r_{top}\right)}{T^2_{532,c}\left(0,r_{top}\right)}\right)\left(\frac{T_{1064,c}{}^2\left(r_{top},r\right)}{T_{532,c}{}^2\left(r_{top},r\right)}\right). \quad (7.6)$$

Because the spectral transmission differences due to non-molecular constituents are assumed to be negligible, and because the extinction coefficients for cirrus particles are assumed to be spectrally independent, the two-way transmittance terms in Eq. 7.6 vanish (that is, $T^2_{1064,c}\left(0,r_{top}\right)\big/T^2_{532,c}\left(0,r_{top}\right)\approx 1$ and $T^2_{1064,c}\left(r_{top},r\right)\big/T^2_{532,c}\left(r_{top},r\right)\approx 1$). Under these assumptions, the 1064 nm calibration equation becomes

$$\frac{\tilde{X}_{1064}(r)}{\tilde{X}_{532}(r)} \approx \frac{C_{1064}}{C_{532}}\left(\frac{\chi_c\beta_{532,c}(r)+\chi_m\beta_{532,m}(r)}{\beta_{532,c}(r)+\beta_{532,m}(r)}\right) = \frac{C_{1064}}{C_{532}}\left(\frac{\chi_c\tilde{R}_{532}(r)+\chi_m}{\tilde{R}_{532}(r)+1}\right) \quad (7.7)$$

where

$$\tilde{R}_{532}(r) = \frac{\beta_{532,c}(r)}{\beta_{532,m}(r)} \quad (7.8)$$

and is related to the previously defined scattering ratio (i.e., Eq. 4.6) by

$$\tilde{R}_{532}(r) = R_{532}(r) - 1\,. \quad (7.9)$$

The corresponding estimate of the 1064 nm calibration constant is

$$C_{1064} \approx C_{532}\left(\frac{\tilde{X}_{1064}(r)}{\tilde{X}_{532}(r)}\right)\left(\frac{\tilde{R}_{532}(r)+1}{\chi_c\tilde{R}_{532}(r)+\chi_m}\right) \quad (7.10)$$

Substituting 7.9 into the far right-hand expression in 7.10 produces

$$\frac{\left(R_{532}(r)-1\right)+1}{\chi_c\left(R_{532}(r)-1\right)+\chi_m} = \frac{R_{532}(r)}{\chi_c R_{532}(r)-\chi_c+\chi_m}, \quad (7.11)$$

and this expression converges to χ_c^{-1} as $R_{532}(r)\to\infty$. Thus in the limit,

$$C_{1064} \approx \chi_c^{-1} C_{532}\left(\frac{\tilde{X}_{1064}(r)}{\tilde{X}_{532}(r)}\right). \quad (7.12)$$

This equation (i.e., 7.12) is the one used to compute the CALIOP calibration constant at 1064 nm.

When using equation (7.12) to estimate the 1064 nm calibration constant, we are neglecting those contributions to the signals made by molecular backscatter. However, assuming the value of χ_c is reasonably well known, the errors incurred by neglecting the molecular component of the backscatter are small (less that 2%). Furthermore, as illustrated by Figure 7.1, these errors largely independent of the actual value of χ_c. Figure 7.1 plots the relative errors in the 1064 nm

calibration coefficient as a function of the 532 nm scattering ratio for those cases when molecular scattering is ignored for both wavelengths. Calculations for two different cloud backscatter color ratios are shown: χ_c = 1.0, plotted using a solid line; and χ_c = 0.85, plotted using symbols (empty circles). For 532 nm scattering ratios larger than 50, the relative error is seen to be less than 2% for both cases. This indicates that if the threshold value is properly chosen, the effect of molecular scattering is very small.

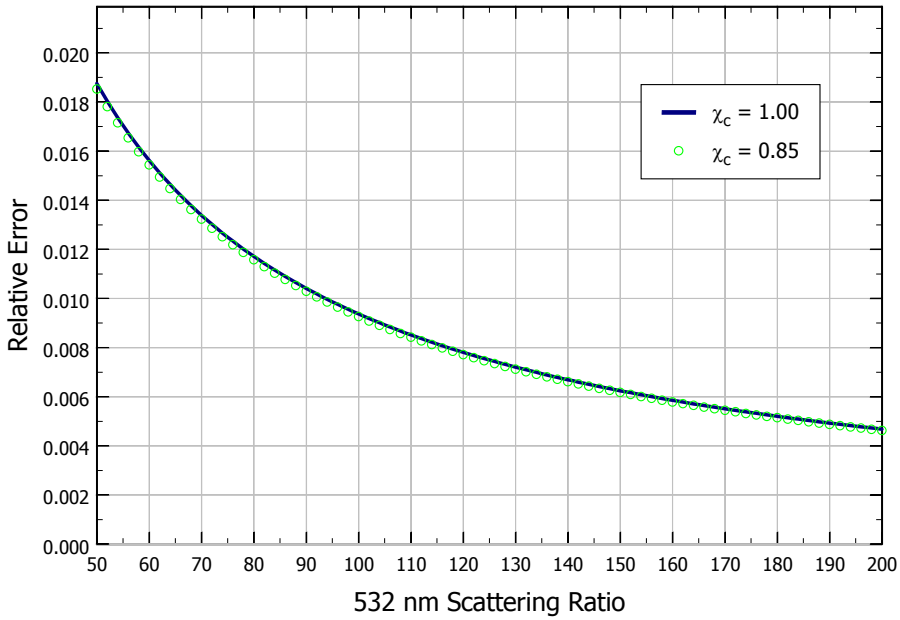


Figure 7.1 Relative error in 1064 calibration coefficient vs. 532 scattering ratio when molecular scattering is ignored.

7.1.2.2. Practical Implementation

In the CALIPSO calibration processing timeline, the 532 nm signal is calibrated prior to performing the 1064 nm calibration. Because of this, the calibrated total attenuated backscatter (i.e., parallel plus perpendicular backscatter) is used instead of the energy-normalized range-corrected 532 nm signals, and Eq. (7.12) becomes

$$C_{1064} \approx \chi_c^{-1} \left(\frac{\tilde{X}_{1064}(r)}{\tilde{\beta}'_{532}(r)} \right). \tag{7.13}$$

In this expression, $\tilde{\beta}'_{532}$ is the total attenuated backscatter at 532 nm (i.e., as computed using Eq. 6.6) that has been corrected for the attenuations due to molecules and ozone; that is

$$\tilde{\beta}'_{532}(r) = \frac{\beta'_{532}(r)}{T^2_{m,532}(r) T^2_{O_3,532}(r)}. \tag{7.14}$$

To locate cirrus clouds suitable for performing the 1064 nm calibration, a threshold signal for identifying strong cloud returns is determined by computing the 532 nm attenuated backscatter that is equivalent to a predetermined threshold scattering ratio, R_t. For the initial data

processing we have set R_t = 50; however, this limit is implemented as a runtime parameter, and hence can be changed during the mission, as circumstances dictate. The attenuated backscatter threshold is computed as below. Since (from Eq. 4.6)

$$R_t = \frac{\beta_{c,t}(z) + \beta_m(z)}{\beta_m(z)} \tag{7.15}$$

it follows that

$$\beta_{c,t}(z) + \beta_m(z) = R_t\ \beta_m(z) \tag{7.16}$$

and the threshold signal becomes

$$\beta'_{532,t}(z) = R_t\ \beta_{m,532}(z) T^2_{m,532}(z) T^2_{O_3,532}(z). \tag{7.17}$$

Note that the quantities $\beta_{m,532}(z)$, $T^2_{m,532}(z)$, and $T^2_{O_3,532}(z)$ are all derived from the available meteorological information.

Prior to beginning the search for clouds, the 532 nm attenuated backscatter and 1064 nm normalized signal profiles are averaged horizontally over an integer number of major frames to increase SNR. A single major frame (i.e., a 5-km horizontal average) is used for nighttime calibration, and two major frames are averaged for the day side of the orbit. The search for cloud returns is restricted to the altitude region from 17-km down to 8.2-km. This restriction helps eliminate non-cirrus cloud returns, and facilitates the modeling of the transmission terms in the above equations. To be suitable for use in the calibration scheme, a cloud must contain a minimum of three consecutive points that all exceed the attenuated backscatter threshold. If there are multiple regions within a profile that meet this criterion, only the highest is used.

Operationally, the 1064 nm calibration algorithm proceeds as follows:

1) Average the 532 nm attenuated backscatter coefficients and the normalized 1064 nm signal profiles horizontally over the specified number of major frames. Compute the corresponding molecular density array to be used for molecular backscatter and extinction coefficients, and ozone density array to be used for estimating ozone extinction coefficients.
2) Compute signal thresholds for 532 nm equivalent to a scattering ratio of R_t.
3) Search for strong cloud returns between 8.2 and 17 km using thresholds calculated in Step 2.
4) Determine the highest altitude cloud segment with at least 3 consecutive signals above threshold. (The bottom of the cloud segment occurs the first time a signal does not exceed threshold.)
5) Correct the data within the cloud for the attenuations due to molecules and ozone (i.e., calculate $\tilde{\beta}'_{532}(r)$ and $\tilde{X}_{1064}(r)$).
6) Calculate intermediate C_{1064} values using Eq. (7.13) at each altitude in the cloud segment. Calculate the mean of these constants. Report the mean C_{1064} estimate for the profile

along with time, location, peak cloud scattering ratio, altitude of the peak, cloud depth, and boresight/HV change information in calibration data product file.

7) Steps 1 – 5 are repeated for the nighttime and daytime portion of each orbit. The mean, standard deviation, uncertainty using NSF, and number of samples of C_{1064} for each ½ orbit are calculated and reported in the calibration data product file.

8) Perform outlier rejection using mean and standard deviation. The outlier rejection procedure computes the mean, $\overline{C}_{1064}$, and the standard deviation, ΔC_{1064}, of C_{1064} for ½ orbit (i.e., a granule), and then removes those 1064 nm calibration coefficient estimates for which $\left|C_{1064} - \overline{C}_{1064}\right| > k_{th}\Delta C_{1064}$, where k_{th} is the threshold (k_{th} is equal to 2 currently and will be updated after the launch). We note that, however, for this procedure to work properly, an underlying distribution of cloud color ratios should probably be symmetric. This issue will be checked onboard. A flag is reported in the calibration data product file to indicate whether an intermediate C_{1064} estimate computed from a particular profile will be incorporated into the average used to estimate the 1064 nm calibration coefficient.

9) Calibrate 1064 nm data using granule-averaged calibration coefficients using outlier rejection stored in calibration data product file.

7.2. Error Analysis

The uncertainty in the 1064 nm calibration coefficient consists of both random and systematic errors.

$$\left(\frac{\Delta C_{1064}}{C_{1064}}\right)^2 = \left(\frac{\Delta C_{1064}}{C_{1064}}\right)_{RAN}^2 + \left(\frac{\Delta C_{1064}}{C_{1064}}\right)_{SYS}^2 \tag{7.18}$$

7.2.1. Systematic Error

The systematic component of error is given by

$$\left(\frac{\Delta C_{1064}}{C_{1064}}\right)_{SYS}^2 = \left(\frac{\Delta C_{532,\|}}{C_{532,\|}}\right)^2 + \left(\frac{\delta_c}{1+\delta_c}\right)^2 \left(\frac{\Delta K_p}{K_p}\right)^2 + \left(\frac{\Delta T_{1064}(r_{ct})}{T_{1064}(r_{ct})}\right)^2 + \left(\frac{\Delta T_{532}(r_{ct})}{T_{532}(r_{ct})}\right)^2 + \left(\frac{\Delta \chi'}{\chi'}\right)^2, \tag{7.19}$$

Where δ_c is the depolarization ratio of the selected cloud layer. For estimating the systematic error, δ_c is assumed to be 0.5 for cirrus clouds. $\Delta\chi'$ represents the potential error due to the possible systematic wavelength dependence of attenuated backscatter color ratio $\beta'_{1064} / \beta'_{532}$. While the current version of the algorithm assumes that $\beta_{1064} = \beta_{532}$ for cirrus particles, this assumption requires further investigation. Measurements exist for color ratios with significant departures from unity for some types of cirrus cloud (Beyerle, 2001). Determining the cirrus types having the most consistent color ratio values will be the subject of study through the first year of the mission. Data may have to be reprocessed to refine 1064 nm calibration coefficients should the outlier rejection technique fail to produce adequate screening of clouds with color ratios that significantly differ from unity.

A preliminary estimate for the systematic error is broken out in Table 7.1.

Estimates of the error terms in Table 7.1 will be improved as more knowledge is gained on the accuracy of the products used to compute them. The statistics of the wavelength dependence of χ' in high cirrus clouds will be investigated when more data is acquired after launch, though an estimated of $\Delta\chi'/\chi'$ of 0.04 has been temporarily made. We note that events may occur that will drastically affect the estimate of the systematic uncertainties. For instance, a volcanic eruption may drastically affect the error in both $\hat{R}_{\|}(z_c)$ and $\hat{T}^2_{532}(z_c)$. In such cases, estimates for these error terms will have to be updated based on expert analysis of all available data.

Table 7.1. Preliminary estimates of systematic error in the calculation of the 1064 nm calibration coefficient.

$\left(\frac{\Delta C_{1064}}{C_{1064}}\right)_{SYS}$	$\left(\frac{\Delta C_{532,\|}}{C_{532,\|}}\right)$	$\left(\frac{\delta_c}{1+\delta_c}\right)\left(\frac{\Delta K_p}{K_p}\right)$	$\left(\frac{\Delta T_{1064}(r_{ct})}{T_{1064}(r_{ct})}\right)$	$\left(\frac{\Delta T_{532}(r_{ct})}{T_{532}(r_{ct})}\right)$	$\left(\frac{\Delta\chi'}{\chi'}\right)$
0.07	0.05	0.003	.002	.02	0.04

7.2.2. Random Error

The random error in the 1064 calibration coefficient is dominated by detection noise. The random error in the 1064 calibration coefficient for each averaged profile (over a number of major frames) is estimated as follows:

$$\left(\Delta C_{1064}(y_i)\right)_{RAN} \approx \frac{C_{1064}(y_i)}{\left(j_{high}-j_{low}+1\right)}\left\{\sum_{j=j_{low}}^{j_{high}}\left[\frac{\left(\frac{r_j^2 NSF_{1064}^2 X_{c,1064}(y_i,z_j)}{E_{1064}}+\frac{\left(r_j^2\Delta P_{bd,1064}(y_i)\right)^2}{E_{1064}^2 G_{A,1064}^2}\right) f_{1064}{}^2(N_{range})}{\left(X_{c,1064}(y_i,z_j)\right)^2}\right.\right.$$
$$\left.\left.\frac{\left(\frac{r_j^2 NSF_{\|}^2 X_{c,\|}(y_i,z_j)}{E_{532}}+\frac{\left(r_j^2\Delta P_{bd,\|}(y_i)\right)^2}{E_{532}^2 G_{A,\|}^2}\right) f_{\|}^2(N_{range})+\left(\frac{r_j^2 NSF_{\perp}^2 X_{c,\perp}(y_i,z_j)}{K_p^2 E_{532}}+\frac{\left(r_j^2\Delta P_{bd,\perp}(y_i)\right)^2}{K_p^2 E_{532}^2 G_{A,\perp}^2}\right) f_{\perp}{}^2(N_{range})}{\left(X_{c,\|}(y_i,z_j)+\frac{1}{K_p}X_{c,\perp}(y_i,z_j)\right)^2}\right]\right\}^{1/2} \quad (7.20)$$

$N_{\text{range}}=(r_{\text{high}}-r_{\text{low}})/15$ is the number of 15-m bins within the cloud found for 1064 nm calibration. $NSF_{\|}$, $NSF_{\perp}$ and NSF_{1064} are the noise scale factor (refer to Section 8) for 532 nm parallel and perpendicular and 1064 nm channel, respectively, in the G_A-normalized digitizer reading domain with range resolution ΔR and average shot number N_{shot}. ΔP_{bd} denotes the RMS noise (standard deviation) of background signal (in digitizer counts) including background radiation and dark current, i.e., the variance of baseline, and this parameter is measured on board for each channel. Both NSF and ΔP_{bd} must be converted using Eqs.(4.25) and (4.26), respectively. E and G_A are the laser energy and gain used to normalize lidar profiles. The random error in the averaged 1064 calibration coefficient for half orbit (day or night side) is then given by

$$\left(\Delta\bar{C}_{1064}\right)_{RAN} = \frac{1}{N_c}\sqrt{\sum_{i=1}^{N_c}\left(\Delta C_{1064}(y_i)\right)_{RAN}{}^2} \tag{7.21}$$

where N_c is the total number of profiles in a ½ orbit that are found with cloud and used for statistical computation of C_{1064}.

The method for estimating the 1064 nm noise scale factor, NSF_{1064}, is still being investigated and the value has been set to zero in the baseline code. This should not significantly affect the result since the term with NSF_{1064} is added to a background noise term which dominates the sum.

Again, for comparison, the standard deviation computed from all intermediate 1064 calibration coefficient from each average frame is also computed using

$$\left(\Delta C_{1064}\right)_{RAN} = \frac{1}{N_c}\sqrt{\sum_{i=1}^{N_c}\left(C_{1064}(y_i) - \bar{C}_{1064}\right)^2} \tag{7.22}$$

7.3. Simulation Results

Three orbits of LITE data (orbits 23, 24, and 27), acquired at low gain settings, had a sufficient number of unsaturated calibration quality clouds to test the 1064 nm calibration algorithm. The data for these orbits were rescaled using the CALIPSO simulator to obtain data with the correct averaging resolution and expected noise characteristics. Figure 7.2 shows the result of applying the 1064 nm calibration algorithm to CALIPSO simulated data from LITE orbit 23. Figure 7.2a plots the 1064 nm calibration coefficient versus peak 532 nm scattering ratio for each calibration cloud profile. It shows that the calibration coefficient is relatively stable with cloud intensity, although there is more variability with weaker clouds. Likewise, Figure 7.2b shows that the calibration coefficient is relatively independent of cloud altitude. Figure 7.2c shows the latitudes where the calibration clouds were obtained.

The results from all three orbits are summarized in Table 7.2. The mean calibration coefficient for each orbit is within about 4% of the simulator calibration value of 2.49E+20. The standard deviations of the calibration coefficients are on the order of 5%. These results are consistent with the expected uncertainties discussed in the previous section. Outlier rejection (rejection of all calibration points outside a preselected number of standard deviations of the mean) is applied to further improve results.

Table 7.2 C_{1064} using CALIPSO Simulated Data

LITE Orbit #	No Outlier Rejection	Outlier Rejection
Orbit 23	2.47E20 ± 1.19E19 (87 profiles)	2.49E20 (66 profiles)
Orbit 24	2.36E20 ± 1.24E19 (38 profiles)	2.39E20 (34 profiles)
Orbit 27	2.48E20 ± 9.97E18 (160 profiles)	2.50E20 (115 profiles)

Truth ~2.49E20

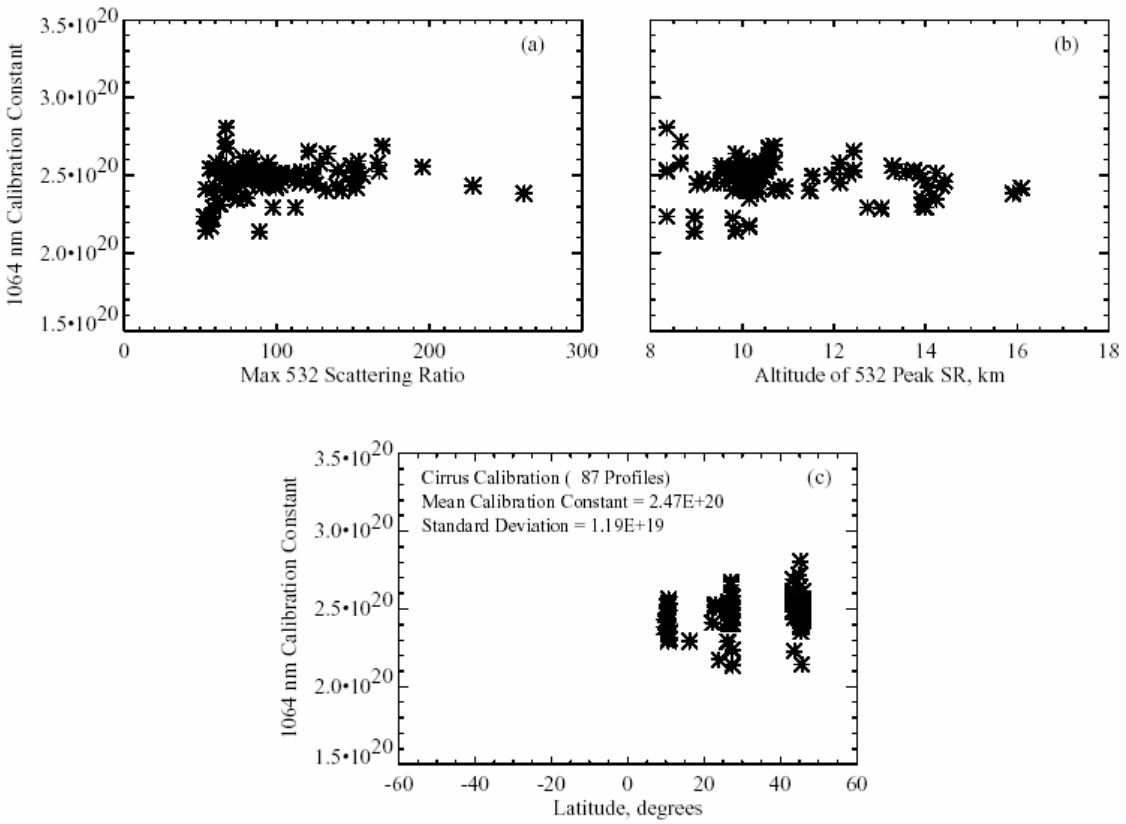


Figure 7.2 C_{1064} versus: a) maximum 532 nm cloud scattering ratio; b) altitude of peak 532 nm cloud scattering ratio; and c) latitude of calibration cloud profiles for CALIPSO simulation of LITE orbit 23.

7.4. 1064 Channel Attenuated Backscatter Profiles

Computation of the attenuated backscatter profiles from the 1064 nm input profiles is very simple and is accomplished via

$$\hat{\beta}'_{1064}(k,z) = \frac{1}{\hat{C}_{1064}(k)} X_{1064}(k,z) \tag{7.23}$$

where,

$\hat{\beta}'_{1064}(k,z)$ = Level 1 output profile product for the kth attenuated backscatter profile

$X_{1064}(k,z)$ = Level 1A output for the kth 1064 nm profile

$\hat{C}_{1064}(k)$ = 1064 nm calibration coefficient computed for that profile from the calibration coefficient time series in the Calibration Data Product File

8. Computing Noise Scale Factor (NSF)

This section describes the algorithm used to estimate the uncertainty due to random detection noise (i.e., the random error) in the attenuated backscatter coefficients produced in the Level 1 processing. The random error due to the background radiation, detector dark current, and thermal noise, referred to as the background noise, can be determined from the baseline signal (totalling 1000 samples from 60.3-75.3 km altitudes) measured onboard for each shot by computing the standard deviation of the baseline signal samples. The random error due to the lidar scattering signal noise (shot noise) is quantified via the Noise Scale Factor (NSF), which, combined with the background noise, is used in the Level 1 and Level 2 algorithms to generate uncertainty estimates and propagate those uncertainties into uncertainty estimates for Level 1 calibration products and Level 2 derived products (e.g., backscatter and extinction).

The uncertainty due to random detection noise can be estimated statistically by (1) computing the standard deviation of a series of data samples or by (2) applying NSF to each individual measurement. The former, more conventional method can be used only for segments of data where the expected value of the signal is very stable; otherwise significant overestimation of the error would result due to real (i.e., geophysical) variations as opposed to random variations in the signal. Figure 8.1 presents profiles of standard deviations computed from 100 cloud-free profiles (conventional method) and from one profile using the NSF. The data was acquired by the Cloud Physics Lidar (CPL) (McGill, 2002) which is an air-borne, nadir-looking system using photon-counting detection and can provide measurements of photon counts. Significant overestimation (~25%) is seen in the planetary boundary aerosol layer below ~1.5 km, though this layer looks well mixed and relatively stable, whereas good agreement appears in the "clear air" region above ~1.5 km where molecular scattering dominates the lidar return signal.

With the NSF method, on the other hand, a value of the random error due to noise can be estimated for each data sample, and thus does not rely on the statistics of a number of the measured samples. The estimate of the uncertainty in the sample attenuated backscatter coefficient is based on values of the sample itself, and is calculated essentially by scaling the observed attenuated backscatter coefficient by the NSF, with some correction for differences in averaging scales. As shown by the example in Figure 8.1, the standard deviation profile computed using NSF was derived from each single bin sample at an altitude of the profile, whereas the profile derived using the conventional method is computed from 100 samples for each altitude separately in the 100 lidar return profiles.

The use of the NSF for calculating uncertainties due to shot noise is based on the facts that (a) the magnitude of the observed attenuated backscatter coefficient at any point in the profile can be related to the number of photons effectively averaged into the data sample from which the attenuated backscatter are computed, and (b) the shot noise follows Poisson statistics which has a proportional relationship between the standard deviation (a measure of error) and the square root of the mean. The number of photons that arrive in a fixed time period fluctuates even when emitted by a light source having constant intensity. This fluctuation is due to the quantum nature of photons and constitutes the shot noise (also called quantum noise). We note that, although the distribution of the multiplied photoelectrons by a PMT or APD is no longer Poisson and the multiplication process introduces an excess noise when analog detection is used, the standard deviation still has a proportional relation to the square root of mean number of the multiplied photoelectrons. Calculation of the NSF amounts to computing the number of photons per

digitizer count of measured signal, and in some sense can be viewed as a continuous calculation of the transfer function of the lidar. The calculation is based on the observed variance in the measured daytime background signal and is based on the background monitor reading. Only the 532 nm channels have background monitors, hence the calculation of the NSF is limited to those channels. No method has been devised for estimating random signal error for the 1064 nm data on a point-by-point basis. More detailed discussions are given in the following subsections.

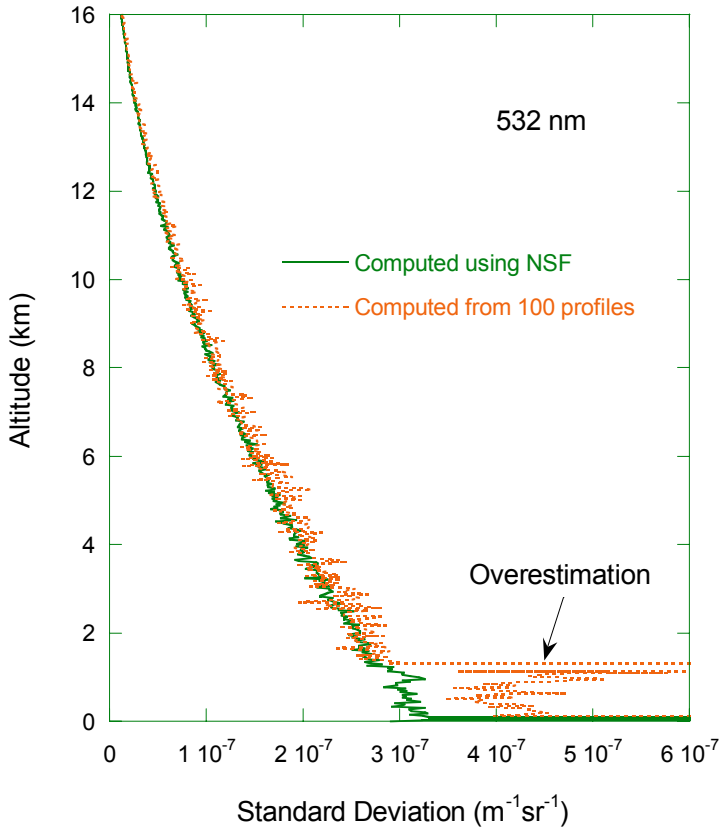


Figure 8.1 Example of standard deviations computed for each altitude from 100 cloud-free lidar return profiles (green curve) and from one profile using the NSF (read curve).

8.1. Theoretical Basis

For statistically stationary light (constant intensity), the photon numbers counted in equal sample times are Poisson distributed and the probability distribution can be described by

$$P(n_p) = \frac{\overline{n}_p{}^n}{n_p!} e^{-\overline{n}_p}, \tag{8.1}$$

where n_p is the photon number, $\overline{n}_p$ is the mean (or the rate of Poisson function). The standard deviation, Δn_p, which is a measure of the random error, is related to the mean by

$$\Delta n_p = \sqrt{\overline{\left(n_p - \overline{n}_p\right)^2}} = \sqrt{\overline{n}_p}\ . \tag{8.2}$$

This is due to the well known quantum nature of light (Oliver, 1965; Saleh, 1978). Figure 8.2 presents examples of photon-count distributions of the lidar scattering signal and solar background signal as observed by CPL. The samples (~ 6,000) used to derive the scattering signal distribution are from 6 bins around 13 km (a "clear air" region) of 1000 cloud-free profiles. The samples (~100,000) used to derive the solar background signal distribution are from 100 subsurface bins of the same 1000 cloud-free profiles. Also shown are Poisson curves having the same mean as those of the measured distributions. Both scattering signal and solar background signal are shown to have a Poisson distribution.

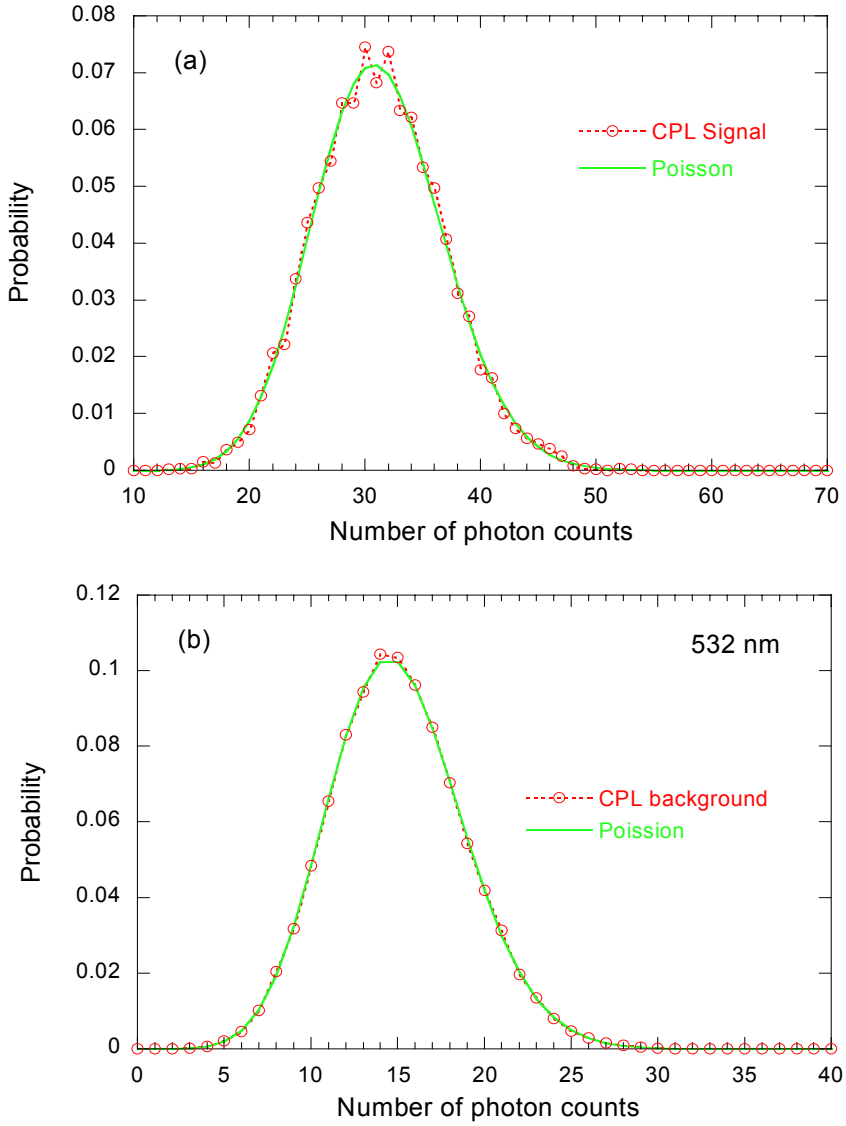


Figure 8.2 Photon-count distribution of (a) lidar scattering signal and (b) solar background signal observed by CPL.

Note that for a light source having a varying intensity, such as a thermal emission source, the arrival of photon counts at random times is a doubly stochastic Poisson process, and a "photon-bunching" noise due to the variation of the light source will be superimposed on the shot noise

(Saleh, 1978). This "photon-bunching" noise is, however, not a primary error source for the lidar measurement, because the laser light is a coherent source and its variation is generally very small. Additionally, in CALIOP the laser output is monitored shot-by-shot and the acquired lidar return profiles are normalized by the laser pulse energy.

When the collected photons are recorded in a domain $p=k\ n_p$ (e.g., the digitizer-reading domain for analog detection), where k is a transfer function (constant), we have

$$\Delta p = \begin{cases} k\sqrt{\overline{n}_p} = \sqrt{k}\sqrt{\overline{p}}, & \text{without excess noise;} \\ k\sqrt{F_m \cdot \overline{n}_p} = \sqrt{kF_m}\sqrt{\overline{p}}, & \text{with excess noise.} \end{cases} \tag{8.3}$$

where, F_m is the excess noise factor quantifying the extra noise introduced in the multiplication process due to the variability of multiplication gain when a PMT or APD operated in the analog detection mode is used. The multiplication in a typical PMT is a multiply stochastic Poisson process and can be described by a compound Poisson formula [Liu and Sugimoto, 2002]. The distribution of the multiplication gain of an APD has a more complicated form [McIntyre, 1972]. However, both PMT and APD type detectors have a proportional relationship between the standard deviation and the square root of the mean of the multiplied photoelectrons. And, they both introduce an extra noise due to the variability of multiplication gain which can be quantified by the excess noise factor F_m. Then we can introduce NSF by

$$NSF = . \begin{cases} \sqrt{k}, & \text{without excess noise} \\ \sqrt{k \cdot F_m}\ . & \text{with excess noise} \end{cases} \tag{8.4}$$

The random error due to the shot noise can then be estimated from a measurement p (if not very noisy) using

$$\Delta p = NSF\sqrt{\overline{p}} \approx NSF\sqrt{p} \tag{8.5}$$

Averaging over some number of range bins or shots is usually applied to lidar measurement data to reduce the noise so that high quality data products can be produced. The random error in data products derived from the averaged data can be fairly well estimated using Eq.(8.5). In practice, however, the background noise including solar radiation, dark current, and thermal noise, is also imposed in the lidar measurement. The total random error can then be estimated by

$$(\Delta p)^2 = (\Delta p_s)^2 + (\Delta p_{BG})^2 \approx NSF^2\, p + (\Delta p_{BG})^2\ . \tag{8.6}$$

Where, Δp_{BG} is the RMS noise of the background signal. It can be measured from samples where there is no scattering signal (e.g., below the surface or in very high altitudes).

For lidar measurements using an analog PMT or APD, the NSF in the digitizer-reading domain P is given by

$$NSF = \sqrt{2eBF_m G_m G_A} = \sqrt{eF_m G_m G_A \frac{1}{\Delta T_0}} = \sqrt{eF_m G_m G_A \frac{c}{2\Delta Z_0}}\ , \tag{8.7}$$

where

c = light speed;

e = electron charge;

G_m = multiplication gain;

G_A = gain of preamplifier converting the anode current to digitizer readings;

B = spectral bandwidth, $B \approx 1/2\Delta T_0$;

ΔT_0 = $2\Delta Z_0/c$ integration time, and

ΔZ_0 = range resolution.

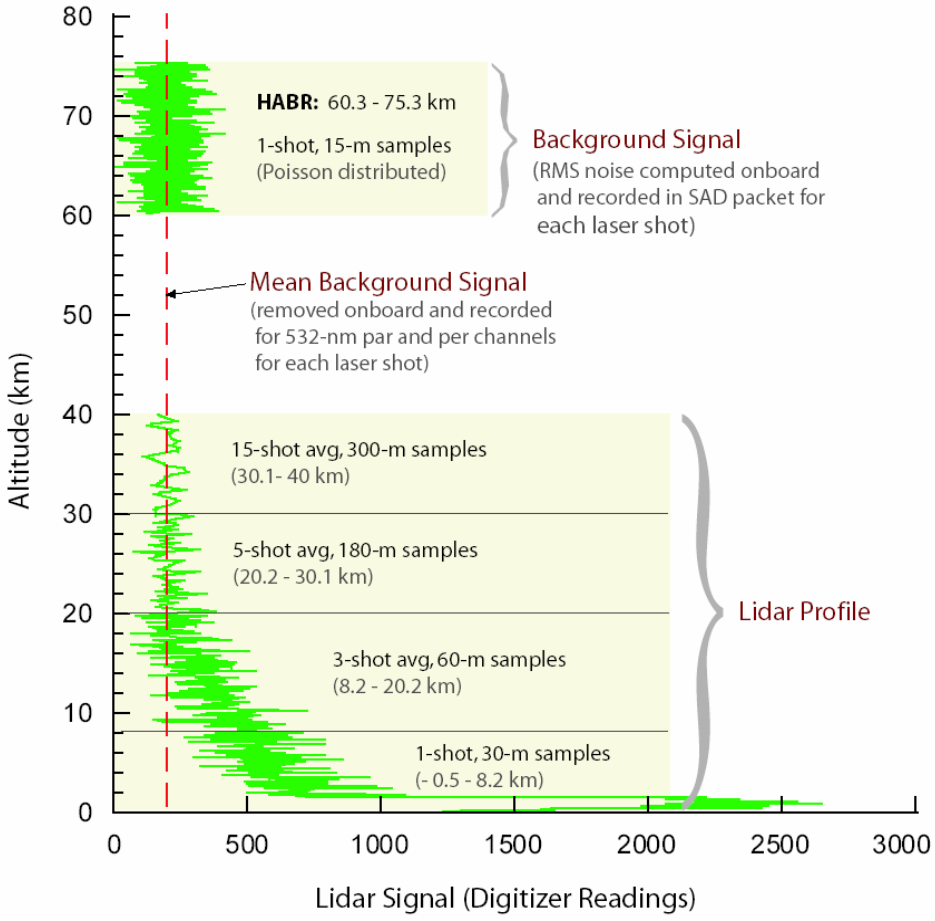


Figure 8.3 An example of simulated CALIOP profiles.

It is seen that the NSF is proportional to the square root of the product of a number of detector and electronic parameters/constants. If all these parameters/constants are known, the NSF can be computed directly using Eq. (8.7). In practice, however, the multiplication gain G_m and excess noise factor F_m may change over the course of a long observation period , or when the hardware undergoes environmental changes. CALIPSO, for example, will conduct a 3-year global observation and the NSF may change significantly over the CALIPSO observation period. Estimation of the NSF from on-orbit observations is therefore required.

The NSF is determined from solar background signals. Figure 8.3 presents an example of simulated lidar samples for the 532 nm parallel channel. Lidar returns are sampled at a range interval of 15 meters for each laser shot. Vertical averages are then performed on-board corresponding to different altitude ranges (20 bins for 30.1-40 km; 12 for 20.2-30.1km; 4 for 8.2-20.2km; and 2 for -0.5 – 8.2 km). Horizontal averages over a number of laser shots are computed for some altitude ranges (15 shots for 30.1-40 km; 5 for 20.2-30.1km; 3 for 8.2-20.2 km). 1000 15-m samples are collected in the High Altitude Background Region (HABR, 60.3-75.3 km) and used to compute RMS noise of the background signal for each shot onboard. In the HABR the scattering signal is negligibly small and the variation of lidar samples in this region is due mostly to the background noise. The mean background signal, which is estimated averaging samples from an altitude range of 97-112 km, is removed onboard but recorded by the background monitor for the two 532 nm channels. This quantity as well as the RMS noise determined from the HABR samples are stored in the a science ancillary data (SAD) packet for downlink. We note that, the mean background signal is measured by a separate digitizer (i.e., the background monitor), whereas the RMS noise is acquired by the science digitizers (consisting of two 12-bit digitizers). These two quantities (mean and standard deviation) acquired during daytime will be used to compute NSF for the 532 nm channels as discussed in below.

During daytime the solar radiation will dominate the background signal in the 532 nm channels, and the samples (in photon counts) in the HABR will follow Poisson statistics, as shown by the example in Figure 8.2. The sample time of a lidar return profile is very short (< 1 ms). During such a short time period the solar radiation is almost constant and the perturbation from the atmosphere is very small (the atmosphere is "frozen").

8.2. NSF Algorithm

8.2.1. Algorithm Description

For CALIOP, the range-scaled, energy-normalized, gain-normalized lidar signal is defined by Eq. (3.7) ($X(r) = r^2 \cdot P(r) / E \cdot G_A$). Because G_A will have different values for the daytime and nighttime observations, the NSF in the digitizer-reading domain (see Eq. (8.7)) will have different nominal values for the daytime and nighttime segments. For this reason,the NSF will be computed in the G_A-normalized domain ($V=P/G_A$) for single shots and a 15-m sample interval. In this domain, the NSF has a constant value if the detector parameters (G_m and F_m) do not change. During daytime, the NSF can then be determined from the mean background signal and RMS background noise derived from the HABR samples (Figure 8.3) using

$$NSF = \frac{\Delta P_{BG,1shot}}{\sqrt{V_{BG,1shot}}} = \frac{\Delta P_{BG,1shot} / G_A}{\sqrt{V_{BG,1shot}}}, \tag{8.8}$$

where

$\Delta P_{BG,1shot}$ = single-shot background RMS noise, in science digitizer readings, including the noise due to the background radiation and detector dark current. This quantity is computed onboard from the 1000 15-m samples in the HABR for each single shot using the following equation:

$$\Delta P_{BG,1shot} = \sqrt{\frac{1}{1000}\sum_{j=1}^{1000}\left(P_{HABR,1shot}(r_j) - \bar{P}_{HABR,1shot}\right)^2} \tag{8.9}$$

$V_{\mathrm{BG,1shot}}$ = background signal, in G_A-normalized digitizer counts, for each single shot, converted from the background energy monitor using:

$$V_{BG,1shot} = (C0 + N_{BGMon,1shot} \times BGMonSens) \times TIAGain \times PostAmpGain \times SciDigSens \tag{8.10}$$

$N_{\mathrm{BGMon,1shot}}$ = background signal, in background monitor digitizer counts, recorded by the background energy monitor for each shot

$C0$ = equivalent offset in the anode current domain, in amps, applied to the background monitor,

default value: - 0.1782756 x10^{-6} amps/count, for P channel;

+0.1844652 x10^{-6} amps/count, for S channel

BGMonSens = background monitor sensitivity,

default value: 0.000760019x10^{-6} amps/count, for P channel;

0.000759954x10^{-6} amps/count, for S channel

TIAGain = TIA gain, 2.49x10^3 volts/amp

PostAmpGain = gain of the post amplifier, 1.25 volts/volt

SciDigSens = science digitizer sensitivity, 8192 counts/volt.

8.2.2. Simulations and Algorithm Tests Using LITE Data

To estimate the average number of laser shots required to achieve accurate NSF determination, retrieval simulations were conducted. Table 8.1 lists relative errors in the NSFs retrieved from simulated CALIPSO 532 nm parallel and perpendicular data. Retrievals using single shot data and averages over 5, 20, 80, and 320 km have been considered. Simulations have been conducted for both the parallel and perpendicular channels. The results indicate that similar accuracy for the NSF estimate can be achieved using equivalent data averaging for the two channels. Also, due to the much higher background signals, daytime measurements provide much more accurate NSF estimates than nighttime measurements. Simulations with daytime measurements show that the NSF estimate can be determined with less than 10% relative error using data from a single shot. To achieve the same accuracy, nighttime measurements must be averaged over horizontal distances greater than 80 km (240 shots).

The NSF algorithm has also been tested using LITE data. Figure 8.4 presents NSFs retrieved from single shot, 10, and 100 shot averaged profiles of the daytime portion of orbit 117 observations. The results show that NSF varies by more than 10% over this segment of LITE data. The peak around index 22 in Figure 8.4 is due to saturation of the background monitor digitizer.

Table 8.1 Relative error of NSFs retrieved from simulated CALIPSO data

Average	1 shot		5 km		20 km		80 km		320 km	
	night	day	night	day	night	day	night	day	night	day
Parallel (%)	---	5.80	72.3	1.50	28.1	0.74	13.5	0.34	6.2	---
Perpend (%)	---	5.70	---	1.46	26.0	0.76	13.6	0.35	---	---

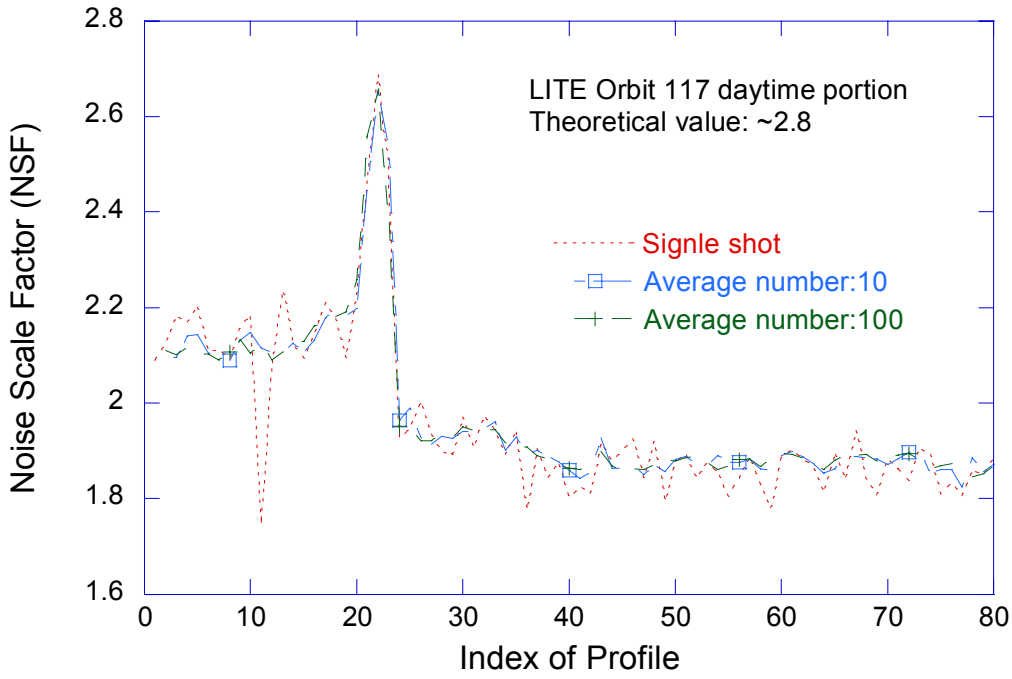


Figure 8.4 NSF values computed from LITE data. The horizontal axis is the index of every 100th single-shot profile.

8.2.3. Operational Procedures

8.2.3.1. 532 nm Day Orbit Segment

For a 5 km frame (15 shots) the following steps are applied respectively to 532 nm parallel and perpendicular channels:

1) Average (a) the background RMS noise and (b) the background signal measured by the background energy monitor over 15 shots within in a frame, using

$$\Delta P_{BG,1\,fraome\_avg} = \sqrt{\frac{1}{15}\sum_{k=1}^{15}\left(\Delta P_{BG,1shot,k}\right)^2} \qquad (8.11)$$

$$P_{BGMon,1frame\_avg} = \frac{1}{15}\sum_{k=1}^{15} P_{BGMon,1shot,k} \tag{8.12}$$

2) Convert $P_{\mathrm{BGMon,1frame\_avg}}$ to $V_{\mathrm{BG,1frame\_avg}}$ using Eq.(8.10)

3) Compute NSF using

$$NSF = \frac{\Delta P_{BG,1frame\_avg} / G_A}{\sqrt{\overline{V}_{BG,1frame\_avg}}} \tag{8.13}$$

4) Assign the computed NSF using Eq. (8.13) to each profile in the frame.

NOTE: Because the solar background signal is not high enough for an accurate NSF measurement in day-night transition regions, an offset can be set to the two ends of the daytime portion. That is, a number of frames in the beginning and the end of the daytime portion data are classified as nighttime data. In addition, for the daytime data, the digitizer reading is checked for each frame so that only data having reliable readings are used.

8.2.3.2. 532 nm Night Orbit Segment

Compute the mean value of NSFs derived from each daytime frame and apply this mean value to all nighttime profiles.

8.2.3.3. 1064 nm

Because the CALIOP 1064 nm channel does not provide measurements of background signals, the procedure described above for 532 nm channels cannot be used in this channel. A method may be determined in the future using data from ground-based system tests or on-orbit observations. Initially the NSF will be set to 0 for the 1064 nm channel. Alternative methods for estimating NSF for the 1064 nm channel will be investigated during/following the on-orbit check-out period.

8.3. Application of Algorithm and Averaging Issues

The NSF is used to estimate the uncertainty in calibration coefficients determined by the Level 1 algorithms and is also used in the determination of uncertainties in data products produced by the Level 2 algorithms. This section addresses the mechanics of applying the NSF to the basic attenuated backscatter product. In particular, it concerns scaling the NSF to account for the difference between the vertical resolution at which the NSF is computed and that at which the products are averaged. The NSF is computed at the native, on-board resolution of the lidar samples (15-meter range interval and single shots), but is applied to attenuated backscatter profiles that are computed at various vertical resolutions determined by the onboard averaging scheme. Moreover, the attenuated backscatter coefficients may be further averaged to produce Level 2 products. In order to apply the NSF, compensation must be made for the averaging scale of the product in question: if data are vertically averaged, then the noise estimated via the NSF must be reduced appropriately.

Typically, this compensation is performed by dividing by a simple ratio of the averaging scales; i.e., the square root of the bin number of 15-m samples averaged, $N_{\mathrm{bin}}=\Delta r/\Delta r_{\mathrm{samp}}$, (where Δr_{samp} =

15 meters is the sample time interval and Δr is the averaging range) and the shot number, N_{shot}, for average. For example, in the G_A-normalized digitizer-reading domain, $V=P/G_A$,

$$\Delta V = \frac{1}{\sqrt{N_{bin}N_{shot}}}\sqrt{NSF^2\bar{V} + \overline{\left(\Delta V_{BG}\right)^2}} = \sqrt{\frac{NSF^2}{N_{bin}N_{shot}}\bar{V} + \frac{\overline{\left(\Delta V_{BG}\right)^2}}{N_{bin}N_{shot}}}\,. \tag{8.14}$$

This effect is taken into account using Eqs. (4.25) and (4.26). However, the correction for varying averaging scales is complicated by the fact that the on-board sampling interval (Δr_{samp}=15 meters) is smaller than the effective range resolution due to the electronic bandwidth of the detection system ($\Delta r_{res}\approx$30 meters). (We note that the highest vertical resolution of the downlinked data is 30 m, due to averaging that takes place on-board the satellite before downlink. However, some of the parameters from which the NSF is estimated are computed from data at the raw instrument resolution of 15 meter.) Because of this, each lidar sample in a downlinked full-resolution profile will be partially correlated to the nearest 2 or 3 neighbors. Figure 8.5a shows an example of the autocorrelation coefficient derived from topmost 2500 samples of a subset (6000 profiles) of the LITE orbit 117 data. Because of the partial correlation between the neighbor samples, vertical averaging does not reduce the variance in lidar samples as much as it would if the samples were independent. For independent samples, the standard deviation of the ensemble can be reduced by a factor of $(N_{bin})^{1/2}$, where N_{bin} is the number of range bins over which the data are averaged. This is shown in Figure 8.5b.

As mentioned earlier, the downlinked CALIOP profiles have different range resolutions for different altitude regions (averaged onboard over a different number of 15-m samples). Further vertical averaging may also be required during science data processing to improve the quality of retrievals. In estimating the random error due to noise, a correction to the sample correlation must be applied to the vertically averaged samples or the parameters retrieved from the averaged samples (e.g., extinction and backscatter coefficients). This correction is also required for estimating random errors in layer-averaged products (e.g., mean attenuated backscatter, integrated attenuated color ratio, and total depolarization ratio). A correction function $f(N_{bin})$ is therefore introduced. It can be derived straightforwardly from the measured standard deviations with different average bin number, such as those in Figure 8.5b, using

$$f(N_{bin}) = \frac{\Delta V(N_{bin})}{\Delta V(N_{bin}=1)/N_{bin}^{1/2}} \tag{8.15}$$

The correction function computed using Eq. (8.15) from the standard deviations shown in Figure 8.5b is presented in Figure 8.5c (dashed curve). This correction function can also be derived from the autocorrelation coefficient based on the error propagation theory (e.g., Bevington and Robinson, 1992):

$$f(N_{bin}) = \left[1 + 2\sum_{m=1}^{N_{bin}-1}\left(\frac{N_{bin}-m}{N_{bin}}\right)R(m)\right]^{1/2}, \tag{8.16}$$

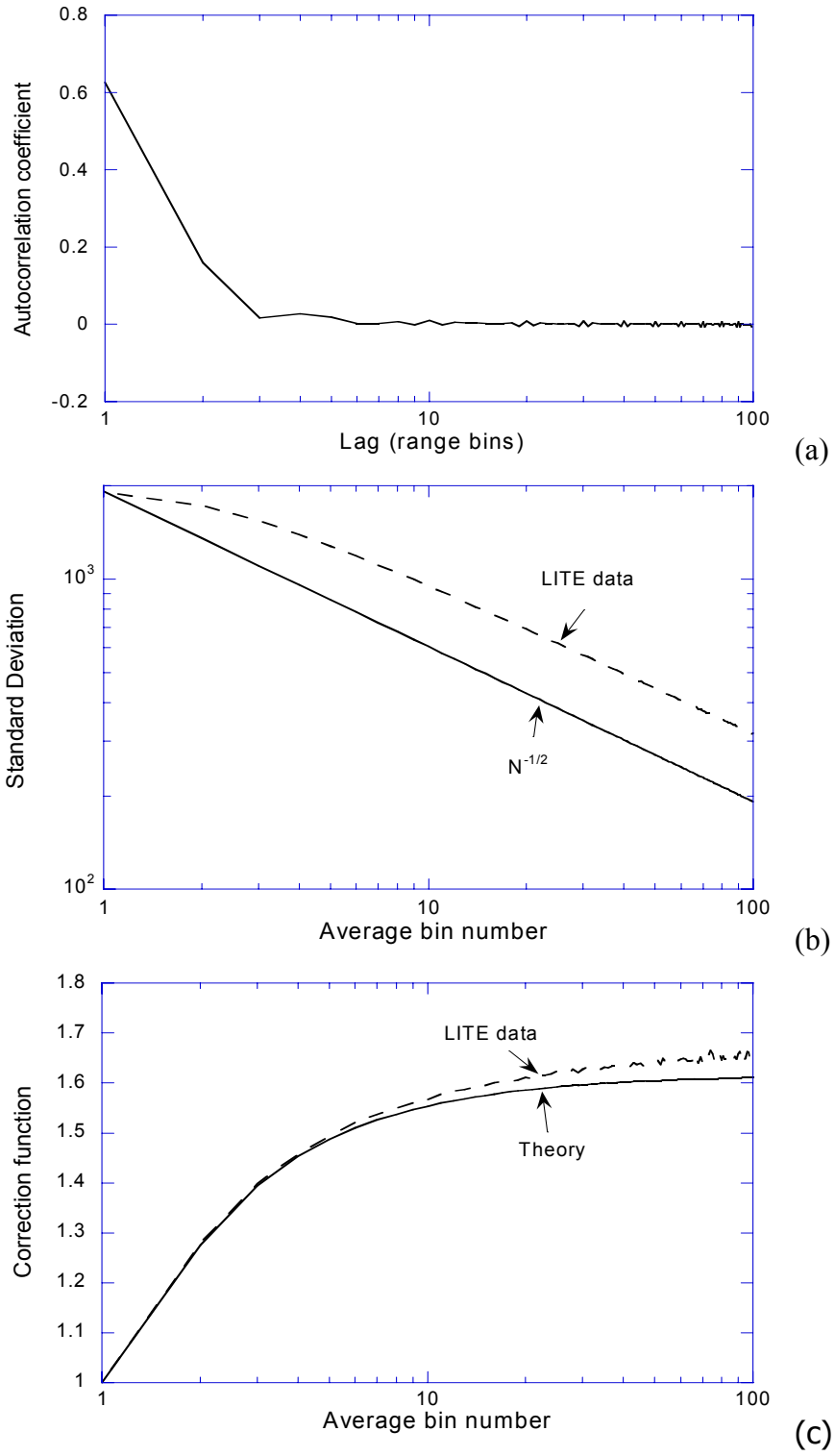


Figure 8.5. (a) Autocorrelation coefficient derived from topmost 2500 samples of a subset (6000 profiles) of the LITE orbit 117 data. (b) Reduction in standard deviation with averaging. (c) Averaging correction function as a function of the number of vertical range bins averaged.

where $R(m)$ is the autocorrelation coefficient for a lag of m range bins (Figure 8.5a). The correction function computed from the measured autocorrelation coefficient shown in Figure 8.5a using Eq.(8.16) is also presented in Figure 8.5c (solid curve). It is seen that the f-function computed using Eq.(8.15) is consistent with that using Eq.(8.16) for small averaging intervals. It is however larger than that using Eq.(8.16) for large averaging intervals. This is most likely due to systematic error such as baseline ripple or oscillation (the baseline slope has been removed in the computation). Note that when the sample range interval is larger than the range resolution determined by the electronics, the autocorrelation coefficient is zero for lag ≥ 1, and $f = 1$. In this case, no correction is required.

For Level 1 processing, a look-up table for the f function is constructed using Eqs. (8.15) and (8.16). The table will be updated based on measured values after the CALIPSO launch. The random error in attenuated backscatter with the vertical averaging correction is given by

$$\begin{aligned}\Delta\beta'_{corr}(r) &= \left[\frac{NSF^2}{N_{bin}\cdot N_{shot}}\frac{r^2}{E\cdot C}\beta'(r)+\left(\frac{r^2}{E\cdot C}\right)^2\frac{\left(\Delta P_{BG,1shot}\right)^2}{N_{bin}\cdot N_{shot}}\right]^{1/2} f(N_{bin}) \\ &= \Delta\beta'(r)\cdot f(N_{bin})\ ,\end{aligned} \tag{8.17}$$

where N_{shot} is the average number of shots, (i.e., the ratio of range resolution to the sample range interval $N_{\text{bin}}=\Delta r/\Delta r_{\text{samp}}$, where $\Delta r_{\text{samp}}=15$ meters in the CALIOP case), and $\Delta\beta'$ is the random error in attenuated backscatter without the vertical average correction. We note that, no correction is needed for horizontal averaging because the random errors of samples at a given range bin in different profiles are statistically independent.

9. References

Ansmann, A., et al.: "Lidar network observations of cirrus morphological and scattering properties during the International Cirrus Experiment 1989: The 18 October 1989 case study and statistical study", *J. Appl. Meteo*, Vol. 32, 1608-1622, (1993).

Bates, D. R.: "Rayleigh scattering by air", *Planet. Space Sci.*, **32**, 785-790 (1984).

Bevington, P. R and D. K. Robinson, 1992: Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, 328 pp.

Beyerle, G., et al.: "A lidar and backscatter sonde measurement campaign at Table Mountain during February –March 1997: Observations of cirrus clouds", *J. Atmos. Sci*. **58**, 1275–1287, (2001).

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

Bohren C. F. and D. R. Huffman, Absorption and scattering of light by small particles, (Wiley, New York, 1983).

Bucholtz, A., "Rayleigh-scattering calculations for the terrestrial atmosphere", *Appl. Opt.*, **34**, 2765-2773 (1995).

Cairo, F., G. Di Donfrancesco, A. Adriani, L. Pulvirenti, and F. Fierli: "Comparison of various linear depolarization parameters measured by lidar", *Appl. Opt.* **38**, 4425-4432, (1999).

Collis, R. T. H., and P. B. Russell: "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.

Elterman, L.: "UV, Visible, and IR Attenuation for Altitudes to 50 km", Rep. AFCRL-68-0153, Environmental Research Papers, No. 285, (Air Force Cambridge Res. Lab., Bedford, MA, 1968).

Jäger H., and T. Deshler: "Lidar backscatter to extinction, mass and area conversions for stratospheric aerosols based on midlatitude balloonborne size distribution measurements", *Geophys. Res. Lett.*, **29**, doi:10.1029/2002GL015609 (2002).

Jäger, H., T. Deshler, and D. J. Hofmann: "Midlatitude lidar backscatter conversions based on balloonborne aerosol measurements", *Geophys. Res. Lett.* **22**, 1729-1732 (1995).

Li, X., and H. J. Götze: "Tutorial: Ellipsoid, geoid, gravity, geodesy, and geophysics", *Geophysics*, **66**, 1660-1668 (2001).

Liu, Z., and N. Sugimoto: "Simulation study for cloud detection with space lidars by use of analog detection photomultiplier tubes", *Appl. Opt.* **41**, 1750-1759 (2002).

Liu, Z., M. McGill, Y. Hu, C. Hostetler, M. Vaughan, and D. Winker: "Validating lidar depolarization calibration using solar radiation scattered by ice clouds", *IEEE Geoscience and Remote Sensing Letters*, Vol. 1., 157-161 (2004).

McGill, M., et al.: "Cloud Physics Lidar: instrument description and initial measurement results", *Appl. Opt.* **41**, 3725-3734 (2002).

McGuire, J. P. and R. A. Chapman, "Analysis of spatial pseudo depolarizers in imaging systems", *Opt. Eng.*, **29**, 1478-1484 (1990).

McIntyre, R. J.: "The distribution of gains in uniformly multiplying avalanche photodiodes: Theory," *IEEE Trans. Electron Devices*, Vol. ED-19, 703-713 (1972).

Oliver, B. M., "Thermal and quantum noise," *Proc. IEEE* **53**, 436-454 (1965).

Powell, K. A., "The Development of the CALIPSO Lidar Simulator", Masters Thesis, College of William and Mary, Williamsburg VA USA 2005.

Powell, K. A., W. H. Hunt, and D. M. Winker, "Simulations of CALIPSO Lidar Data", in *Lidar Remote Sensing in Atmospheric and Earth Sciences*, L. R. Bissonnette, G. Roy and G. Vallée (editors), Defence R&D Canada - Valcartier, Québec, Canada, pp. 781-784, 2002.

Reagan, M. J., X.Wang, and M. J. Osborn: "Spaceborne lidar calibration from cirrus and molecular backscatter returns," *IEEE Trans. Geosci. Remote Sens.*, 40, 2285– 2290, (2002).

Saleh, B., Photoelectron Statistics with Application to Spectroscopy and Optical Communication, Vol.6 of Springer Series in Optical Sciences (Springer-Verlag, Berlin,1978), Chapter 5.

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

SPARC, Assessment of Stratospheric Aerosol properties (ASAP), SPARC Report No. 4, WMO/TD-No. 1295 (2006).

Trepte. C. R., and M. H. Hitchman: "Tropical stratospheric circulation deduced from satellite aerosol data", *Nature*, **335**, 626-628 (1992).

Van de Hulst, H. C., Light Scattering by Small Particles, Dover Publications, Inc., (New York, 1981), p107.

Winker, D. M., R. Couch, and M. P. McCormick: "An overview of LITE: NASA's Lidar In-space Technology Experiment", *Proc. IEEE*, **84**, 164–180 (1996).

Young, A. T.: "Revised depolarization corrections for atmospheric extinction", *Appl. Opt.*, **19**, 3427-3428 (1980).