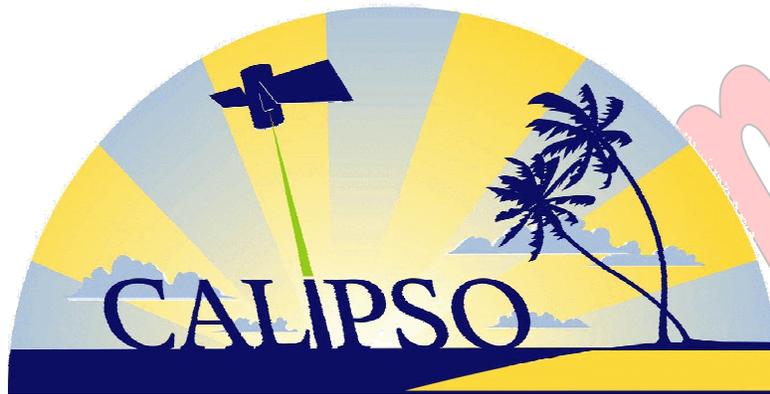


# CALIOP Algorithm Theoretical Basis Document

## Part 4: Extinction Retrieval Algorithms



### Primary Authors:

Stuart A. Young, CSIRO Marine & Atmospheric Research, Aspendale, Victoria, Australia  
David M. Winker, NASA Langley Research Center, Hampton, Virginia, USA  
Mark A. Vaughan, Science Systems & Applications Inc. (SSAI), Hampton, Virginia, USA  
Yongxiang Hu, NASA Langley Research Center, Hampton, Virginia, USA  
Ralph E. Kuehn, Science Systems & Applications Inc. (SSAI), Hampton, Virginia, USA

**PC-SCI-202 Part 4**

**Release 1.0 – DRAFT**

**Last Update: 28 January 2008**

# Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observations

## CALIOP Algorithm Theoretical Basis Document

Document No: PC-SCI-202 Part 4

Prepared By:

---

Stuart A. Young

Date

---

Mark A. Vaughan

Date

---

Yongxiang Hu

Date

---

Ralph E. Kuehn

Date

Approved By:

---

David M. Winker  
CALIPSO Principal Investigator

Date

## Table of Contents

|          |   |    |
|----------|---|----|
| 1.       | Introduction.....   | 4  |
| 1.1.     | Purpose and Scope .....   | 5  |
| 1.2.     | Related Documents .....   | 5  |
| 1.3.     | Revision History .....  | 5  |
| 2.       | Extinction Algorithm .....  | 6  |
| 2.1.     | Background Motivation and Overview.....                                   | 6  |
| 2.2.     | Analysis of CALIPSO Data – Scene and Profile Processes .....              | 9  |
| 2.3.     | Retrieval of Particulate Extinction Profiles .....                        | 13 |
| 2.3.1.   | Important Considerations in the Solution of the Lidar Equation.....       | 14 |
| 2.3.1.1. | The elastic backscatter lidar equation.....                               | 14 |
| 2.3.1.2. | Extinction retrieval formula and algorithms.....                          | 15 |
| 2.3.1.3. | Stability of retrievals and constrained and unconstrained solutions.....  | 16 |
| 2.3.2.   | Mathematical Basis for the CALIPSO Extinction Solver.....                 | 16 |
| 2.3.2.1. | Newton’s (Newton-Raphson) Method algorithm .....                          | 21 |
| 2.3.2.2. | Special considerations in the renormalization of CALIPSO retrievals ..... | 22 |
| 2.3.2.3. | Lidar ratio adjustment to prevent diverging and negative retrievals ..... | 22 |
| 2.3.2.4. | Error / Uncertainty analyses.....   | 24 |
| 2.3.3.   | Sensitivity Analysis .....  | 27 |
| 2.3.3.1. | Calibration and Renormalization Errors .....                              | 28 |
| 2.3.3.2. | Lidar Ratio Errors .....  | 29 |
| 2.3.3.3. | Transmittance Constraint Errors .....                                     | 30 |
| 2.4.     | Implementation of the Hybrid Extinction Retrieval Algorithm (HERA).....   | 32 |
| 2.4.1.   | Introduction.....   | 32 |
| 2.4.2.   | Overview of the Algorithms .....  | 33 |
| 2.4.2.1. | Simple features and clear regions .....                                   | 33 |
| 2.4.2.2. | Complex features .....  | 34 |
| 2.4.2.3. | Complex embedded features.....  | 36 |
| 2.4.3.   | Detailed Description of the Analysis Process.....                         | 36 |
| 2.4.3.1. | Definitions.....  | 36 |
| 2.4.3.2. | Required Input Data.....  | 37 |
| 2.4.3.3. | Internal and External Output Data .....                                   | 40 |

|          |   |    |
|----------|---|----|
| 2.4.3.4. | Flowchart Annotations.....  | 42 |
| 2.4.4.   | Description of the Steps Shown in the Flow Charts .....           | 44 |
| 2.4.4.1. | Simple features.....  | 44 |
| 2.4.4.2. | Complex features .....  | 47 |
| 2.4.4.3. | Complex embedded features.....                                    | 53 |
| 3.       | Accounting for Multiple Scattering in Extinction Retrievals ..... | 55 |
| 3.1.     | Introduction.....   | 55 |
| 3.2.     | Approach.....   | 56 |
| 3.2.1.   | Specifics.....  | 56 |
| 4.       | References.....   | 61 |

## 1. Introduction

CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) is a joint NASA CNES satellite mission that was launched on April 28, 2006 into a sun-synchronous orbit at an altitude of 705 km and an inclination of 98.2° (Winker, et al., 2007). The spacecraft orbits with four other Earth observing satellites as part of the Afternoon or A-Train constellation (Stephens et al., 2002). CALIPSO is designed to provide measurements aimed at improving our understanding of the role of aerosols and clouds in the climate system. The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP, pronounced the same as “calliope”) is the primary instrument on the CALIPSO satellite. CALIOP, built by Ball Aerospace and Technologies Corporation, Fibertek, and NASA’s Langley Research Center (LaRC), is designed to acquire vertical profiles of elastic backscatter at two wavelengths (1064 nm and 532 nm) from a near nadir-viewing geometry during both day and night phases of the orbit. In addition to the total backscatter at the two wavelengths, CALIOP also provides profiles of linear depolarization at 532 nm. Accurate aerosol and cloud heights and the retrieval of extinction coefficient profiles are derived from the total backscatter measurements. The depolarization measurements enable the discrimination between ice clouds and water clouds and the identification of non-spherical aerosol particles. Additional information, such as estimates of aerosol particle size, is obtained from the ratios of the signals obtained at the two wavelengths. In addition to CALIOP, the CALIPSO satellite payload includes an Imaging Infrared Radiometer (IIR), provided by CNES, with three channels in the infrared region optimized for retrievals of cirrus particle size, and a single channel Wide Field Camera (WFC). The WFC is a moderate spatial resolution camera operating in the visible regime to provide meteorological context for CALIOP and IIR measurements, and as a means of accurately co-registering CALIPSO observations to those from other instruments in the A-Train such as MODIS (MODerate Resolution Imaging Spectroradiometer).

The CALIOP Lidar Level 2 Algorithm Theoretical Basis Document (ATBD) contains five chapters that address four primary classes of retrieval algorithms. The first chapter (PC-SCI 202, Part 1) provides an introduction to the CALIPSO mission, a description of the CALIOP instrument, and an overview of the CALIOP data analysis architecture. Layer detection strategies and the Scene Classification Algorithms (SCA) are described in chapters 2 (PC-SCI 202, Part 2) and 3 (PC-SCI 202, Part 3), respectively. In this, the fourth chapter, we describe the

Hybrid Extinction Retrieval Algorithm (HERA), which extracts range-resolved estimates of particulate backscatter and extinction coefficients from the attenuated backscatter profiles measured by CALIOP. As the name implies, HERA is a hybrid collection of procedures that have been combined to form a single, fully automated retrieval scheme. Operating interactively with the SCA modules, HERA derives optimized profiles of layer optical properties for all cloud and aerosol layers identified by the Selective Iterated Boundary Locator (SIBYL) algorithm. In the remaining sections of this document we review the science objectives and the motivation underlying the development of HERA; derive the physical and mathematical bases for HERA's extinction profile-solving engine; develop a comprehensive error propagation scheme; illustrate the data flow within the processing stream; and fully describe all necessary and contingent relationships between the different processing modules incorporated into the HERA architecture.

## 1.1. Purpose and Scope

This document is the first release of the CALIOP Algorithm Theoretical Basis Document Part 4: Extinction Retrieval Algorithms and describes the extinction retrieval algorithms used in the production of CALIPSO's Level 2 data products. It joins the series of CALIOP ATBD documents that collectively provide an overview of the CALIPSO mission and describe the lidar, on-board data reduction procedures, Level 1 data handling and calibration procedures, and Level 2 data production algorithms. This document provides an overview of the extinction algorithm structure and mathematical basis.

## 1.2. Related Documents

- PC-SCI-201: CALIOP Algorithm Theoretical Basis Document: Calibration and Level 1 Data Products
- PC-SCI-202 Part 1: CALIOP Algorithm Theoretical Basis Document: CALIOP Instrument, and Algorithms Overview
- PC-SCI-202 Part 2: CALIOP Algorithm Theoretical Basis Document: Feature detection and layer properties algorithms
- PC-SCI-202 Part 3: CALIOP Algorithm Theoretical Basis Document: Scene classification algorithms
- PC-SCI-202 Part 4: CALIOP Algorithm Theoretical Basis Document: Extinction retrieval algorithms (this document)
- PC-SCI-202 Part 5: CALIOP Algorithm Theoretical Basis Document: Post-extinction processing and particle property algorithms
- PC-SCI-202 Part 6: CALIOP Algorithm Theoretical Basis Document: Appendices
- PC-SCI-503: CALIPSO Data Products Catalog

## 1.3. Revision History

The algorithms are described in this document as they have been implemented in Version 2.0 of the CALIPSO production data processing system. Subsequent changes, if any, will be

documented in Table 1.

Table 1: Revision history of the CALIOP Hybrid Extinction Retrieval Algorithms

| Version | Release Date    | Comments   |
|---------|-----------------|--|
| 1.0     | 28 January 2008 | This baseline release for this ATBD supplants all descriptions of the extinction and post-extinction algorithms found in the preliminary drafts of the CALIOP Level 2 ATBD released prior to January 15, 2008. |

## 2. Extinction Algorithm

### 2.1. Background Motivation and Overview

The retrieval of particulate extinction from measurements by space-borne elastic-backscatter lidars such as CALIOP faces a number of challenges not encountered in the analysis of measurements made by ground-based lidars. These difficulties are the result a combination of factors, including the large distances between the lidar and the targets of interest (typically 500-700 km), the high speed at which the lidar sweeps across the target space ( $\sim 7 \text{ km sec}^{-1}$ ), constraints placed on the pulse energy of the laser transmitter by eye-safety requirements, the relatively low firing rate of the laser ( $\sim 20 \text{ Hz}$ ) relative to the velocity of the satellite, and vertical and horizontal variations in the composition of the layers being measured. Taken together, these factors can reduce the measurement signal-to-noise ratio (SNR) within clouds and aerosols to levels where the uncertainties in the retrieved data products become unacceptably high. Furthermore, applying the conventional averaging techniques traditionally used by the lidar community to enhance SNR is generally inappropriate, due both to the spatial and temporal separation between consecutive CALIPSO profiles, and to the varying and non-linear relationship between extinction and backscatter within the lidar measurements.

For CALIPSO, the routine analysis of a large volume of lidar profile data received each day ( $\sim 1.74$  million shots) introduces additional practical and computational challenges that require a robust, on-the-fly processing strategy to handle many different retrieval situations. Automated analysis techniques must be able to identify different optical properties of cloud and aerosol features (e.g., the particulate extinction-to-backscatter ratio, also commonly known as the lidar ratio) over different regions of the globe as well as within the same column. The processing system must be able to handle both features easily detectable on a single-shot basis and also faint ones that may require averaging a large number of profiles. In addition to these issues, the retrieval method used to derive particulate extinction and backscatter coefficients may depend upon the structure of the lidar signal. For example, while it is common to process ground-based lidar measurements using the so-called backward analysis direction in order to ensure stability of the solution (e.g. Klett, 1985; Young, 1995), this approach is typically not possible for the CALIOP data set, and hence the processing system must be able to identify and pursue alternate paths to derive tractable solutions.

To address many of these issues, a sophisticated, fully autonomous processing system has been developed for CALIPSO. Standard lidar processing typically averages a fixed number of profiles together, and performs retrievals of the atmospheric column using this averaged profile. A key aspect of CALIPSO analysis system is that it uses an innovative nested multi-grid

averaging scheme to process a composite lidar ‘scene’. Scenes are segments of what are commonly referred to as CALIPSO curtain files, which, as illustrated in Figure 2.1, are two-dimensional, height versus along-track distance, vertical swaths through the atmosphere. Through this approach, the processing system is able to identify and extract high SNR profile data from each scene and derive optical properties of clouds and aerosol layers. The CALIOP Level 2 processing system is composed of three modules, which have the general function of detecting layers, classifying these layers by type, and performing extinction retrievals. These three modules are the Selective Iterated Boundary Locator (SIBYL) and Scene Classifier Algorithm (SCA), and Hybrid Extinction Retrieval Algorithms (HERA), respectively. The rationale and design for the CALIPSO averaging scheme are reported in the layer detection ATBD (PC-SCI-202 Part 2) and in Vaughan et al. (2004).

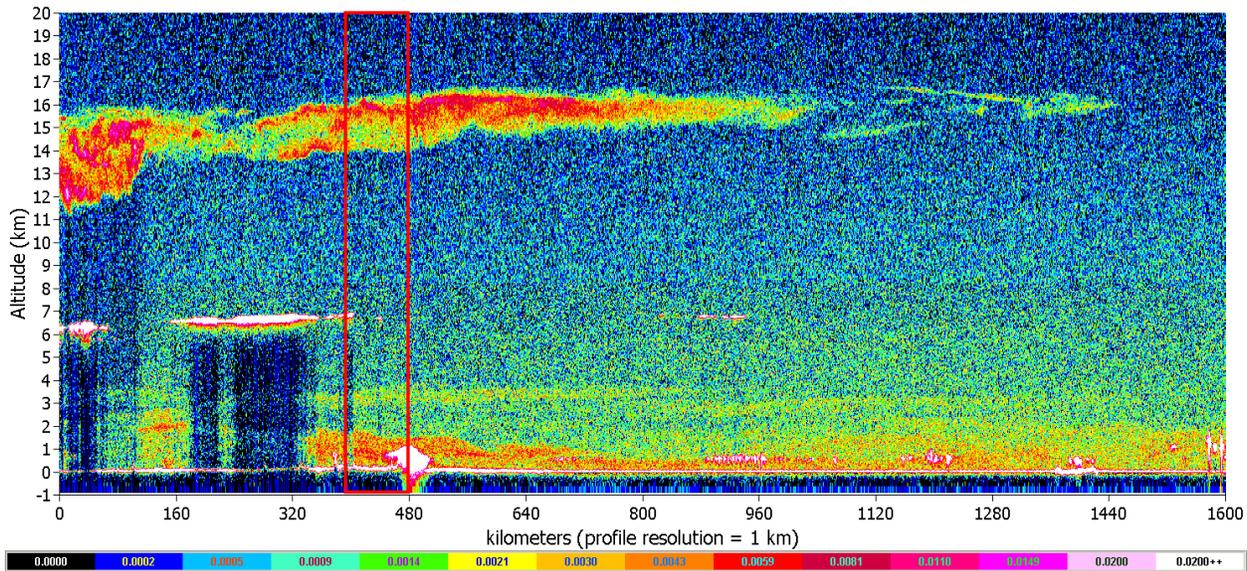


Figure 2.1: A CALIPSO “curtain file” image showing synthetic CALIPSO data derived from measurements acquired by the Lidar In-space Technology Experiment (LITE; Winker et al., 1996) and degraded to the CALIPSO horizontal and vertical sampling resolutions using the CALIPSO lidar simulator (Powell et al., 2002). The red rectangle extended horizontally from 400 km to 480 km demarcates a single 80 km CALIPSO “scene”.

An overarching view of the Level 2 algorithms and their relationship to one another is shown in Figure 2.2. Level 2 lidar processing begins with SIBYL operating on a sequence of scenes, which consist of segments of Level 1 data covering 80 km in along-track distance, corresponding to a collection of 240 consecutive, single-shot profiles. The module averages these profiles to horizontal resolutions of 5, 20, and 80 km, consisting of averages of 15, 60, and 240 profiles, respectively, and detects features at each of these resolutions. The SCA module then classifies these generic features as clouds or one of five aerosol types, based primarily on scattering strength and the spectral dependence of the lidar attenuated backscatter [Liu et al., 2004]. The module can further discriminate between ice/water clouds by relying on the layer-averaged lidar depolarization ratio and ancillary information such as altitude and temperature (Hu et al., 2001). The SCA module also uses a combination of observed parameters and *a priori* information to select appropriate values for the initial lidar ratios and multiple scattering factors required for retrieving extinction and optical thickness by HERA.

The HERA module that is described herein is a hybrid collection of procedures that have been combined to form a single, fully automated scene process responsible for retrieving profiles of optical properties from the features identified by SIBYL. The constituent algorithms include a routine that recreates the averaged profiles originally created by SIBYL’s layer detection scheme. These reconstituted profiles are subsequently used by both the layer classification algorithms and the extinction retrieval algorithms. When permitted by the structure of the signal profile, SIBYL makes an initial estimate of the 532-nm attenuation through a feature, and uses this value to correct the underlying signal for transmission losses. HERA, on the other hand, calculates and corrects for the attenuation within all features in order to produce profiles of extinction and backscatter at both 532 nm and 1064 nm. HERA corrects for attenuation effects beneath all features, including those where SIBYL is unable to make an accurate assessment of the feature transmittance (e.g. where the SNR is too low).

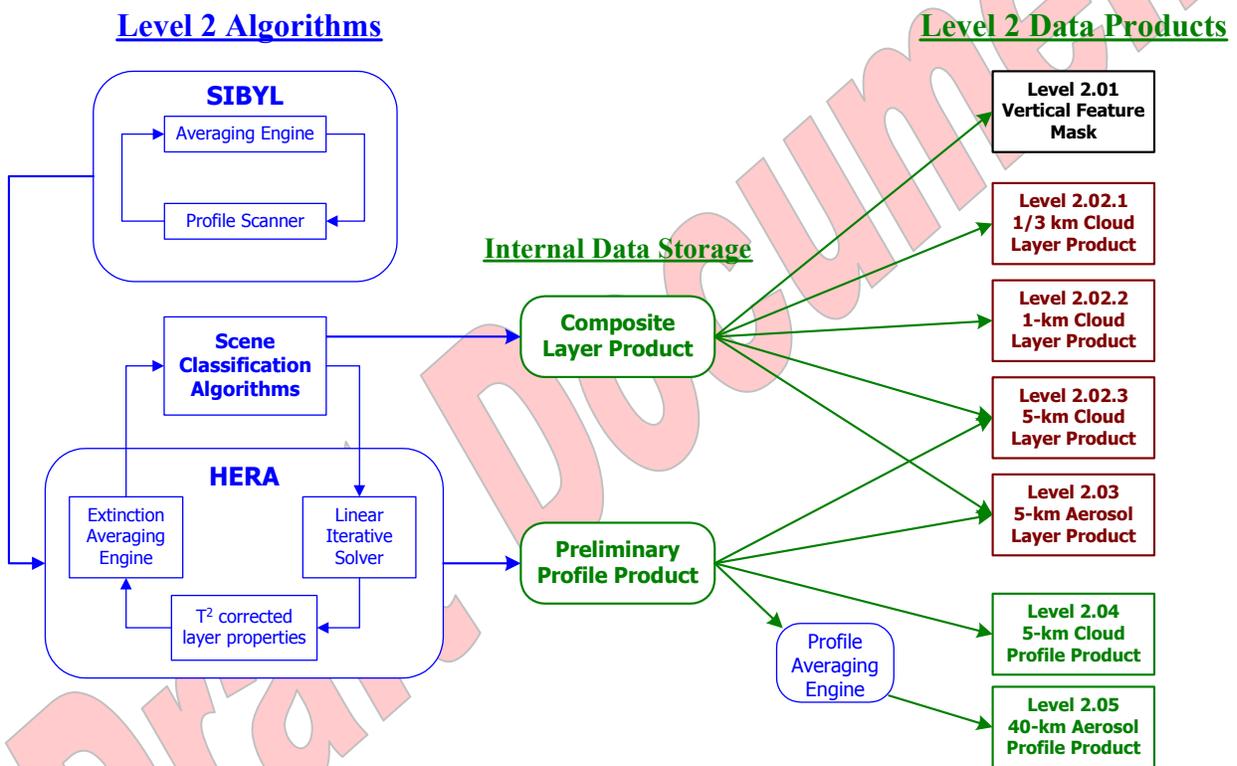


Figure 2.2: The production pathway for the CALIOP Level 2 data products.

The output from the CALIOP Level 2 processing system, shown schematically in Figure 2.3, falls into several major types of products: the vertical feature mask, layer products at varying resolutions, and profile products for clouds and aerosols. The layer products and profile products are reported at uniform spatial resolutions, which are derived via averaging and/or interpolation from the varying resolutions used in generating the intermediate products. Horizontal resolutions of the different products are identified in the figure. A detailed description of SIBYL is presented in the CALIPSO Feature Detection and Layer Properties ATBD (PC-SCI-202 Part 2); a description of SCA can be found in the Scene Classification Algorithms ATBD (PC-SCI-202 Part 3).

The HERA module has several preferred, or default, processing decision paths. For CALIPSO's downlooking viewing geometry, the probability of encountering significant particulate scattering increases with increased penetration distance, and hence a forward solution direction is initiated by default. When possible, the accuracy of the retrieval is maximized by constraining the retrieved optical thickness to agree with measurements provided by SIBYL of the reduction in the clear-air signal below a vertically isolated feature. The constraint is achieved by automatic, iterative adjustment of the lidar ratio in the case of features or, in the case of "feature-free" or apparently clear regions, of the factor we call here the "renormalization" factor. (The terms "normalization" and "calibration" have been used historically in the lidar community to refer to the process of relating the raw, measured lidar signal to the calculated total atmospheric backscatter at some range from the lidar. In the calibrated data used in the CALIPSO Level 2 analyses, this factor is the product of the transmittances of the layers and regions between the lidar and the range of interest. This is discussed in greater depth in Section 2.3.) There are situations where these preferred processing paths are not possible, and in these cases the algorithms select the analysis settings that seek to give the best result for the particular situation encountered.

## 2.2. Analysis of CALIPSO Data – Scene and Profile Processes

The fundamental algorithms that operate within HERA can be considered as profile processes and are rooted in conventional lidar data analysis techniques. They take as inputs either a single profile of lidar signals, or, in the case of the scene classification routines (e.g., ice/water phase), a collection of integrated quantities derived from some segment of a profile. Similarly, the outputs from these profile processes are either a derived profile of optical parameters (e.g., particulate extinction coefficients), the integral of such a profile (e.g., optical thickness), or some additional information about the content and/or structure of the input profile (e.g., as provided by base and top altitudes). As noted previously, prior to execution of HERA, SIBYL has generated a comprehensive spatial analysis that identifies all features present within each CALIOP scene. Since the application of a 'scene concept' is novel, the following discussion provides an illustrative example of the SIBYL output product and how HERA operates on it.

A representative lidar scene obtained from the SIBYL module is depicted in Figure 2.3. It consists of sixteen 5 km horizontal resolution profiles generated from consecutive blocks of fifteen single-shot profiles. Since this example is provided for illustrative purposes, the altitude bins are restricted to heights below 18 km and their thickness is provided at 0.5 km. For CALIOP there are actually 583 altitude bins extending from approximately 2 km below the surface to about 40 km above.

The feature objects in Figure 2.3 are numbered consecutively using a top down, left to right scheme and are color coded according to the amount of averaging required for their detection. Red, yellow, and green blocks signify along-track averaging distances of 5, 20, and 80 km, respectively. Depending upon their relationship to one another, features are further denoted as being either simple, complex, or embedded. Regions devoid of features are identified as composed of clear-air or background conditions. Definitions of these features and regions are as follows:

*Simple features* contain only one type of particulate scatterer. For some vertically isolated, simple features, the transmittance (e.g., optical thickness) can be measured from the reduction in

the clear-air signal below the feature and can be used to constrain the retrieval of the extinction profile. As there is only one type of scatterer in a simple feature and, therefore, one lidar ratio, the constraint is achieved by varying this ratio until the measured and retrieved optical thicknesses agree. In Figure 2.3, features F2 – F7 and F11 are examples of simple, vertically isolated features.

|                                 |    | 5 km | 10 km | 15 km | 20 km | 25 km | 30 km | 35 km | 40 km | 45 km | 50 km | 55 km | 60 km | 65 km | 70 km | 75 km | 80 km |
|---------------------------------|----|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| <b>R<sub>0</sub> (18.0 km)</b>  |    | A    | B     | C     | D     | E     | F     | G     | H     | I     | J     | K     | L     | M     | N     | O     | P     |
| <b>R<sub>1</sub> (17.5 km)</b>  | 1  |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>2</sub> (17.0 km)</b>  | 2  |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>3</sub> (16.5 km)</b>  | 3  |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>4</sub> (16.0 km)</b>  | 4  |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>5</sub> (15.5 km)</b>  | 5  | F2   | F3    | F4    | F5    | F6    |       |       |       |       |       | F7    |       |       | F8    |       |       |
| <b>R<sub>6</sub> (15.0 km)</b>  | 6  |      |       |       |       |       |       |       |       |       |       |       |       |       |       | F9    |       |
| <b>R<sub>7</sub> (14.5 km)</b>  | 7  |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       | F10   |
| <b>R<sub>8</sub> (14.0 km)</b>  | 8  |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>9</sub> (13.5 km)</b>  | 9  |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>10</sub> (13.0 km)</b> | 10 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>11</sub> (12.5 km)</b> | 11 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>12</sub> (12.0 km)</b> | 12 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>13</sub> (11.5 km)</b> | 13 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>14</sub> (11.0 km)</b> | 14 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>15</sub> (10.5 km)</b> | 15 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>16</sub> (10.0 km)</b> | 16 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>17</sub> (9.5 km)</b>  | 17 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>18</sub> (9.0 km)</b>  | 18 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>19</sub> (8.5 km)</b>  | 19 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>20</sub> (8.0 km)</b>  | 20 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>21</sub> (7.5 km)</b>  | 21 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>22</sub> (7.0 km)</b>  | 22 |      |       |       | F12   |       |       |       |       | F11   |       |       |       |       |       |       |       |
| <b>R<sub>23</sub> (6.5 km)</b>  | 23 | F14  | F15   |       |       |       | F16   |       |       |       | F17   | F18   |       |       |       | F12   |       |
| <b>R<sub>24</sub> (6.0 km)</b>  | 24 |      |       | F19   |       |       | F20   |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>25</sub> (5.5 km)</b>  | 25 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>26</sub> (5.0 km)</b>  | 26 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>27</sub> (4.5 km)</b>  | 27 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>28</sub> (4.0 km)</b>  | 28 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>29</sub> (3.5 km)</b>  | 29 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>30</sub> (3.0 km)</b>  | 30 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>31</sub> (2.5 km)</b>  | 31 |      |       |       |       |       |       |       | F21   |       |       |       |       |       |       |       |       |
| <b>R<sub>32</sub> (2.0 km)</b>  | 32 |      |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>33</sub> (1.5 km)</b>  | 33 |      |       |       |       |       |       | F22   |       |       |       |       |       |       |       | F23   |       |
| <b>R<sub>34</sub> (1.0 km)</b>  | 34 |      |       |       |       | F27   |       |       |       |       |       |       |       |       |       |       |       |
| <b>R<sub>35</sub> (0.5 km)</b>  | 35 |      | F24   |       |       |       | F28   | F29   | F30   | F31   | F32   | F33   | F34   | F35   | F36   | F37   | F38   |
| <b>R<sub>36</sub> (0.0 km)</b>  | 36 | SFC  |       | F25   | F26   |       |       |       |       |       |       |       |       |       |       |       |       |

Figure 2.3: A simplified representation of an 80 km CALIPSO scene with feature objects identified by SIBYL. Individual features are labeled sequentially from F1- F38 beginning from the top of the image. Additional details are provided in the text.

*Complex features* are identified as being physically or optically inhomogeneous and may contain different types of particulate scatterers. Effectively they are composed of several simple features that are vertically adjacent with other features, or contain embedded features. (Vertically adjacent features have spatial boundaries that are contiguous in the along-track direction.) There are three complex features in Figure 2.3. The first is composed of the vertically adjacent features F1, F8, F9, and F10. The second complex feature occupies altitude bins F22 - F24 and extends across the whole 80 km block, as does the third complex (boundary-layer) feature. Note that F11 is not included in the second complex feature as it is not vertically adjacent to another feature, nor does it contain an embedded feature.

*Embedded features* are completely contained within the boundaries of a feature of coarser horizontal resolution. They share no common vertical boundary with the coarser feature. Features F17, F18, and F20 are examples of embedded features.

*Clear-air or background regions* are, by definition, regions with no detectable aerosol or cloud features. They are used to identify clear air regions below optically thin features and enable a direct calculation of transmittance or provide a constraint to the extinction retrieval. The boundaries of these regions are defined with respect to the presence of the surrounding features.

The ‘scene’ input to the HERA module consists of a block of data containing sixteen profiles of attenuated backscatter averaged by SIBYL to 5 m horizontal resolution, along with information on the location of all features detected at 5 km, 20 km and/or 80 km, and other feature properties including values of the multiple scattering correction function,  $\eta(z)$ , and the lidar ratio,  $S$ , selected by the SCA. The extinction retrieval works from the greatest altitude at which data are recorded down to the surface (i.e. from smallest to greatest range from the lidar). HERA processes features in order of decreasing altitude, regardless of the horizontal resolution at which they were detected. In principle, the “clear air” regions between the features can also be processed using assumed values of the lidar ratio appropriate for background aerosols. These supposedly clear regions may contain aerosol layers that are below the detection threshold of SIBYL. Faint aerosol layers having high lidar ratios may go undetected, yet still contain a small but perhaps significant fraction of the column aerosol optical thickness. However, low SNR presents a formidable barrier to the practical implementation of a fully automated extinction solution in arbitrarily sized regions of clear air. There are three main obstacles to surmount. First, errors in the derived optical depths are nonlinearly related to the SNR of the profile being analyzed, and even when solved with the correct lidar ratio, low SNR profiles can generate very large overestimates of the layer optical depth. Because errors in the optical depth estimates for overlying layers are necessarily propagated into the layers below, large errors in overlying regions of low SNR will disproportionately distort subsequent results derived in underlying, high SNR regions. Second, the spatial distribution of clouds and aerosol layers in the Earth’s atmosphere limits the available opportunities for obtaining averages of clear air computed over extended horizontal and vertical ranges. The third obstacle is closely related to the second, and arises from the technological and engineering constraints imposed on the design and construction of space-based lidars. When compared to their ground-based counterparts, these instruments are inherently photon starved, and hence must average much larger amounts of data to obtain a similar “clear air” SNR. In light of these concerns, the current implementation of HERA has been constructed so that solutions in “clear air” can be disabled. When disabled, HERA does not attempt to retrieve extinction solutions in the clear air regions, and the aerosol optical depths for each of these regions are set to zero. In this document, we present a complete description of the CALIOP hybrid extinction retrieval algorithm, and thus solutions in the clear regions are described as being an integral part of the analysis for any scene. However, the reader is warned that in this initial release of the CALIOP extinction data products, **clear air solutions are disabled**.

The first step for the HERA module is to average clear regions in the scene (columns A to P) to produce a single profile from the top altitude down to the highest base of all the unsolved features that are in contact with a clear region in at least one column along their bases. (Note that this rule defines the altitude range to be *searched* for a clear region. If a feature at a lower

altitude occupies all sixteen columns, then the clear region cannot extend to its base). In the present example, the base of the clear region 1 corresponds to row 6, which contains the bases of features F1 to F7 and is delineated by the blue line across columns F-H. The reason for the selection of row 6 as the lower limit of the averaged clear region 1 is that the clear air signal viewed across the width of the scene at lower altitudes (clear region 2) is attenuated in at least one column by an overlying feature. Analysis of these clear regions cannot be performed until the attenuation of the overlying features is known. Particulate extinction is retrieved from the averaged attenuated backscatter profile and the square of the particulate transmittance is calculated over its vertical extent. To correct for attenuation by undetected overlying aerosols in clear region 1, the input attenuated backscatter profiles in all sixteen columns are divided by the square of the overlying transmittance in each column. Through this correction process, the transmittance correction that results from that part of the clear region that occupies any column is applied to all rows in that column from the base of the clear region in that column down to the surface.

The complex feature containing F1 and F8 – F10 is analyzed next, using the method described below for complex features. Underlying regions of columns M – P are now corrected (rescaled) for the attenuation caused by these features in the same way as all columns were previously corrected for the overlying background aerosol attenuation. Note that each column is corrected by a different amount depending on how much each of the features contributes to the attenuation.

Any clear region between the last retrieved clear altitude and the next highest feature is solved next, and the underlying columns rescaled as before. In the present example, there is no unprocessed clear region overlying the next highest feature, and each of the simple features F2 to F7 is solved and the underlying regions (level 7 down to the surface) in columns A – E and I – L rescaled.

The next step is to create a clear profile from level 7 down to the top of F12. As each range step has differing numbers of columns contributing the average profile, and as each of these has suffered different degrees of attenuation by overlying features that results in different SNRs in different columns, a mean profile is calculated that accounts for the different numbers of contributing signals at each altitude. Once this is solved, and all 16 columns rescaled from the base of clear region 2 downwards, the simple feature 11 is solved. Then follows the retrieval of all the features in the middle complex feature, the clear region below it and finally the complex boundary layer feature.

#### *Analysis of Simple Features*

Wherever possible, simple, elevated, vertically isolated features (e.g. F2 and F11) are solved using a retrieval that is constrained by the feature transmittance determined from the ratio of the clear-air signal above and below the feature. However, if the SNR of the feature is too low, the feature is totally attenuating, or the feature is in contact with the surface, then the transmittance of the feature cannot be determined independently and a constrained solution is not possible. An unconstrained retrieval is employed in this situation using appropriate tabulated values for some of the unknown parameters ( $S$  and  $\eta(z)$  defined above).

#### *Analysis of Complex Features*

Complex features, by definition, contain more unknowns (lidar ratios) than constraints (optical thickness or transmittance measurements), so a unique, constrained solution is not possible. Instead, HERA seeks a solution (a set of  $S$  values) that yields particulate optical thickness,  $\tau$ ,

averaged over the full vertical and horizontal extent of the feature and is consistent with that determined by an estimate of the two-way transmittance obtained from the clear air beneath the feature averaged over the same horizontal extent.

Consider the complex feature on the top right of Figure 2.3. First, the averaged profile representing F1 is calculated from columns M – P. This profile is solved over its whole vertical extent using the lidar ratio supplied by the SCA module. Each underlying, adjacent feature is temporarily rescaled by the transmittance of that proportion of F1 that overlies it, so F8 – F10 are all rescaled differently. F8 – F10 are then solved individually, using lidar ratios supplied by the SCA routine. The average particulate  $\tau$  is calculated from the retrieved extinction in columns M – P and compared with the value determined from the ratio of the clear-air signals above and below the complex feature (e.g., from rows 1 – 3 above and 10 – 21 below). If the values do not agree, then the lidar ratio of the feature with the largest integrated attenuated backscatter,  $\gamma'$ , is adjusted and the whole complex feature solved again. The lidar ratio of any feature is adjusted within previously defined, acceptable limits. If further adjustment is needed, then the lidar ratios of the other features are adjusted in order of decreasing  $\gamma'$  until consistency is reached between the measured and calculated values of  $\tau$ . Note that this procedure ensures consistency between the measured and calculated values of optical thickness. It cannot guarantee that each extinction profile is the correct one as the problem is underdetermined.

Embedded features like F20 pose an extra degree of difficulty. F20 cannot be solved until the feature in which it is embedded, F16, is solved. However, F20 cannot be solved properly until the attenuation of the part that lies beneath F20 is corrected for the attenuation of F20. An iterative solution is employed whereby the attenuation by F20 is initially set to zero allowing for the solution of F16. F20 can then be corrected for the attenuation caused by the part of F16 that overlies it (provided that the lidar ratios are correct). The part of F16 below F20 can then be corrected for the attenuation of F20 just calculated, the average attenuated backscatter profile for F16 recalculated (now correctly) and the correct extinction profile for F16 retrieved. Thus, the attenuation by F20 is improved iteratively until the retrieval of F16 and F20 converges.

The simplified descriptions of the retrieval of extinction and backscatter in a CALIPSO scene given above are intended to outline the relationship between the various scene and profile processes in the CALIPSO Level 2 analysis. The definition of a scene as a combination of features and feature-free regions of various scales, the identification of the types of the various features and the assignment of analysis parameters, the creation of average profiles of attenuated backscatter from signals in various regions of the scene and, finally, the analysis of these average profiles are the essential components of the multi-pass, multi-scale analysis. In the following sections, the mathematical basis for the analysis of the average profiles is given and this is followed by a more-detailed description of the steps involved in the calculation and analysis of simple, complex and embedded features. This latter section is supported by detailed flow charts.

### 2.3. Retrieval of Particulate Extinction Profiles

There are a number of effective methods for deriving particulate extinction and backscatter coefficients from the signals measured by elastic-backscatter lidars. For CALIOP, effective approaches need to consider situations where the particulate attenuation is usually significant and variable. Some early retrieval algorithms attempted to account for attenuation using simple, iterative techniques (e.g. Elterman, 1966, Gambling and Bartusek, 1972a, b), in which the uncalibrated, signal profiles were normalized to produce profiles of attenuated backscatter by

scaling the measured signal to the modeled, molecular signal at a distant point high in the atmosphere where particulate scattering was assumed to be negligible. As the focus of that work was often elevated aerosol layers with backscattering comparable to the ambient molecular values, the authors necessarily used two-component solutions in which the scattering effects of both particles and molecules were considered along with molecular (ozone) absorption. Closed-form, analytical solutions were also developed (e.g. Barrett and Ben-Dov, 1967; Vizee et al., 1969; Davis, 1969; Fernald et al., 1972; Klett, 1981; Fernald, 1984, following the rainfall radar analyses of Hitschfeld and Bordan, 1954).

While all these retrieval algorithms were developed in the context of single scattering, others were further developed to account for multiple scattering situations (e.g. Platt, 1973; Sassen and Cho, 1992) by the introduction of a multiple scattering efficiency factor that modified the extinction at any range. However, as explained in Section 2.3.1.2, below, and Section 3, the CALIOP analysis framework adopts a range-dependent multiple scattering factor,  $\eta(r)$ , which modifies the optical depth (rather than extinction) and is a function of penetration depth into the layer. This parameterization dictates the use of an iterative retrieval scheme that incorporates profiles of  $\eta(r)$  that are specific to the particular feature types being analyzed.

### 2.3.1. Important Considerations in the Solution of the Lidar Equation

#### 2.3.1.1. The elastic backscatter lidar equation

For an elastic backscatter lidar system as used in CALIPSO, the retrieval of profiles of particulate backscatter and extinction involves the solution of the two-component lidar equation:

$$P(r) = \frac{1}{r^2} E_0 \xi [\beta_M(r) + \beta_P(r)] T_M^2(0, r) T_{O_3}^2(0, r) T_P^2(0, r). \quad (2.1)$$

Here

$P(r)$  is the detected backscattered signal from range  $r$  from the lidar;

$\xi$  is the lidar system parameter (see CALIOP ATBD Part 2, section 3.1.2), where  $\xi = G_A C$ , and  $G_A$  is the amplifier gain and  $C$  is the lidar calibration factor;

$E_0$  is the average laser energy for the single shot or composite profile;

$\beta_M(r)$  is the molecular volume backscatter coefficient, proportional to the molecular number density profile.

$$T_M^2(0, r) = \exp \left[ -2 \int_0^r \sigma_M(r') dr' \right]. \quad (2.2)$$

is the profile of the two-way transmittance between the lidar and range  $r$ ;

$\sigma_M(r) = S_M \beta_M(r)$  is the molecular volume extinction coefficient, where

$S_M$  is the molecular extinction-to-backscatter (or lidar) ratio.

$T_{O_3}^2(0, r) = \exp \left[ -2 \int_0^r \alpha_{O_3}(r') dr' \right]$  is the two-way ozone transmittance, where

$\alpha_{O_3}(r')$  is the ozone absorption coefficient,

$\beta_p(r)$  is the particulate volume backscatter coefficient,

$$T_p^2(0, r) = \exp[-2\eta(r)\tau_p(0, r)] \quad (2.3)$$

is the particulate two-way transmittance, and

$$\tau_p(0, r) = \int_0^r \sigma_p(r') dr' = S_p \int_0^r \beta_p(r') dr' \quad (2.4)$$

is the particulate optical thickness (thickness).

$$\sigma_p(r) = S_p \beta_p(r) \quad (2.5)$$

is the particulate volume extinction coefficient, where

$S_p$  is the particulate extinction-to-backscatter (or lidar) ratio and

$\eta(r)$  is a multiple scattering factor as described in section 3.

After detection, the signal is amplified electronically then digitized in a dual-range digitizer. These processes have their associated gains and offsets. The methods by which they are calculated and their effects removed are described in the CALIOP Lidar Level 1 ATBD (PC-SCI-201) and will not be discussed further here. However, the errors and uncertainties in the determination of these quantities can affect the quality of the retrieval of  $\beta_p(r)$  from equation.

It can be seen that the lidar equation cannot be solved without ancillary information. For operations at the CALIPSO wavelengths (532 nm and 1064 nm), the molecular number density and ozone absorption coefficient profiles must be known. For CALIPSO, these are obtained from meteorological analyses produced by NASA's Global Modeling and Assimilation Office (GMAO). Furthermore, four additional quantities need to be determined. These are the zero offset, which is measured and removed from the signal on board the satellite; the calibration ( $C$ ), which is calculated as part of CALIOP Level 1B processing; the particulate lidar ratio ( $S_p$ ); and the multiple scattering factor profile  $\eta(r)$ . Although the estimation of these quantities is not part of the extinction analysis, any errors or uncertainties in them lead to errors and uncertainties in the retrieved extinction products.

### 2.3.1.2. Extinction retrieval formula and algorithms

The algorithms and formulae that have been used to solve the lidar equation fall broadly into two classes. The iterative methods that were developed for the analysis of searchlight data on aerosols in the upper atmosphere (e.g. Elterman, 1966) were adopted for lidar studies, often of the same features (e.g. Gambling *et al.*, 1970; Gambling and Bartusek, 1972a, b; Platt, 1973). Analytical formulae developed for the analysis of rainfall radar data (e.g. Hitschfeld and Bordan, 1954) and based on the solution to the Bernoulli equation were also adopted by lidar researchers (e.g. Barrett and Ben-Dov, 1967; Viezee *et al.*, 1969; Davis, 1969; Klett, 1981). These, so-called, single-component solutions only consider one atmospheric scattering component and are not applicable to situations where molecular and particulate scattering might be comparable in magnitude. Two-component analytical solutions were developed for these situations (e.g. Fernald *et al.*, 1972; Fernald, 1984).

A two-component solution that incorporated Platt's (1973, 1979)  $\eta(r)$  parameterization of multiple scattering in an analytical formula was developed by Sassen and Cho (1992). Although  $\eta(r)$  varies with range, it is usually replaced with layer-effective value,  $\bar{\eta}$ , that is assumed to remain constant throughout the layer. However, the approach used in the CALIPSO analysis differs in two respects. In the CALIPSO analysis framework,  $\eta(r)$  modifies the cumulative particulate optical thickness of a feature (equations (2.3) and (2.4)), rather than the extinction coefficient at a particular range as in the earlier papers. This influences the selection of the retrieval algorithm for HERA. The iterative method can easily be adapted to incorporate the CALIPSO multiple scattering parameterization, and has been selected for deriving profiles of volume backscatter and extinction coefficients from both clouds and aerosols (Young et al., 2003). Secondly, because  $\eta(r)$  is derived from Monte Carlo computations of multiply-scattered lidar signals specific to each CALIOP aerosol and cloud subtype, it is known as a function of range and this range-dependent function can be used instead of an averaged value.

### 2.3.1.3. Stability of retrievals and constrained and unconstrained solutions

It has been known for some time that solutions to the lidar equation using a near-field boundary condition can be unstable (e.g. Vizee *et al.*, 1969). This instability exists in both the analytical and iterative forward solutions. Klett (1981) showed that stability could be achieved by using a far-field boundary condition, setting the total extinction in the far-field to some assumed or known value. (This actually had been the usual practice in the analysis of stratospheric lidar data, although using total backscatter instead of extinction, largely because of the possibility of using an accurate calibration or boundary value in an aerosol-free region of the atmosphere above the stratosphere.) However, if a solution can be constrained by a measurement of optical thickness over some region, then stable solutions can be obtained in either the forward or backward directions as pointed out by Young (1995). This constraint is achieved by setting the lidar ratio so that the retrieved optical thickness matches the measured value.

For the CALIPSO lidar data analysis, solutions are constrained by measurements of optical thickness made by SIBYL using the transmittance method where these are available. HERA has the option of choosing either forward or backward solutions as appropriate for the profile being processed.

### 2.3.2. Mathematical Basis for the CALIPSO Extinction Solver

The HERA module processes the sixteen 5 km horizontal resolution profiles in an 80 km scene, working from the top altitude down to the surface, or to the altitude of the last valid atmospheric signal. Using the locations of the features identified by SIBYL, the HERA module creates averaged profiles in the scene and produces retrieved profiles of particulate backscatter and extinction and layer-averaged values of these quantities.

#### *Modification of the attenuated backscatter coefficients for use in extinction processing*

During Level 1B processing of the data, a lidar calibration factor is calculated (see CALIOP ATBD Part 2). Next, the various numbers of profiles are averaged and the resulting signal processed to extract information on any detected features. Following the symbols defined in Section 2.3.1.1, the sixteen profiles supplied to HERA by SIBYL for each 80 km block of data of the attenuated backscatter signal are corrected for ozone transmittance:

$$\beta'(0, r) = [\beta_M(r) + \beta_P(r)] T_M^2(0, r) T_P^2(0, r) = \frac{P(r)r^2}{E_0 G_A C T_{O_3}^2}. \quad (2.6)$$

Note that, due to the ozone correction, this definition of  $\beta'$  differs slightly from that in the other parts of the CALIOP ATBD.

The uncertainty in this quantity,  $\Delta\beta'$ , is supplied, along with profiles of the multiple scattering factor and the associated uncertainties. The molecular backscatter and transmittance profiles, ozone transmittance profiles and all their uncertainties are provided by the Meteorological Manager Module (Met Manager). Statistics of each layer detected by SIBYL, including the altitude of the base and top, optical thickness and lidar ratio, and their uncertainties are also supplied.

For any region, the retrieval can be either constrained or unconstrained. A constrained retrieval is one where the optical thickness calculated by integrating the retrieved particulate extinction coefficient over the retrieval interval is adjusted to match some independently measured value. The constraint available within HERA is the particulate optical thickness,  $\tau_P$ , determined from the reduction in the signal from the clear air below the feature. For features, the particulate lidar ratio,  $S_P$ , is adjusted until the retrieved  $\tau_P$  matches that measured by the SIBYL module, either within the combined uncertainties of the measured and calculated values, or to within a predefined default tolerance. Although it is possible to calculate a value of the required lidar ratio using analytical formulae, various factors, including noise on the data, the specification of  $\eta(r)$  as a function of range and an imperfect knowledge by the SCA routine of the attenuation by overlying regions, it has been found that  $S_P$  is best refined by iteration. This only takes a small number of iteration cycles. However, as explained in Section 2.2, retrievals in features are only iterated when the feature is simple, that is, an elevated, vertically isolated 532-nm feature with no other embedded features.

It has been found in testing the algorithm on simulated data, that retrievals in many feature-free regions can diverge in the positive or negative directions largely because of the low SNR in the region used for initiating the retrieval. This can occur, for example, if the retrieval is initiated in the stratosphere near 30 km altitude where the signal is weak, or in apparently clear regions below detected features of moderate attenuation. To attempt to correct a diverging solution in these regions of very low particulate optical thickness by adjusting the lidar ratio is a poor strategy and can lead to widely varying, physically unrealistic and even negative lidar ratios. As the principal cause of the divergence in these feature-free regions is an incorrect boundary (i.e., the “normalization factor”) value resulting from either a low SNR in the normalization region or an incorrect evaluation of the attenuation of overlying regions, the boundary value is adjusted instead of the lidar ratio. In the case of a constrained retrieval, the normalization factor is adjusted until the retrieved and specified particulate optical thicknesses agree within calculated uncertainties. However, where there is no available constraint and an unconstrained retrieval is specified, then a slightly different approach is adopted. In these situations, the normalization factor is adjusted so that the retrieved optical thickness is greater than zero, but less than the minimum value that could be detected by the feature finder. It is reasoned that regions having particulate optical thicknesses greater than this value would be identified as features and not as clear regions. If the region in which the normalization factor needs adjustment is the top, stratospheric, region, then the renormalization adjustment is interpreted as being the result of a

statistical variation in the calibration factor. (The averaged profile used in the analysis of clear regions may have been calculated over a different horizontal extent from that used to calculate the calibration factor supplied for the whole 80 km block.) In this situation, all underlying data are rescaled by this calibration adjustment.

Once the section of the averaged profile is ready for analysis, various retrieval parameters are determined by the HERA module based on settings in a run-time file and characteristics of the data. These parameters include the minimum and maximum ranges, the retrieval direction, the  $\tau_P$  constraint (when available), the normalization range and normalization factor and their uncertainties where appropriate.

Before an averaged profile can be analyzed, it must be rescaled by a factor that accounts for the molecular and particulate transmittance losses down to the altitude at which the retrieval is to be initiated. For a forward retrieval, this rescaling (effectively a renormalization) occurs at the top of the profile and the renormalization factor is simply:

$$C_N(r_t) = T_M^2(0, r_t) T_P^2(0, r_t), \quad (2.7)$$

with relative uncertainty

$$\frac{\Delta C_N(r_t)}{C_N(r_t)} = \left[ \left( \frac{\Delta T_M^2(0, r_t)}{T_M^2(0, r_t)} \right)^2 + \left( \frac{\Delta T_P^2(0, r_t)}{T_P^2(0, r_t)} \right)^2 \right]^{1/2}, \quad (2.8)$$

where  $r_t$  is the range to the top of the feature or region being analyzed. Although the two-way particulate transmittance is included here in equation (2.7), this is only to illustrate how an incorrect estimate of the overlying particulate transmittance losses can bias the retrieval. As it is written, equation (2.7) describes the renormalization at the first point at the top of the atmosphere (near 40 km). In practice, the renormalization factor used when analyzing lower regions does not include the particulate transmittance. As explained below, because any scene can be complicated by the presence of complex and adjacent features, it is more convenient, after a feature has been processed, to then rescale underlying attenuated backscatter data by the retrieved particulate two-way transmittance. This is preferable to including this factor in the renormalization factor.

If a backward retrieval is to be performed then the renormalization factor must include the particulate two-way layer transmittance loss,  $T_L^2$ , through the feature:

$$C_N(r_b) = T_M^2(0, r_b) T_L^2, \quad (2.9)$$

with relative uncertainty

$$\frac{\Delta C_N(r_b)}{C_N(r_b)} = \left[ \left( \frac{\Delta T_M^2(0, r_b)}{T_M^2(0, r_b)} \right)^2 + \left( \frac{\Delta T_L^2}{T_L^2} \right)^2 \right]^{1/2}, \quad (2.10)$$

where  $r_b$  is the range to the base of the feature or region being analyzed. (If a backward retrieval is to be performed in the stratosphere using some independently measured value of stratospheric particulate optical thickness as a constraint on the retrieval, then the layer transmittance  $T_L^2$  in

equations (2.9) and (2.10) must be replaced by the by the two-way transmittance loss at the base of the region to be analyzed,  $T_{Aref}^2(0, r_b)$ .)

Because any error in the renormalization factor will produce a bias in the retrievals of the profiles of  $\beta_P(r)$  and  $\sigma_P(r)$ , the uncertainty  $\Delta C_N$  in equations (2.8) and (2.10) will be included in the estimation of the overall bias error in the retrievals, even though it may be random in origin. Random errors and systematic errors have different effects on the retrieved profiles; the former mainly superimpose “noise” on the result while the latter produce biases. (The non-Gaussian statistics of the photon detection process can lead to biases in some circumstances, but these can be made small – compared with the bias errors considered here – with sufficient averaging.) Therefore, random and systematic errors will be calculated separately and combined in quadrature in the result. In this work, random errors and uncertainties will be signified by the notation  $\Delta x$ , bias errors by  $\varepsilon_b(x)$  and total, combined errors by  $\varepsilon_t(x)$ .

Once the renormalization factor and minimum and maximum ranges are selected by the HERA module, the appropriate section of the profile and accompanying data are sent to the extinction analysis subroutine. In this subroutine, the first step is to rescale the profiles of attenuated backscatter and its uncertainty:

$$\beta'_N(r) = \beta'(0, r) / C_N(r_N), \quad (2.11)$$

$$\Delta\beta'_N(r) = \Delta\beta'(0, r) / C_N(r_N). \quad (2.12)$$

Strictly, the quantity  $\beta'_N(r)$  represents a family of curves and should be written as  $\beta'_N(r_N, r)$  to reflect the dependence on the renormalization range ( $r_N$  in this case) but to simplify the notation, the inclusion of the normalization factor in the normalized attenuated backscatter is signified by the subscript  $N$ .

For the forward analysis direction,  $r_N < r$ , so all the transmittances can be factorized as shown:

$$T^2(0, r) = T^2(0, r_N)T^2(r_N, r). \quad (2.13)$$

Then, using (2.7) and (2.11), the normalized attenuated backscatter coefficient becomes

$$\beta'_N(r) = [\beta_M(r) + \beta_P(r)]T_M^2(r_N, r)T_P^2(r_N, r). \quad (2.14)$$

This equation can be rearranged to give the particulate backscatter at range  $r$ :

$$\beta_P(r) = \frac{\beta'_N(r)}{T_M^2(r_N, r)T_P^2(r_N, r)} - \beta_M(r). \quad (2.15)$$

However, as

$$T_P^2(r_N, r) = \exp\left[-2\eta(r)S_P \int_{r_N}^r \beta_P(z)dz\right] = \exp[-2\eta(r)\tau_P(r_N, r)], \quad (2.16)$$

the particulate transmittance factor in equation (2.15) includes the unknown,  $\beta_P(r)$ , being sought. ( $S_P$  is the lidar ratio for the aerosols in the region being analyzed and  $\tau_P$  is the particulate optical thickness.) We are, therefore, seeking a solution to an equation of the form  $x = F(x)$ . Equations of this form are commonly solved using one of the fixed-point iteration schemes (e.g. Fröberg,

Section 2.5). Some lidar researchers, (e.g. Elterman, 1966; Gambling and Bartusek, 1972 a, b and others) have used an iterative technique (also referred to in the literature as simple, linear, fixed-point, or Picard iteration, or repeated substitution) where, as a first step in the iterative solution,  $\beta_p(r)$  in the particulate transmittance factor equation (2.16) is assigned an initial value of zero. The value of  $\beta_p(r)$  calculated from equation (2.15) with this approximation is then used to refine the particulate transmittance in equation (2.16). This iteration between equations (2.15) and (2.16) is continued until changes in consecutive values of  $\beta_p(r)$  are less than a predetermined tolerance. Once the particulate backscatter has been retrieved, the particulate extinction is obtained by multiplying by the appropriate lidar ratio:

$$\sigma_p(r) = S_p \beta_p(r). \quad (2.17)$$

The solution then proceeds to the next range where the process is repeated.

Where the backward analysis direction is used, a similar procedure is followed. Now however,  $r_N > r$ , so all the transmittance factors can be expressed as

$$T^2(0, r_N) = T^2(0, r) T^2(r, r_N). \quad (2.18)$$

The normalized, attenuated backscatter for the backward direction becomes

$$\beta'_N(r) = \frac{\beta_M(r) + \beta_P(r)}{T_M^2(r, r_N) T_P^2(r, r_N)}. \quad (2.19)$$

Equations (2.15) and (2.16) become, for the backward direction,

$$\beta_P(r) = \beta'_N(r) [T_M^2(r, r_N) T_P^2(r, r_N)] - \beta_M(r), \quad (2.20)$$

and

$$T_P^2(r, r_N) = \exp \left[ -2\eta(r) S_p \int_r^{r_N} \beta_p(z) dz \right]. \quad (2.21)$$

In the HERA analysis of the CALIPSO data, once a region or feature has been processed, the attenuated backscatter data in all regions that underlie it are renormalized by dividing by the  $T_p^2$  of the region just processed. At first it might seem that this  $T_p^2$  should be included in the renormalization factor  $C_N$  (see equation (2.7)). However, by studying again the representative complex scene depicted in Figure 2.3, it is obvious that the value of  $T_p^2$  that overlies a feature is different in each column and differs from feature to feature. (Note, for example, the different features that overlie each column of feature F12 in that figure.) It is, therefore, far more correct (and convenient) to rescale the attenuated backscatter data in those columns, and only those columns, that are beneath the last region analyzed:

$$\beta'_N(r) = \frac{\beta'(r)}{T_P^2(r_t, r_b)}, \quad (2.22)$$

where  $r_b$  and  $r_t$  are, respectively, the ranges to the top and base of the region processed. (The  $\beta'(r)$  here is a general representation of the attenuated backscatter array at the current stage of the

analysis process and may have been rescaled previously.) The random component of the uncertainty in the underlying data points is also increased:

$$\Delta\beta'_N(r) = \Delta\beta'(r) / T_P^2(r_t, r_b). \quad (2.23)$$

It can be seen then that, as the analysis proceeds down through the atmosphere, the uncertainty in the retrievals is increased, as is expected, by the correction for the attenuation of the region just processed. Errors in the attenuation correction can have similar effects to errors in the renormalization factor and produce biased retrievals as will be shown below. This emphasizes the requirement that the initial calibration and in the specification of constraints or lidar ratios be as accurate as possible.

It is possible for errors in the retrieval of features higher in a scene to be transferred down to lower features that are not directly below the upper feature. This will occur if there is an intervening region or feature that occupies columns that are common to both upper and lower features. In the Representative Complex Scene (Figure 2.3) errors in the retrieval of feature F10, for example, will lead to bias errors in the retrieval of feature F24, even though this feature is not directly below F10. The error propagation occurs in the processes of rescaling below the upper feature and of forming the average profiles for the intervening clear regions and features. In this example, the average profiles for clear regions 2 and 3 and features F12 and F21 will be affected by retrieval errors in F10 when the region in column P that underlies it is rescaled by the retrieved two-way transmittance of F10. As F24 is below both clear regions and both F12 and F21, it will also be rescaled after these features are solved and any errors in the retrieved transmittance will be propagated as biases to the analysis of F24. Propagation of bias errors is estimated by calculating the contribution of errors in the retrieved transmittances of upper features to the renormalization factors of lower features or clear regions. In the example at hand, when F10 is solved, the combined systematic and random components of the uncertainty in the transmittance are added to column 16 of an array of bias estimates. The bias error in the renormalization factor for feature F12 (or clear region 2 below F10) is calculated by averaging the bias estimates in all of the columns that overlie it. This bias error estimate is thus propagated to the estimate of the bias error in the transmittance of F12 and the whole complex feature of which it is a part, and thence to clear region 3, F21 and, ultimately, F24.

#### 2.3.2.1. Newton's (Newton-Raphson) Method algorithm

The Hybrid Extinction Retrieval Algorithm used in the analysis of CALIPSO lidar data uses different methods depending on the circumstances. It has been found in tests with simulated data that much faster convergence is achieved in regions of moderate to high particulate extinction using a Newton's Method iterative scheme rather than the simple scheme described above. (Newton's method is also known as the Newton-Raphson method. See, for example, Fröberg, section 2.2). As highly attenuating regions are always analyzed in the forward direction in the extinction algorithm, the Newton-Raphson method is employed exclusively in the forward branch of the algorithm. (The extra complexity involved in calculating the derivative of the function and the extra logical branching associated with testing for zero-valued derivatives led to the decision to retain the simple iteration for backward retrievals where the convergence in the more tenuous regions analyzed is always rapid.)

In the Newton-Raphson method, successive estimates,  $k$ , of the particulate backscatter at any range,  $r$ , from the lidar are obtained from the familiar formula:

$$\beta_{P,k+1}(r) = \beta_{P,k}(r) - \frac{f(\beta_{P,k}(r))}{f'(\beta_{P,k}(r))}. \quad (2.24)$$

In the implementation of the method in HERA,

$$f(\beta_{P,k}(r)) = \frac{\beta'_N(r)}{T_M^2(0,r)} \exp[2\eta(r)\tau_P(r_N,r)] - \beta_M(r) - \beta_{P,k}(r), \quad (2.25)$$

and

$$f'(\beta_{P,k}(r)) = \frac{\eta(r)S_P\delta r\beta'_N(r)}{T_M^2(0,r)} \exp[2\eta(r)\tau_P(r_N,r)] - 1.0, \quad (2.26)$$

is the derivative of  $f(\beta_P(r))$  with respect to  $\beta_P(r)$  at  $\beta_{P,k}(r)$  and  $\delta r$  is the range increment at range  $r$ . The initial value of  $\beta_{P,k}(r)$  at each range step (i.e.  $\beta_{P,0}(r)$ ) is obtained from (2.15). Once convergence of  $\beta_P(r)$  is attained, the particulate transmittance is calculated using (2.16) and the analysis proceeds to the next range step.

### 2.3.2.2. Special considerations in the renormalization of CALIPSO retrievals

Difficulties arise in the renormalization of attenuated backscatter profiles in two areas. First, as there is likely to be negligible measurable molecular scattering at 1064 nm, normalization to a modeled molecular profile at this wavelength is not possible. The second difficulty arises from the fact that the HERA module will be required to analyze averaged feature profiles where another feature of different horizontal extent and lidar ratio is adjacent at the top of the feature to be analyzed (e.g. features 8-10, 14-20 and 22-38 in the Representative Complex Scene in Figure 2.3). It is not possible, in these situations, to renormalize the averaged, attenuated backscatter profile for these features to a molecular model. Both difficulties are overcome by the following means. First, the normalization of these profiles is always performed at the first point (top) in the feature. Second, when the normalization factor is calculated, particular care must be taken to ensure that the transmittance between the adjacent point in the overlying feature and the first point in the current feature,

$$T_P^2(r_N - \delta r, r_N) = \exp[-\eta(r_N - \delta r)\sigma_P(r_N - \delta r)] \cdot \exp[-\eta(r_N)S_P\beta_P(r_N)], \quad (2.27)$$

is calculated correctly and included.

### 2.3.2.3. Lidar ratio adjustment to prevent diverging and negative retrievals

An inspection of equations (2.15) and (2.16) reveals a potential instability in the analysis in the forward direction. If the normalization factor calculated using equation (2.7) is too small or the lidar ratio in equation (2.16) too large, then the attenuated backscatter in equation (2.14) and retrieved backscatter in equation (2.15) will be too large. As the retrieved backscatter is used in calculating the particulate transmittance correction through equation (2.15), this latter value will be too small and will lead to an even larger value of  $\beta_P(r)$ . Divergence can occur, with the solution eventually becoming undefined. Divergence is determined by testing for an increase in the absolute value of the difference between successive estimates of  $\beta_P(r)$ . The situation is corrected by decreasing the value of the lidar ratio and restarting the retrieval from the top of the region being analyzed.

Another form of error can occur where the lidar ratio is too small. This can lead to negative values of retrieved particulate backscatter. This condition is determined by testing if the retrieved values of  $\beta_p(r)$  at a number of consecutive ranges are negative when the corresponding values of the attenuated backscatter are positive. In these cases the lidar ratio is increased.

In situations of moderate to high optical thickness, it is possible that an adjustment of the lidar ratio in one direction in order to overcome one problem may well cause an additional problem. For example, increasing the lidar ratio to prevent negative retrievals may well cause the solution to diverge. In order to overcome this difficulty the following scheme is adopted, in which the adjustments of the lidar ratio up and down are linked. If a change in one direction (a decrease, for example) is required and if there have been fewer than five adjustments in that or in the opposite direction (an increase), then a one percent adjustment (a decrease in this example) is made to the lidar ratio. Once there has also been an adjustment in the opposite direction, the new value of the lidar ratio is chosen as the mean of the last increased value and the previous decreased value. The result is a lidar ratio that is neither too large nor too small and converges quickly to the correct value. As the optical depth increases, the constraint on the allowable values of  $S_p$  becomes tighter. The scheme is illustrated in Figure 2.4.

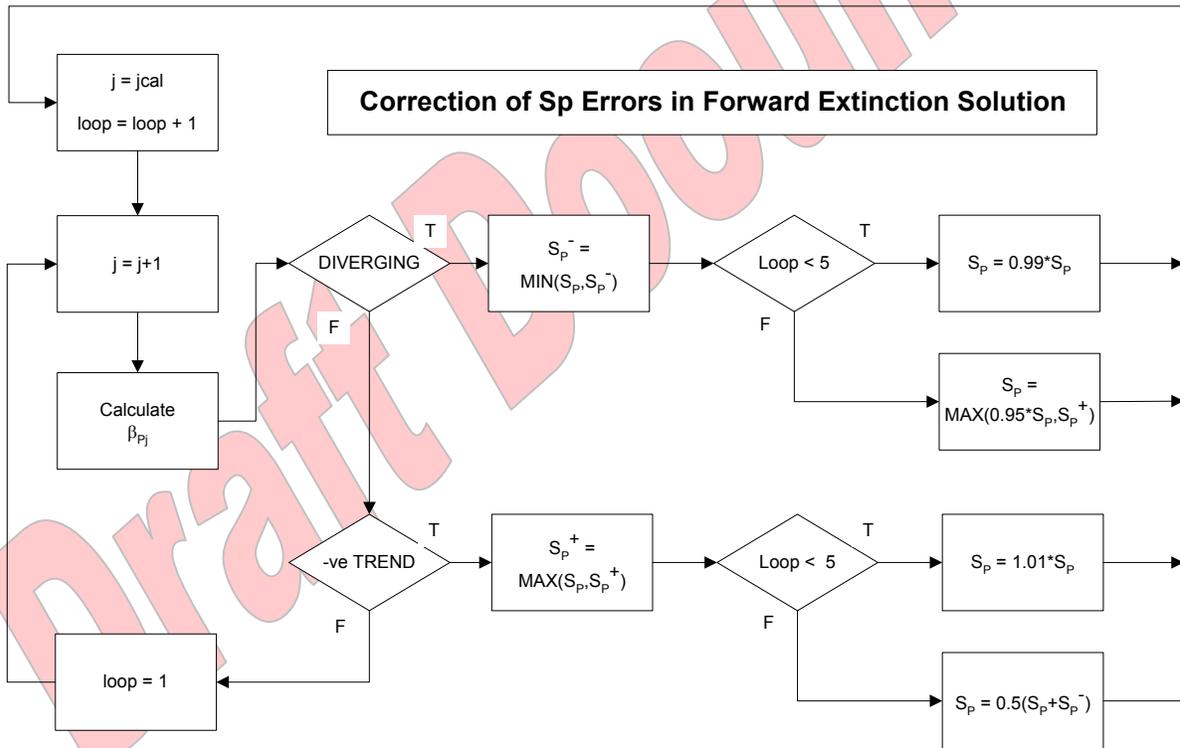


Figure 2.4: Simplified flowchart of the adjustment of the lidar ratio in the CALIPSO extinction solver.  $S^-$  and  $S^+$  are, respectively, the converging least upper and greatest lower bounds to the particulate lidar ratio that is under adjustment. They are initialized to maximum and minimum acceptable lidar ratios as specified in the run-time file. See text for a detailed explanation.

#### 2.3.2.4. Error / Uncertainty analyses

The magnitude of possible errors in the retrieved quantities is estimated by analyzing the propagation of uncertainties in the input variables. The error (or uncertainty) analyses for both the forward and backward directions are virtually identical, as can be deduced by comparing equations (2.15) and (2.16) with (2.20) and (2.21). While the analyses are identical for both forward and backward solutions, because of the quite different magnitudes of the various quantities at near-point and far-point normalization ranges, the growth of the uncertainties with range is quite different. The analysis that follows assumes a retrieval in the forward direction. In the following uncertainty analysis, random and systematic errors will be estimated separately then combined in quadrature in the final result. (The sensitivities of the retrievals of particulate optical thickness and two-way transmittance to a range of systematic errors in profile calibration and lidar ratios are examined in detail in Section 2.3.3.)

##### *Random errors*

We shall now examine the effects of random errors in the attenuated backscatter coefficient profile, the molecular backscatter profile and the molecular, two-way transmittance profile on the estimated random errors in the particulate backscatter, extinction and two-way transmittance. It will be assumed that the random uncertainties in the different quantities are uncorrelated. While this may not always be strictly true, it is expected that any errors incurred in this approximation will be much smaller than the main contributors to the overall uncertainties.

From equation (2.15), the uncertainty in the retrieved particulate backscatter coefficient can be written

$$(\Delta\beta_p(r))^2 = \beta_T^2(r) \left[ \left( \frac{\Delta\beta'_N(r)}{\beta'_N(r)} \right)^2 + \left( \frac{\Delta T_M^2(r_N, r)}{T_M^2(r_N, r)} \right)^2 + \left( \frac{\Delta T_P^2(r_N, r)}{T_P^2(r_N, r)} \right)^2 \right] + (\Delta\beta_M(r))^2. \quad (2.28)$$

Here,  $\beta_T(r)$  is the total backscatter coefficient at range  $r$ . The uncertainty  $\Delta\beta'_N(r)$  is the standard deviation of the average of the values of  $\beta'_N(r)$  contributing to the average profile, scaled by the renormalization factor. The uncertainty  $\Delta\beta_M(r)$  results from the uncertainty in the molecular density profile supplied by the Met Manager. This uncertainty is used to calculate the uncertainty in the molecular two-way transmittance profile  $\Delta T_M^2(r_N, r)$ . Obviously, this is an example where the uncertainties are not completely independent. The correlation between the molecular backscatter and transmittance, and the values of these quantities at consecutive ranges, is dependent on the vertical resolution of the molecular density model and that of the lidar data. Corrections for these correlations will be made when more information is obtained on the final format of the Met Manager data. Currently, no uncertainties in the meteorological data are supplied by the Met Manager so the required uncertainties are calculated from the variability of the meteorological profiles in any 80 km along-track section of data. Meteorological profiles are available, typically, every 55 km along track. The molecular backscatter and two-way transmittance profiles used in the analysis of lidar data in the 80 km section of data are calculated from the averages of the available meteorological profiles. Because there is usually a trend in the changes in the profiles along track, the differences between the average profiles for an 80 km section and the true profiles are likely to be systematic. Therefore, the uncertainties in the molecular backscatter and two-way transmittance profiles are treated as bias errors in the current analysis. Note that the uncertainty at the normalization range,  $\Delta T_M^2(0, r_N)$  will bias the

normalization of the whole profile and is considered in the next section. The second term in equation (2.28) refers to random errors in the molecular transmittance between  $r_N$  and  $r$ . The uncertainty in the particulate transmittance term in equation (2.28) is derived from equation (2.16):

$$\left( \frac{\Delta T_P^2(r_N, r)}{T_P^2(r_N, r)} \right) = -2\eta(r)\Delta\tau_P(r_N, r). \quad (2.29)$$

The optical thickness can be expressed as the product of the lidar ratio and the integrated backscatter:

$$\tau_P(r_N, r) = S_P\gamma_P(r_N, r), \quad (2.30)$$

from which we can derive an expression for the uncertainty:

$$\Delta\tau(r_N, r) = S_P\Delta\gamma_P(r_N, r). \quad (2.31)$$

In the software implementation of the extinction retrieval algorithm, the integrated particulate backscatter coefficient is calculated at each range increment using trapezoidal integration:

$$\gamma_P(r_N, r) = \gamma_{Pj} = 0.5 \sum_{i=j_{cal}}^{j-1} (\beta_{Pi} + \beta_{Pi+1})(r_{i+1} - r_i) = 0.5 \sum_{i=j_{cal}}^{j-1} (\beta_{Pi} + \beta_{Pi+1})\delta r_{i+1}, \quad (2.32)$$

where  $\delta r_j$  is the  $j$ th range increment. (These range increments are not all the same.) Because the uncertainties in the particulate backscatter are primarily a result of the statistical noise in the input profiles of attenuated backscatter and are, therefore, uncorrelated, the uncertainty in  $\gamma_{Pj}$  can now be calculated by summing the variances:

$$(\Delta\gamma_{Pj})^2 = 0.25((\delta r_j\Delta\beta_{Pj})^2 + (\delta r_{j_{cal}}\Delta\beta_{Pj_{cal}})^2 + \sum_{i=j_{cal}+1}^{j-1} (\delta r_i + \delta r_{i+1})^2(\Delta\beta_{Pi})^2). \quad (2.33)$$

However, as the first term in equation (2.33) is the unknown in equation (2.28), this term is evaluated by initially setting  $\Delta\beta_j$  to  $\Delta\beta_{j-1}$  and iterating using equations (2.28) and (2.33) until convergence is achieved following a similar procedure to that used in the iteration of equations (2.15) and (2.16). Finally, the uncertainty in the aerosol extinction coefficient can be derived from equation (2.17):

$$\Delta\sigma_P(r) = S_P\Delta\beta_P(r). \quad (2.34)$$

### *Biases produced by systematic errors*

We now consider the propagation of biases due to errors in the initial calibration and in the effective lidar ratio. (The effective lidar ratio is the product of the lidar ratio and the multiple scattering function, both of which may be in error. In the analysis below, the errors will be combined.) It will be shown how errors in these quantities cause bias errors in the retrieved particulate extinction profile and thence in the particulate transmittance that is used to renormalize signals from lower regions of the atmosphere.

If a retrieval is initiated at the first point in the profile near 40 km altitude, then the normalization factor is calculated from equation (2.7) in which it is assumed that the transmittance losses due to aerosols above the normalization height, and aerosol scattering at the normalization height, are both zero. (It can be shown that the two assumptions are equivalent in the current analysis.) If

this is not the case and either the molecular or particulate two-way transmittance is specified incorrectly, then a bias error will occur. If  $\hat{T}^2$  represents the estimated value of the two-way transmittance, then the estimated, erroneous, retrieved particulate backscatter can be calculated from equation (2.15) as

$$\hat{\beta}_p(r) = \frac{\beta'(r)}{\hat{T}_M^2(0,r)\hat{T}_P^2(0,r)} - \beta_M(r) = \frac{F_{N0}\beta'(r)}{T_M^2(0,r)T_P^2(0,r)} - \beta_M(r), \quad (2.35)$$

where

$$F_{N0} = \frac{T_M^2(0,r)T_P^2(0,r)}{\hat{T}_M^2(0,r)\hat{T}_P^2(0,r)} \quad (2.36)$$

is the factor that describes the normalization bias error. (This factor can include the estimate of the calibration error.) The bias error in the retrieved particulate backscatter is then

$$\varepsilon_b(\beta_p(r)) = \hat{\beta}_p(r) - \beta_p(r) = (F_{N0} - 1)\beta_p(r) = (F_{N0} - 1)\beta_p(r) + (F_{N0} - 1)\beta_M(r) \quad (2.37)$$

so that the normalization error produces an apparent error in the molecular backscatter as well as in the particulate backscatter. If we then define the relative error in the effective particulate lidar ratio that we select for the analysis as

$$R_S = \hat{S}_p / S_p, \quad (2.38)$$

then the estimated, erroneous, particulate extinction coefficient can be written as

$$\hat{\sigma}_p(r) = \hat{S}_p \hat{\beta}_p(r) = R_S S_p F_{N0} \beta_p(r) + R_S S_p (F_{N0} - 1) \beta_M(r). \quad (2.39)$$

The bias error in the retrieved particulate extinction is then

$$\varepsilon_b(\sigma_p(r)) = \hat{\sigma}_p(r) - \sigma_p(r) = (R_S F_{N0} - 1) S_p \beta_p(r) + R_S S_p (F_{N0} - 1) \beta_M(r). \quad (2.40)$$

The biased particulate optical thickness over the next range increment  $\delta r$  can then be derived:

$$\hat{\tau}_p(r, r + \delta r) = \delta r \hat{\sigma}_p(r) = \delta r R_S S_p F_{N0} \beta_p(r) + \delta r R_S S_p (F_{N0} - 1) \beta_M(r), \quad (2.41)$$

with bias error

$$\varepsilon_b(\tau_p(r, r + \delta r)) = \delta r (R_S F_{N0} - 1) S_p \beta_p(r) + \delta r R_S S_p (F_{N0} - 1) \beta_M(r). \quad (2.42)$$

Assuming that the second term in (2.42) is the only systematic error in  $\tau_M(r, r + \delta r)$ , we can write

$$\tau_T(r, r + \delta r) = \tau_p(r, r + \delta r) + \tau_M(r, r + \delta r), \quad (2.43)$$

and

$$\varepsilon_b(\tau_T(r, r + \delta r)) = \hat{\tau}_p(r, r + \delta r) - \tau_p(r, r + \delta r) = \varepsilon_b(\tau_p(r, r + \delta r)). \quad (2.44)$$

By combining equations (2.36) and (2.44) we can write

$$F_N = T^2 / \hat{T}^2 = \exp(2(\hat{\tau}_p - \tau_p)) = \exp(2\varepsilon_b(\tau_p)), \quad (2.45)$$

which then leads to the expression for the bias error in the retrieval at the next range increment:

$$F_{N1} = \frac{T_M^2(r, r + \delta r) T_P^2(r, r + \delta r)}{\hat{T}_M^2(r, r + \delta r) \hat{T}_P^2(r, r + \delta r)} \quad (2.46)$$

$$= \exp\{2\delta r S_p (R_S F_{N0} - 1) \beta_p(r)\} \exp\{2\delta r S_p R_S (F_{N0} - 1) \beta_M(r)\}.$$

This bias then feeds into the calculation of the particular backscatter at the next range:

$$\hat{\beta}_p(r + \delta r) = \frac{\beta'(r + \delta r)}{\hat{T}_M^2(0, r) \hat{T}_P^2(0, r) \hat{T}_M^2(r, r + \delta r) \hat{T}_P^2(r, r + \delta r)} - \beta_M(r, r + \delta r), \quad (2.47)$$

or

$$\hat{\beta}_p(r + \delta r) = \frac{F_{N1} F_{N0} \beta'(r + \delta r)}{T_M^2(0, r + \delta r) T_P^2(0, r + \delta r)} - \beta_M(r + \delta r). \quad (2.48)$$

The bias error is

$$\varepsilon_b(\beta_p(r + \delta r)) = (F_{N1} F_{N0} - 1) \beta_p(r + \delta r) + (F_{N1} F_{N0} - 1) \beta_M(r + \delta r). \quad (2.49)$$

It can be shown then that the propagated bias error in the retrieved quantities at  $n$  range increments from the starting range can be described by the general equations:

$$\varepsilon_b(\beta_p(r + n\delta r)) = \left(\prod_{j=1}^n F_{Nj} - 1\right) \beta_p(r + n\delta r) + \left(\prod_{j=1}^n F_{Nj} - 1\right) \beta_M(r + n\delta r), \quad (2.50)$$

$$\varepsilon_b(\sigma_p(r + n\delta r)) = S_p (R_S \prod_{j=1}^n F_{Nj} - 1) \beta_p(r + n\delta r) + S_p R_S (\prod_{j=1}^n F_{Nj} - 1) \beta_M(r + n\delta r), \quad (2.51)$$

$$\varepsilon_b(\tau_p(r + n\delta, r + (n+1)\delta r)) = \delta r \varepsilon_b(\sigma_p(r + n\delta r)), \quad (2.52)$$

and

$$F_{Nn+1} = \exp\left\{2\delta r S_p (R_S \prod_{j=1}^n F_{Nj} - 1) \beta_p(r + n\delta r)\right\} \exp\left\{2\delta r S_p R_S (\prod_{j=1}^n F_{Nj} - 1) \beta_M(r + n\delta r)\right\}. \quad (2.53)$$

Although the random and bias errors in the retrieved quantities are calculated and propagated separately at each successive range step, they are combined in the quoted total error estimate at each range:

$$\varepsilon_t(x) = \left((\varepsilon_b(x))^2 + (\Delta x)^2\right)^{1/2}. \quad (2.54)$$

### 2.3.3. Sensitivity Analysis

The error and uncertainty analysis presented in Section 2.3.2.4 studied the propagation of errors or uncertainties in various input quantities and derived general expressions for the resulting uncertainties in the output data products. This section examines the sensitivity of the retrievals to systematic errors in the calibration and renormalization factors, the lidar ratio used in the retrieval and errors in the transmittance values used to constrain the retrieval. The sensitivities are presented over a range of values of atmospheric parameters and systematic errors.

### 2.3.3.1. Calibration and Renormalization Errors

Using the integrated attenuated backscatter equation (Fernald et al., 1972),

$$\gamma'_p = (1 - T_p^2) / (2\bar{\eta}S), \quad (2.55)$$

and rearranging it to give an equation for the estimated two-way particulate transmittance that results from an incorrectly calibrated attenuated backscatter profile (and hence integrated attenuated backscatter) we have:

$$\widehat{T}_p^2 = 1 - 2\bar{\eta}S\widehat{\gamma}'_p. \quad (2.56)$$

Here we omit the range interval of integration for clarity and use the layer-effective multiple scattering factor  $\bar{\eta}$ . Calibration and renormalization errors can result from either an incorrect specification of the transmittance between the lidar and the calibration (or renormalization) range or an incorrect assessment of the scattering ratio at the normalization range. Either way, the attenuated backscatter profile is either too low or too high by the ratio  $R_N = \widehat{\gamma}'_p / \gamma'_p$  that takes the value unity if the calibration or renormalization is correct. (Strictly,  $R_N = \widehat{\gamma}'_T / \gamma'_T$ , but here we assume that  $\gamma'_p \gg \gamma'_M$  as it is in these conditions that the errors become significant.) Therefore, in an unconstrained retrieval,

$$\widehat{T}_p^2 = 1 - R_N(1 - T_p^2), \quad (2.57)$$

and we can see that

$$R_N = \frac{(1 - \widehat{T}_p^2)}{(1 - T_p^2)}. \quad (2.58)$$

The relative error in the retrieved  $T^2$  can then be written:

$$\frac{\varepsilon(T_p^2)}{T_p^2} = \frac{\widehat{T}_p^2 - T_p^2}{T_p^2} = \frac{1 - R_N(1 - T_p^2) - T_p^2}{T_p^2} = \frac{(1 - T_p^2)}{T_p^2}(1 - R_N). \quad (2.59)$$

Using (2.3) we can write the error in the retrieved particulate optical thickness as

$$\widehat{\tau}_p - \tau_p = -\frac{1}{2\bar{\eta}} \ln \left( \frac{1 - R_N(1 - T_p^2)}{T_p^2} \right) = -\frac{1}{2\bar{\eta}} \ln \left( \frac{1 - R_N [1 - \exp(-2\bar{\eta}\tau_p)]}{\exp(-2\bar{\eta}\tau_p)} \right), \quad (2.60)$$

or, as a relative error:

$$\frac{\varepsilon(\tau_p)}{\tau_p} = \frac{\widehat{\tau}_p - \tau_p}{\tau_p} = \frac{1}{2\bar{\eta}\tau_p} \ln \left( \frac{\exp(-2\bar{\eta}\tau_p)}{1 - R_N [1 - \exp(-2\bar{\eta}\tau_p)]} \right). \quad (2.61)$$

It can be seen that the relative error becomes undefined when

$$R_N \geq \frac{1.0}{1 - \exp(-2\bar{\eta}\tau_p)}. \quad (2.62)$$

### 2.3.3.2. Lidar Ratio Errors

Equation (2.55) above can also be used to explore the effects of errors in the specified lidar ratio on unconstrained retrievals. The incorrectly retrieved, two-way particulate transmittance resulting from an incorrectly specified lidar ratio is then

$$\hat{T}_p^2 = 1 - 2\bar{\eta}\hat{S}_p\gamma'_p. \quad (2.63)$$

Substituting for  $\gamma'_p$  using (2.55) we have

$$\hat{T}_p^2 = 1 - 2\bar{\eta}\hat{S}_p \left[ \frac{1}{2\bar{\eta}S_p} (1 - T_p^2) \right], \quad (2.64)$$

or

$$\hat{T}_p^2 = 1 - \frac{\hat{S}_p}{S_p} (1 - T_p^2). \quad (2.65)$$

The similarity with (2.57) is obvious and it is also apparent that

$$\frac{\hat{S}_p}{S_p} = \frac{1 - \hat{T}_p^2}{1 - T_p^2}, \quad (2.66)$$

which is of the same form as (2.59). Therefore, realizing that the ratios  $(\hat{S}_p / S_p)$  and  $R_N$  are equivalent, we can write the relative errors in the retrieved particulate transmittance and optical thickness as

$$\frac{\varepsilon(T_p^2)}{T_p^2} = \frac{(1 - T_p^2)}{T_p^2} \left( 1 - \frac{\hat{S}_p}{S_p} \right), \quad (2.67)$$

and

$$\frac{\varepsilon(\tau_p)}{\tau_p} = \frac{\hat{\tau}_p - \tau_p}{\tau_p} = \frac{1}{2\bar{\eta}\tau_p} \ln \left( \frac{\exp(-2\bar{\eta}\tau_p)}{1 - (\hat{S}_p/S_p)[1 - \exp(-2\bar{\eta}\tau_p)]} \right). \quad (2.68)$$

Using the symbol  $r$  to signify either  $(\hat{S}_p / S_p)$  or  $R_N$ , the sensitivities of the retrieved two-way particulate transmittance and optical thickness are plotted in Figure 2.5 and Figure 2.6, respectively. Equation (2.67) can also be written in the form

$$\frac{\varepsilon(T_p^2)}{T_p^2} = \frac{(1 - T_p^2)}{T_p^2} \left( 1 - \frac{\hat{S}_p}{S_p} \right) = \frac{(1 - T_p^2)}{T_p^2} \left( \frac{S_p - \hat{S}_p}{S_p} \right) = \frac{(1 - T_p^2)}{T_p^2} \left( \frac{\varepsilon(S_p)}{S_p} \right). \quad (2.69)$$

Thus, Figure 2.6 also shows the sensitivity of the retrieved particulate transmittance to relative errors in the specified lidar ratio.

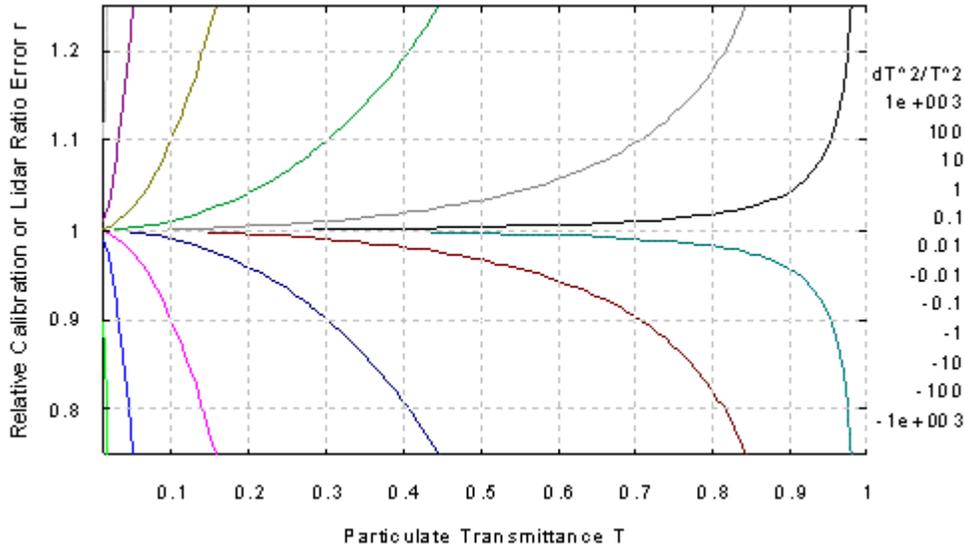


Figure 2.5: Sensitivity of retrieved two-way particulate transmittance to errors in the calibration or specified lidar ratio. Contours are of the relative error in  $T_p^2$  as a function of  $T_p^2$  and  $r$ , where  $r$  is either the relative calibration error  $R_N$  or the ratio of the specified to true lidar ratio. Errors are positive for  $r < 1.0$ .

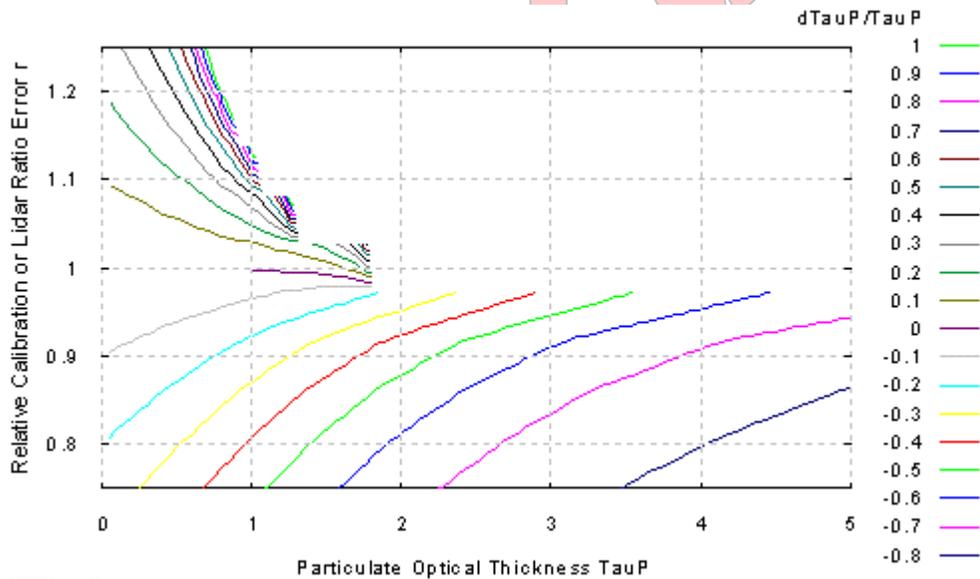


Figure 2.6: Sensitivity of retrieved particulate optical thickness to errors in the calibration or specified lidar ratio. Contours are of the relative error in  $\tau_p$  as a function of  $\tau_p$  and  $r$ , where  $r$  is either the relative calibration error  $R_N$  or the ratio of the specified to true lidar ratio. The multiple scattering factor is taken as unity. Positive errors occur for  $r > 1$  and negative errors for  $r < 1$ . The errors in the top-right region of the figure are undefined as described by equation (2.61).

### 2.3.3.3. Transmittance Constraint Errors

Consider now the situation of constrained retrievals. As the retrieval is constrained by adjusting the lidar ratio to match the retrieved transmittance (or optical thickness) to a specified value, the

error in the lidar ratio that results from the use of a wrong transmittance constraint can be determined from equation (2.69) as

$$\frac{\varepsilon(S_p)}{S_p} = \left( \frac{T_p^2}{1-T_p^2} \right) \frac{\varepsilon(T_p^2)}{T_p^2}. \quad (2.70)$$

It can be seen that relative error in the retrieved lidar ratio increases for values of transmittance close to unity. In these situations of low attenuation, the retrieval of an extinction profile is only relatively weakly dependent on the lidar ratio and, therefore, the relative error is large.

The relative accuracy to which  $S_p$  must be known, in an unconstrained retrieval, to achieve the particulate optical thickness to a specified relative accuracy can be derived from equation (2.70) using equation (2.3):

$$\frac{\varepsilon(S_p)}{S_p} = 2\eta\tau_p \left( \frac{\exp(-2\bar{\eta}\tau_p)}{1-\exp(-2\bar{\eta}\tau_p)} \right) \frac{\varepsilon(\tau_p)}{\tau_p}. \quad (2.71)$$

This is plotted in Figure 2.7 as a function of  $\tau_p$  and  $\varepsilon(\tau_p)/\tau_p$  using a value of unity for  $\eta$ . In Figure 2.8, the same function is plotted using a fixed value of the relative error in the particulate optical thickness of ten percent.

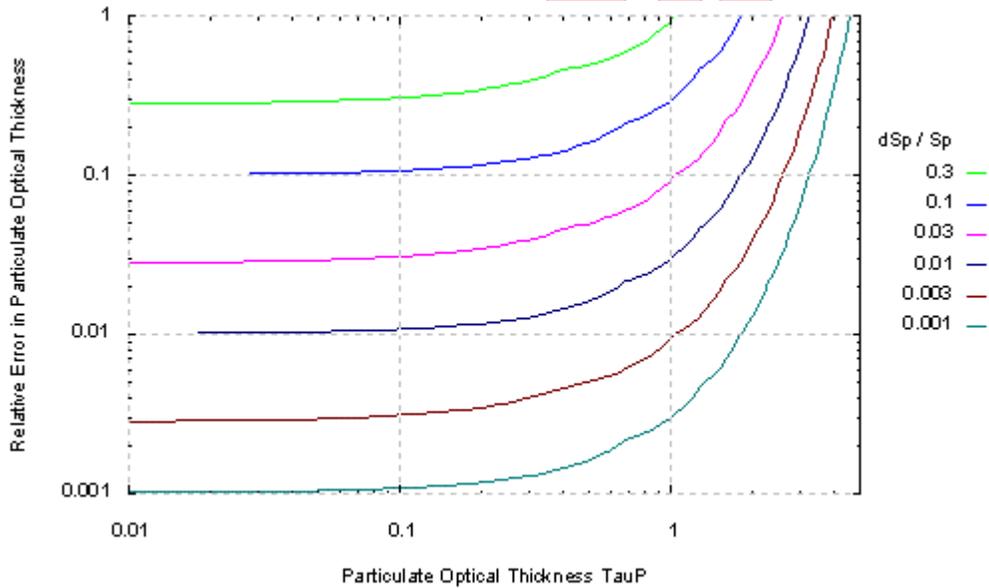


Figure 2.7: The required relative accuracy in the specified lidar ratio in order to retrieve the particulate optical thickness to a specified accuracy, plotted as a function of the particulate optical thickness of the feature.

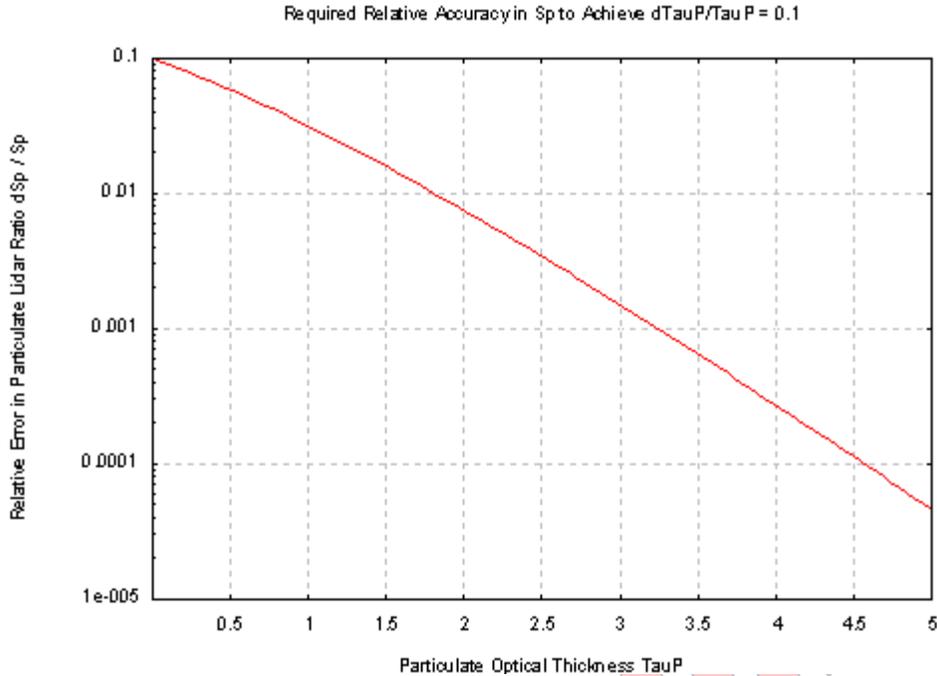


Figure 2.8: The required relative accuracy in the specified lidar ratio in order to retrieve the particulate optical thickness to a specified accuracy of 10% plotted as a function of the optical thickness of the feature.

## 2.4. Implementation of the Hybrid Extinction Retrieval Algorithm (HERA)

### 2.4.1. Introduction

This section describes the algorithms needed for the retrieval of particulate backscatter and extinction information from CALIPSO lidar data. The analysis proceeds top down, across all resolutions, in order of increasing range of the features from the lidar. HERA creates and processes its own averaged profiles using the fine (5 km) resolution input profiles and the layer description data supplied by SIBYL.

As the analysis is quite complicated and involves many different pathways depending on the nature of the signal being processed, whether it be feature or clear air, cloud or aerosol, partially or completely attenuated, among other things, the HERA procedures are described using several annotated flowcharts. The description below begins with an overview of the algorithms, and then follows with more detailed flow charts and explanations for cases of simple features, complex features and embedded features. While the overview is written to describe the procedures in a general manner, which is without reference to a particular software framework, the detailed explanations that follow necessarily refer in places to particular data structures and arrays. This is unavoidable for at least two reasons. First, the actual “extinction engine” or iterative retrieval algorithm already exists and has a defined interface of input and output parameters, and an expectation of the availability of other data through various modules. Secondly, within this document several specific variable names and operations are defined in order to improve the legibility and to reduce ambiguity in the flowcharts.

## 2.4.2. Overview of the Algorithms

An overview of the algorithm structure is presented in Figure 2.9. This flowchart contains three columns, each describing the tasks involved in retrieving extinction for scenes of different degrees of complexity. In the following, reference to variables in the flowcharts will use 12-pt Arial font except where Greek symbols are used.

### 2.4.2.1. Simple features and clear regions

The column on the left considers retrievals from the simplest cases, those from apparently clear regions and from simple features. Clear regions are those regions in which no features have been identified. Simple features are layers bounded everywhere both above and below by clear regions; that is, they are not vertically adjacent to other features at the top or base, nor do they contain other embedded features. Simple features are assumed to be homogeneous, in that they contain particles of only one type that can be correctly characterized by a single lidar ratio.

The process begins by retrieving the analysis parameters stored in a run-time file, and reading the layer descriptors and lidar and meteorological profiles. The layer descriptors contain information on the location of top and base, horizontal extent, two-way transmittance, lidar ratio, integrated attenuated backscatter, and so on, for each detected feature listed in order of increasing range from the lidar (decreasing altitude). The lidar profile data are profiles of attenuated backscatter, calibrated at the top of the atmosphere, and averaged to 5 km horizontal resolution. As described in the layer detection ATBD (PC-SCI-202 Part 2), the profiles retrieved by HERA have been cleared of finer-resolution, boundary layer clouds. The rest of the steps are contained in a loop that begins with the first feature and progresses to the last. Here  $n_f$  is the total number of features detected in the data block. The quantity  $F_{jf}$  that appears in the flowchart means feature number  $jf$ , where  $jf$  is the index of the loop. To improve legibility, use is made of a logical variable `done` that is set to `TRUE` when the relevant feature has been satisfactorily analyzed. So the third box in the column is simply testing whether feature with index  $jf$  has been processed.

The analysis algorithm proceeds by analyzing clear regions above a feature then by analyzing the feature.<sup>1</sup> First, the vertical and horizontal extent of the clear region needs to be defined. Once this is done, the signals from this region are averaged to produce a representative profile that is then analyzed (or solved). The attenuation associated with the clear region is calculated in this step and is then used to correct all the underlying lidar profiles of attenuated backscatter,  $\beta'$ , for this attenuation. As the layer descriptors may also be affected by this attenuation, these are also rescaled where necessary. Situations where this rescaling is not required are those where SIBYL has renormalized below an overlying feature and effectively rescaled all data below this by the attenuation of the overlying feature.

The rest of the column on the left concerns the analysis of features. For simple features, the process is similar to that for clear regions, with the spatial extent of the feature being defined, the signals in this region averaged, the profile solved and the underlying data rescaled. If a

---

<sup>1</sup> Readers are reminded that solutions for “clear air” regions can be disabled using the HERA configuration file, and that in such cases the particulate optical depths and backscatter and extinction coefficients within these regions are set uniformly to zero. **For the initial release of the CALIOP extinction data products, solutions in “clear air” are disabled.**

constrained solution is required, the lidar ratio is adjusted to constrain the retrieved optical thickness to match that supplied by the SCA. Once a satisfactory retrieval has been achieved, `done` is set `TRUE` for that feature. Before a feature is solved, a feature is examined to see if it is complex (vertically adjacent to another feature). This is indicated by the test half way down the column.

#### 2.4.2.2. Complex features

Complex features, defined here as contiguous regions that contain features that are vertically adjacent to other dissimilar features are considered in the central column. As there may be several complex features in an 80 km data block, each is given an index and defined by the indices of its constituent features. Similarly, the layer descriptors for each constituent feature specify the index of the particular complex feature to which the constituent feature belongs. The first step is to identify the complex feature. Here `jcf` is the index of the complex feature that contains the feature identified by the index `jf` in the loop in the left column.

For the case of complex scenes, there are more unknowns (i.e., the lidar ratios of the individual, constituent features) than constraints (transmittance values or optical thicknesses). The problem is, therefore, underdetermined and no unique solution exists. However, HERA attempts to find a solution that produces an overall attenuation of the complex feature that is *consistent* with the value determined by considering the reduction in the signal passing through the complex feature. This attenuation is the effective optical thickness ( $\eta\tau$ ) averaged over the complete horizontal extent of the complex feature. This step is indicated in the second of the three boxes linking the first column to the start of the central column. The central column contains a loop in which each of the `mcf` features in the current complex feature is analyzed, and the lidar ratios of the constituent features adjusted in an attempt to achieve consistency between the calculated effective optical thickness,  $\eta\tau_c$ , calculated from the retrieved extinction profiles and the values of  $\eta$ , and the measured value,  $\eta\tau_m$ .

The solution of the signals from the complex features is comparable to the solution of simple features in the first column, with one significant variation. As the features are vertically adjacent, and some may contain embedded features, it is often not possible to obtain a final solution for each feature and then correct underlying features for the associated attenuation before the underlying features are solved. Because the final values of all the lidar ratios are not known at the start, solutions must be iterated and their contributions to the attenuation of underlying features continually tracked and updated. While there are numerous software strategies that can be successfully applied to this task, within the flowcharts we will describe the required operations with reference to a temporary ‘scratch array’,  $\sigma_{temp}$ . The overlying attenuation for each constituent of a complex feature is calculated by integrating the values in  $\sigma_{temp}$  that lie above the feature in question, and correcting for the multiple scatter function  $\eta$ . This integral (here `ccal`) is used for temporary scaling of the attenuated backscatter profile before an extinction retrieval is performed.

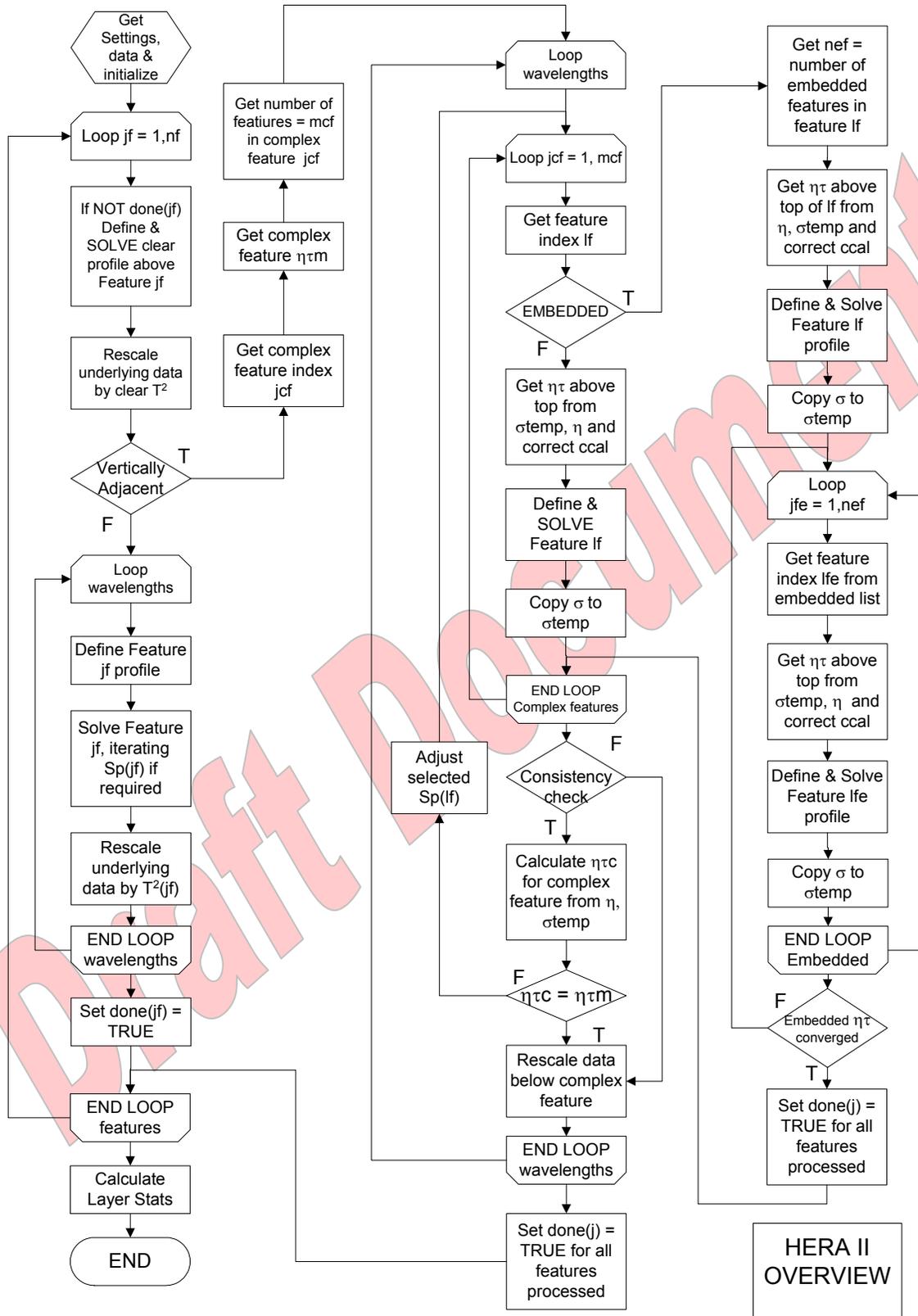


Figure 2.9: Overview of HERA’s analysis pathways.

If a check of the consistency of the retrieved and measured optical thicknesses is feasible, then the calculated value is obtained by integrating  $\sigma_{\text{temp}}$ , scaled by  $\eta$ , and averaging across the horizontal extent of the complex feature. If the measured and calculated values are not consistent within a predefined tolerance, then adjustments are made to the lidar ratio of the feature with the largest  $\gamma'$ . This procedure is repeated as required, with the lidar ratios of progressively weaker features (decreasing  $\gamma'$ ) being adjusted until consistency is achieved. Attenuated backscatter profile data for regions that underlie the complex feature are then rescaled for the attenuation in each column of the complex feature, and all the features that have been processed are marked as having been “done”. Processing then returns to the main loop in the first column.

#### 2.4.2.3. Complex embedded features

A further degree of complexity arises when features contain embedded features. An embedded feature is one that has a smaller horizontal and vertical extent than the feature that contains it. The larger feature cannot be solved independently in this case, as parts of it have been attenuated by some unknown amount by the smaller, embedded feature. Nor can the embedded feature be solved until the attenuation due to overlying regions of the containing feature is known. The problem is solved iteratively, with the larger feature being solved first, then the embedded features. Although the retrieval of the regions of the larger feature that underlie the embedded feature will be incorrect on the first pass, because of the lack of correction for the attenuation caused by the embedded feature, the embedded feature, which is solved after the larger feature, will be solved correctly. On the second pass, the averaged profile of the larger, surrounding feature will be correctly rescaled for the attenuation of the embedded feature and the retrieved extinction profile will be correct. For complex features containing many embedded features, or with multiple levels of embeddedness, more than two passes will be required before all the features are retrieved correctly. The convergence to the correct answer is determined by checking the effective embedded  $\eta\tau$ .

The analysis of features that contain embedded features is shown in the column on the right. The process begins by determining the number of features, *nembed*, that are embedded in the current feature, *lf*. Next, this larger feature is solved in much the same way as the complex features that do not contain embedded features are solved in the central column. The first significant difference is the internal loop in which each of the embedded features, *lfe*, is solved. The other difference is the iteration of the retrieval of the complex, embedded feature until the total effective optical thickness converges. After all the features have been analyzed the processing returns to the complex feature loop in the central column.

### 2.4.3. Detailed Description of the Analysis Process

#### 2.4.3.1. Definitions

Throughout this section, reference will be made to clear air or clear regions, simple features, complex features, and embedded features. These terms are defined here with further explanation.

*Clear regions* are regions in which there are no *identified* features. However, there may be particulate scatterers that cause significant attenuation. Such materials, like biomass-burning aerosols, have a high carbon content and thus a high ratio of extinction to backscatter.

Therefore, while the attenuation may be significant, the backscatter contrast may be too low to enable detection.

*Simple features* are features that are not in contact with other features at their bases or tops (i.e. they are not vertically adjacent) and do not contain embedded features. They are characterized by a single type of scatterer with a single lidar ratio at each wavelength. If an independent measurement of transmittance is available, then this may be used to derive a constrained retrieval of the extinction profile through the simple feature.

*Complex features* are features that are made up of many simple features that are connected to form a continuous area of particulate scattering. Complex features have regions with different lidar ratios. The simplest example is two layers, of the same horizontal resolution but with different lidar ratios, which are vertically adjacent. Complex features can also contain embedded features. Because there are more unknowns (lidar ratios) than available constraints, (transmittance measurements), constrained solutions are not possible with complex features. The best that can be done is that the lidar ratios of the constituent features be adjusted until the average optical thickness, measured across the whole horizontal extent of the feature, agrees with that calculated over the same horizontal extent from the retrieved extinction profiles of the constituent features.

*Embedded features* are features that have a smaller horizontal extent than the feature in which they are embedded, and are contained completely within the vertical limits of the containing feature. Because the transmittance of the embedded feature cannot be determined separately from that of the containing feature, a constrained retrieval is not possible.

#### 2.4.3.2. Required Input Data

##### (a) Settings File Data

To ensure maximum flexibility at runtime, the various constants and computational limits that control the execution of the HERA routines are stored externally in a configuration file.

##### (b) Meteorological Data

Various profiles of meteorological parameters are used in several areas of the algorithms, including signal normalization and extinction, and backscatter profile retrieval. The required quantities are listed below in Table 2.1 (scalar quantities) and Table 2.2 (profile data).

##### (c) Layer Descriptor Data

These data are supplied by SIBYL and/or the SCA, and define the spatial and optical properties of each feature, including all of the following:

- A set of “layer descriptors” that contains an ordered list of all features detected within the data block. These features will be listed in order of decreasing altitude of the top of the feature, regardless of the horizontal averaging used to detect the feature. Features with identical tops are further sorted in order of increasing “adjacency values” (defined below)
- A separate complex feature list that includes an ordered list of the features contained in each complex feature. The number and index of features embedded within each feature will be supplied also. See definitions of *mcfn*, *cffn*, *cfgp*, and *embedlist* below.
- Individual layer descriptors will include the two-way transmittance,  $T^2$ , integrated attenuated backscatter,  $\gamma'$ , particulate lidar ratio,  $S_p$ , and an opacity flag (defined below)

Table 2.3 lists the required parameters.

*(d) Lidar Profile Data*

Profiles of measured lidar backscatter data, averaged to the standard HERA retrieval grid (i.e., 5 km horizontal resolution spanning a continuous distance of 80 km); see Table 2.4 for a complete listing.

Table 2.1: Meteorological Data

|    | Flowchart Parameter | Interpretation   |
|----|---------------------|--|
| 1  | ccal_532            | The 532-nm calibration factor                            |
| 2  | dc532               | Uncertainty in calibration factor at 532 nm              |
| 3  | ccal_064            | The 1064-nm calibration factor                           |
| 4  | dc064               | Uncertainty in calibration factor at 1064 nm             |
| 5  | h_trop              | Height of tropopause (km)                                |
| 6  | i_trop              | Array index of tropopause                                |
| 7  | z_lid               | Altitude of lidar (km)                                   |
| 8  | Sa_532_strat        | Stratospheric Aerosol lidar ratio for clear regions (sr) |
| 9  | dSa_532_strat       | Uncertainty in Stratospheric Aerosol lidar ratio (sr)    |
| 10 | Sa_532_trop         | Tropospheric Aerosol lidar ratio for clear regions (sr)  |
| 11 | dSa_532_trop        | Uncertainty in Tropospheric Aerosol lidar ratio (sr)     |

Table 2.2: Meteorological Profile Data

|    | Flowchart Parameter | Interpretation   |
|----|---------------------|--|
| 1  | nvals               | Number of values (height levels) in profiles   |
| 2  | altitude(nvals)     | Altitude array for lidar and meteorological data (km)  |
| 3  | bm532(nvals)        | Array of 532-nm molecular backscatter values ( $\text{km}^{-1}\text{sr}^{-1}$ )              |
| 4  | dbm532(nvals)       | Array of 532-nm molecular backscatter uncertainty values ( $\text{km}^{-1}\text{sr}^{-1}$ )  |
| 5  | t2m532(nvals)       | Array of 532-nm molecular transmittance values   |
| 6  | dt2m532(nvals)      | Array of 532-nm molecular transmittance uncertainty values                                   |
| 7  | bm064(nvals)        | Array of 1064-nm molecular backscatter values ( $\text{km}^{-1}\text{sr}^{-1}$ )             |
| 8  | dbm064(nvals)       | Array of 1064-nm molecular backscatter uncertainty values ( $\text{km}^{-1}\text{sr}^{-1}$ ) |
| 9  | t2m064(nvals)       | Array of 1064-nm molecular transmittance values  |
| 10 | dt2m064(nvals)      | Array of 1064-nm molecular backscatter uncertainty values                                    |

Table 2.3: Detected Feature Data

|    | Flowchart Parameter | Interpretation  |
|----|---------------------|---|
| 1  | nlayers = nf        | Number of detected layers (features)                                      |
| 2  | h_surf(ncols)       | Altitude of surface (km). This is supplied for each of the nfine columns. |
| 3  | i_surf(ncols)       | Array index of altitude of surface  |
| 4  | h_last(ncols)       | Last valid altitude in signal (km)  |
| 5  | i_last(ncols)       | Array index of last valid altitude  |
| 6  | moflag(j)           | Feature opacity flag for feature j. (See definition below.)               |
| 7  | maflag(j)           | Feature adjacency flag (See definition below.)                            |
| 8  | mcfn(j)             | Index of complex feature containing current feature                       |
| 9  | mhav(j)             | Horizontal averaging resolution (1, 4, 16)                                |
| 10 | mhfst(j)            | First column(horizontal index) in which feature detected                  |
| 11 | mhlst(j)            | Last column(horizontal index) in which feature detected                   |
| 12 | hbase(j)            | Altitude of base of feature number $j \leq nlayers$ (km)                  |
| 13 | mbase(j)            | Array index of base j   |
| 14 | htop(j)             | Altitude of top of feature j  |
| 15 | mtop(j)             | Array index of top j  |
| 16 | t2l532(j)           | Two-way transmittance at 532 nm for feature j.                            |
| 17 | dt2l532(j)          | Uncertainty in two-way transmittance at 532 nm for feature j              |
| 18 | Sp532(j)            | Average lidar ratio at 532 nm for feature j (sr)                          |
| 19 | dSp532(j)           | Uncertainty in average lidar ratio at 532 nm for feature j (sr)           |
| 20 | gp532(j)            | Integrated attenuated backscatter at 532 nm for feature j ( $sr^{-1}$ )   |
| 21 | dgp532(j)           | Uncertainty in 532-nm integrated attenuated backscatter ( $sr^{-1}$ )     |
| 22 | t2l1064(j)          | Two-way transmittance at 1064 nm for feature j                            |
| 23 | dt2l1064(j)         | Uncertainty in two-way transmittance at 1064 nm for feature j             |
| 24 | Sp064(j)            | Average lidar ratio at 1064 nm for feature j (sr)                         |
| 25 | dSp064(j)           | Uncertainty in average lidar ratio at 1064 nm for feature j (sr)          |
| 26 | gp064(j)            | Integrated attenuated backscatter at 1064 nm for feature j ( $sr^{-1}$ )  |
| 27 | dgp064(j)           | Uncertainty in 1064-nm integrated attenuated backscatter ( $sr^{-1}$ )    |

Table 2.4: Lidar and Model Profile Data (j is the range index and k is the 5 km profile index)

|    | Parameter     | Meaning  |
|----|---------------|--|
| 1  | bt2_532(j,k)  | Profile of attenuated backscatter $\beta'(r)$ at 532 nm ( $\text{km}^{-1}\text{sr}^{-1}$ )                     |
| 2  | dbt2_532(j,k) | Uncertainty in attenuated backscatter, $\Delta\beta'(r)$ at 532 nm ( $\text{km}^{-1}\text{sr}^{-1}$ )          |
| 3  | eta532(j,k)   | Profile of multiple scattering factor $\eta(r)$ at 532 nm  |
| 4  | deta532(j,k)  | Uncertainty in multiple scattering factor, $\Delta\eta(r)$ at 532 nm   |
| 5  | bt2_064(j,k)  | Profile of attenuated backscatter $\beta'(r)$ at 1064 nm ( $\text{km}^{-1}\text{sr}^{-1}$ )                    |
| 6  | dbt2_064(j,k) | Profile of attenuated backscatter uncertainty, $\Delta\beta'(r)$ at 1064 nm ( $\text{km}^{-1}\text{sr}^{-1}$ ) |
| 7  | eta064(j,k)   | Profile of multiple scattering factor $\eta(r)$ at 1064 nm   |
| 8  | deta064(j,k)  | Profile of uncertainty in multiple scattering factor, $\Delta\eta(r)$ at 1064 nm                               |
| 9  | t2Aref(j)     | Reference profile of background aerosol transmittance $\tau_{\text{Aref}}^2(r)$ (default=1.0)                  |
| 10 | dt2Aref(j)    | Profile of uncertainty in background aerosol transmittance, $\Delta\tau_{\text{Aref}}^2(r)$                    |

### 2.4.3.3. Internal and External Output Data

#### (a) Extinction Retrieval Quality Flags

On completion of the analysis of each feature, a quality flag is assigned according to the nature of the retrieval.

- mqaflag(nf) = retrieval quality flag (Add bits. )
- = 0 unconstrained retrieval
  - = 1 constrained retrieval
  - = 2 Sp reduced to prevent divergence
  - = 4 Sp increased to reduce negative values of  $\beta_p$
  - = 8 Surface detected ( $\beta_p > \beta_{\text{max}}$ )
  - = 16  $\tau > \tau_{\text{max}}$ . Signal is totally attenuated
  - = 32  $\Delta\tau > \Delta\tau_{\text{max}}$
  - = 64 Sp converges but too many negative values
  - = 128 Retrieval terminated at maximum iterations
  - = 256 Solution not possible with acceptable / supplied Sp
  - = 512 Consistency not possible with acceptable/supplied Sp
  - = 1024 Feature is top feature in column
  - = 2048 Overlying  $\eta\tau < 1.0$

- = 4096 Overlying  $\eta\tau < 2.0$
- = 8192 Overlying  $\eta\tau \geq 2.0$
- = 32768 No extinction retrieval attempted for feature

*(b) Output Layer Variables*

Some of the data values below also form part of the “layer descriptor” data supplied by SIBYL and/or the SCA. The values listed below in Table 2.5 and Table 2.6, however, are updated values of the parameters as calculated by HERA.

Table 2.5: Output Layer Data

|    | Parameter    | Meaning  |
|----|--------------|--|
| 1  | C_n532(j)    | Updated value of the 532-nm normalization factor for feature j.                          |
| 2  | dC_n532(j)   | Uncertainty in the 532-nm normalization factor for feature j                             |
| 3  | Sf532(j)     | Updated (final) value of 532-nm lidar ratio for feature j. (sr)                          |
| 4  | dSf532(j)    | Updated (final) value of 532-nm lidar ratio uncertainty for feature j. (sr)              |
| 5  | tauf532(j)   | Updated (final) value of 532-nm optical thickness for feature j.                         |
| 6  | dtauf532(j)  | Uncertainty in the value of 532-nm optical thickness for feature j.                      |
| 7  | beta1532(j)  | Layer averaged 532-nm particulate backscatter for feature j. ((km.sr) <sup>-1</sup> ).   |
| 8  | dbeta1532(j) | Uncertainty in averaged 532-nm particulate backscatter for feature j.(sr <sup>-1</sup> ) |
| 9  | C_n064(j)    | Updated value of the 1064-nm normalization factor for feature j.                         |
| 10 | dC_n064(j)   | Uncertainty in the 1064-nm normalization factor for feature j.                           |
| 11 | Sf064(j)     | Updated value of 1064-nm lidar ratio for feature j. (sr)                                 |
| 12 | dSf064(j)    | Uncertainty in the value of 1064-nm lidar ratio for feature j. (sr)                      |
| 13 | tauf064(j)   | Updated value of 1064-nm optical thickness for feature j.                                |
| 14 | dtauf064(j)  | Uncertainty in the 1064-nm optical thickness for feature j.                              |
| 15 | beta1064(j)  | Layer averaged 1064-nm particulate backscatter in feature j .((km.sr) <sup>-1</sup> )    |
| 16 | dbeta1064(j) | Uncertainty in averaged 1064-nm particulate backscatter in feature j.                    |
| 17 | cr(j)        | Weighted mean 1064-nm / 532-nm backscatter color ratio through feature j.                |
| 18 | dcr(j)       | Uncertainty in mean 1064-nm / 532-nm backscatter color ratio for feature j.              |

(c) Output Profile Data

Table 2.6: Output Profile Data

|   | Parameter       | Meaning   |
|---|-----------------|---|
| 1 | betap_532(j,k)  | Profile of particulate backscatter $\beta_p(r)$ at 532 nm ( $\text{km}^{-1}\text{sr}^{-1}$ )                  |
| 2 | dbetap_532(j,k) | Uncertainty in 532-nm particulate backscatter $\Delta\beta_p(r)$ at 532 nm ( $\text{km}^{-1}\text{sr}^{-1}$ ) |
| 3 | sigmap532(j,k)  | Profile of particulate extinction $\sigma_p(r)$ at 532 nm ( $\text{km}^{-1}$ )                                |
| 4 | dsigmap532(j,k) | Uncertainty in 532-nm particulate extinction, $\Delta\sigma_p(r)$ at 532 nm ( $\text{km}^{-1}$ )              |
| 5 | betap_064(j,k)  | Profile of particulate backscatter $\beta_p(r)$ at 1064 nm ( $\text{km}^{-1}\text{sr}^{-1}$ )                 |
| 6 | dbetap_064(j,k) | Uncertainty in 1064-nm particulate backscatter, $\Delta\beta_p(r)$ ( $\text{km}^{-1}\text{sr}^{-1}$ )         |
| 7 | sigmap064(j,k)  | Profile of particulate extinction $\sigma_p(r)$ at 1064 nm ( $\text{km}^{-1}$ )                               |
| 8 | dsigmap064(j,k) | Uncertainty in 1064-nm particulate extinction, $\Delta\sigma_p(r)$ ( $\text{km}^{-1}$ )                       |

2.4.3.4. Flowchart Annotations

This section provides a guide to the annotations used in the various HERA flowcharts.

*Layer Descriptor Variables Used in Flow Charts*

- nf = number of features in block
- moflag(nf) = opacity flag
  - = 0 (surface return- do NOT process)
  - = 1 (partially transmissive, unsuitable for constrained retrieval.)
  - = 2 (partially transmissive, suitable for constrained retrieval.)
  - = 3 (opaque, totally-attenuating feature or feature in contact with surface, e.g. a PBL aerosol)
- maflag(nf) = adjacency flag (add bits)
  - = 0 feature is not adjacent & contains no embedded features
  - = 1 feature is adjacent at base
  - = 2 feature is adjacent at top
  - = 3 feature is adjacent at both base and top
  - = 4 feature contains embedded features

*Array Variables related to complex feature structure*

- maxf = maximum number of features in data block
- maxcf = maximum number of complex features

- maxemb = maximum number of embedded features
- cfsize(maxcf) = ordered integer array of number of features in each complex feature.
- cffn(maxf,maxcf) = array of feature numbers for each complex feature in order of increasing feature number (i.e. highest and lowest maflag first.)
- cfgp(maxf,maxcf) = array of feature numbers for each complex feature in order of decreasing gamma-primed (i.e. highest gamma-primed first)
- nembf(maxemb) = number of embedded features in each feature in order of feature number
- embedlist(maxemb,nf) = integer array listing index numbers of the embedded features in each feature in order of increasing feature number.

*Array Variables on Standard Analysis Grid*

The standard analysis grid here is the array defined by `nfine` columns of fine-resolution profiles, each with `nrow` range values. (In the current implementation, `nrow` = 583 and `nfine` or `ncol` = 16.)

- bt2(nrow) = copy of attenuated backscatter for retrieval algorithm
- dbt2(nrow) = copy of uncertainty in above for retrieval algorithm
- sigma(nrow,ncol) = sigma retrieval temporarily mapped to grid
- sigma(nrow,ncol) = uncertainty in above item
- map(nrow,ncol) = integer array assigning each element to a feature number, or to zero if no feature, -1 for invalid atmospheric data

where:

- nrow = number of range values (`nvals`) in array
- ncol = number of fine-resolution profiles (columns) in array

*Logical Variables Used in Flow Charts*

- iterate = If TRUE iterate solution to match constraint.
- iterateC = If TRUE iterate complex feature retrieval to achieve consistency with measured average optical thickness of complex feature.
- match532 = If TRUE replace 1064-nm cloud profile with 532-nm profile (not implemented in current version of algorithm).
- done(j) = If TRUE feature j has been analyzed.

## 2.4.4. Description of the Steps Shown in the Flow Charts

### 2.4.4.1. Simple features

Figure 2.10 shows the procedure followed for simple features and for the clear-air regions that lie above them. Each cycle of the loop defined by steps 7 to 24 completes the analysis of a clear region and a feature.

1. Read the settings specified in the run-time file.
2. Read the Meteorological data.
3. Read the Layer Descriptor data.
4. Generate the feature map on the standard data grid (range (i.e. row), columns).
5. Generate the complex feature and embedded data and make a list of all the features in the scene.
6. Read the whole (currently 80 km) block of lidar data consisting of 16 profiles at 5 km horizontal resolution. Assign data to the attenuated backscatter and uncertainty arrays. Initialize parameters where required. Initialize top row to and feature number zero.
7. Begin processing loop. The steps 7 – 24 form a loop with the index incremented from 1 to  $nf$ , the number of features in the block. Assign the bottom row for analysis initially to the level of the last valid atmospheric signal in any column.
8. If the number of features is not greater than zero, or if all the features have been processed (all “done”) then bypass the following tests and go straight to step 9.1. Otherwise check to see if feature  $jf$  is a member of a complex feature AND if any members of that complex feature are beneath another feature, which is itself not a member of the same complex feature, AND which is yet to be solved. If so, then select the next feature  $jf$  in the feature list  $flist$  and repeat the test until the new feature is acceptable
9. The bounds of the clear region above the current feature are identified. Note that the topmost clear region may be divided into a stratospheric region and a tropospheric region, each with a different lidar ratio for background aerosols. The regions should be analyzed separately.
  - 9.1 The lower bound of the clear region is the altitude of the highest base (lowest index  $mbase$ ) of all unsolved features ( $done = FALSE$ ) that are in contact with a clear region in at least one column at their base.
  - 9.2 If the top row of the region is not greater than the bottom row, then there is no clear region to solve so move straight to solving next feature at step 14.
  - 9.3 Initialize retrieval arrays to zero, set renormalization indicator  $renorm(col)$  for each column to  $FALSE$ .
10. The data in the clear region are averaged to form profiles of attenuated backscatter, and uncertainty at the coarsest horizontal resolution. Set  $renorm(col) = TRUE$  for each column containing clear points within vertical limits of region.
11. Using the just-calculated profile of the number of averaged columns at each row of the averaged clear profile, identify those regions for which the SNR is sufficient for retrieving reliable results. Separate stratospheric and tropospheric regions.

12. (12.0 to 12.7) describe the process of solving the acceptable sub-regions ( $n_{\text{regions}}$ ) of the current clear profile.
  - 12.0 Initialize the particulate transmittance to unity and the region counter  $j_{\text{region}}$  to zero.
  - 12.1 Loop through the sub-regions
  - 12.2 Select the index of the calibration (normalization) range depending on the solution direction, and select the normalization factor according to the total overlying transmittance and other factors. The top-most clear region will possibly require retrieval in the backward direction.
  - 12.3 Adjust the re-normalization factor  $c_{\text{cal}}$  to allow for the transmittance of overlying sub-regions.
  - 12.4 Call the profile extinction solver with the appropriate parameters to calculate profiles of backscatter, extinction and effective optical thickness ( $\eta\tau$ ), and their uncertainties. Rescale data for calibration correction if the region is the top region of the atmosphere with its base above the tropopause and if the retrieved optical thickness is significantly greater than its uncertainty in absolute terms.
  - 12.5 Update the cumulative transmittance loss by scaling by that of the just-solved sub-region.
  - 12.6 If all sub-regions have been processed, end analysis of the clear profile.
  - 12.7 Copy the temporary output profiles to all rows and columns in the output arrays where the feature map indicates that there are no features.
13. For each column, calculate the square of the transmittance of the clear region,  $T^2(k)$  and divide the attenuated backscatter for all underlying regions by this quantity.

This process is illustrated in steps 13.1 to 13.8. Note that, although the clear profile just solved is the average of all (up to 16) columns,  $T^2(k)$  will vary from column to column depending on the vertical extent of the clear region in each column.
14. Check the value of the adjacency flag for the feature with the current loop index.
15. If the adjacency flag for the feature is zero, the feature is a simple feature and has neither vertically adjacent nor embedded features. Carry out the analysis of the simple feature first at 532 nm then at 1064 nm, looping through the steps 15 to 20 for each wavelength. Define the limits of the current feature as determined by the map array. Average those points where the value of the map array is the same as the loop index ( $\text{map}(j,k) = jf$ ) to form profiles of mean attenuated backscatter and uncertainty.
16. Recalculate the layer descriptors  $\gamma'(j)$ ,  $T^2(j)$  and their uncertainties for the current feature from the corrected attenuated backscatter data where this is necessary.
17. Select the index of the calibration (normalization) range depending on the solution direction, wavelength and adjacency, and select the normalization factor according to the total overlying transmittance. Note that all features in complex features will be solved in the forward direction, but some simple features may be solved in the backward direction. Get value of  $S_p$  and uncertainty after considering  $T^2$ ,  $\gamma'$  and overlying attenuation. These values should have been corrected in step 16.

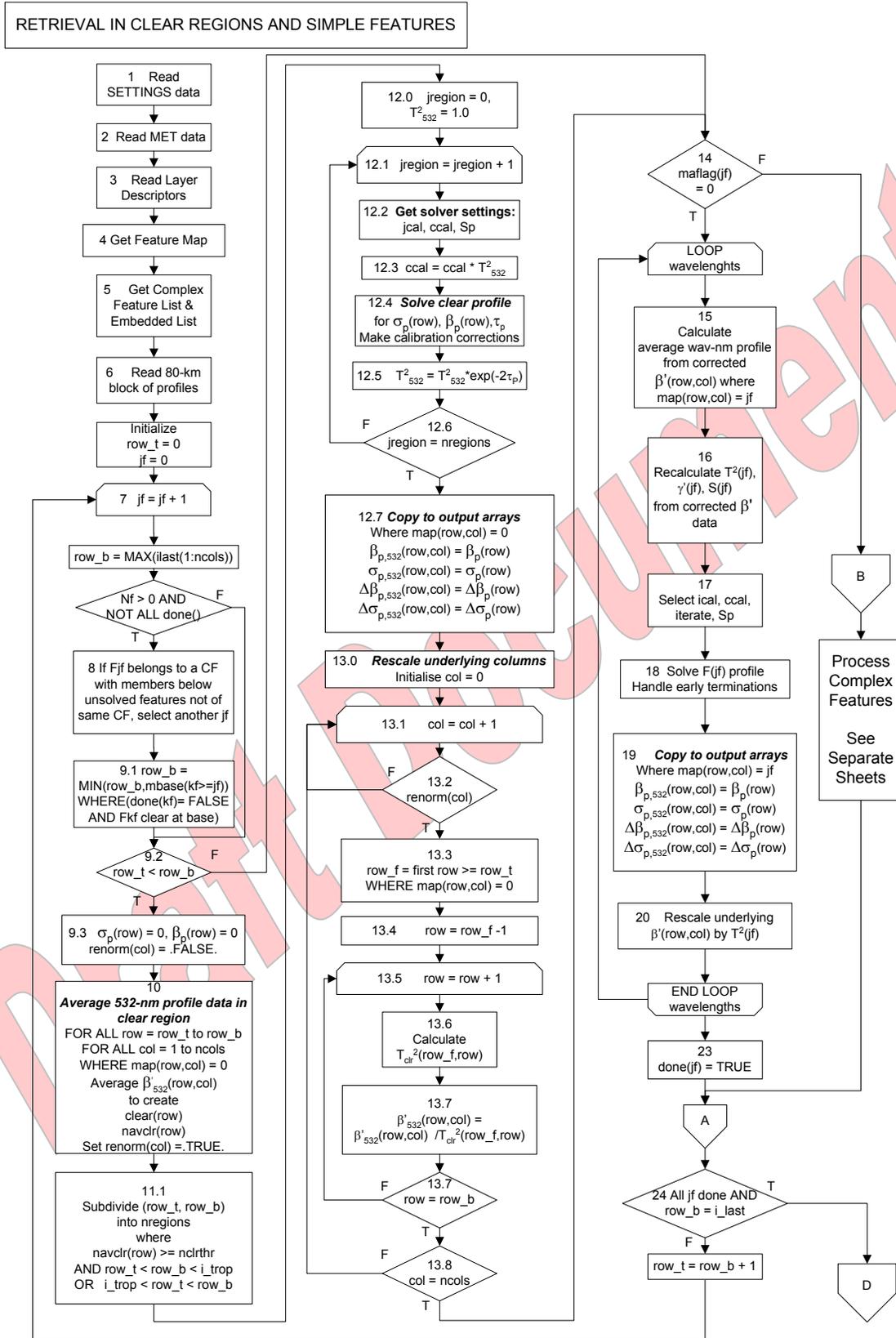


Figure 2.10: Retrievals in simple features and clear regions

18. Call the profile extinction solver with the appropriate parameters to calculate profiles of backscatter, extinction and effective optical thickness ( $\eta\tau$ ) and their uncertainties. If the solution attempt terminates before base reached, fill values are assigned beyond the termination range. All underlying **map** data are set to -1 to indicate to averaging routines that data in this region are not to be used.
19. As in step 12.7, copy the temporary output profile to all rows and columns in the output array where the feature map indicates that feature **jf** is present.
20. For each column occupied by the feature **jf**, divide the attenuated backscatter for all underlying regions by the square of the feature transmittance. Update the uncertainties of the underlying data using equation (2.23).
21. Set the flag **done(jf)** to indicate that the current feature has been analyzed satisfactorily.
22. The last step in the loop. If all the features have been analyzed and the current lower row limit is the lowest valid row, **i\_last**, (i.e. if all valid rows have been analyzed), exit to point D. If not set the top row of the next region to be processed to the row below the last one processed.
  - A. The branch describing the analysis of complex features, which separated at step 14, returns here.
  - B. The branch describing the analysis of complex features begins here.

#### 2.4.4.2. Complex features

The flow chart in Figure 2.11 is organized into three main tasks.

The first column (steps 30 to 44) illustrates the process by which it is decided whether a suitable consistency check can be made on the solution or not. This is done by measuring the mean signal attenuation through the complex feature using the transmittance method (comparing the signals above and below the feature). If the measured mean optical thickness of the complex feature, **taucfm**, is negative, then no consistency check is possible and the **iteratec** variable is set to **FALSE**. The rest of the column is concerned with identifying the complex feature, getting the number and indices of its members, and analyzing each one in order of decreasing altitude. Unlike the case for simple features, where each individual feature is analyzed at 532 nm then immediately afterward at 1064 nm, with the analysis of complex features, all the features in the current complex feature are processed (in steps 34 to 62) at 532 nm, and a consistent result achieved (if required), before the complex feature is processed at 1064 nm.

Once all features have been analyzed, a check is made to see if the retrieval has produced an overall attenuation through the complex feature that is consistent with the measured value. If it is not, then the lidar ratios of the various features are adjusted, one at a time and the complete complex feature is reanalyzed. The lidar ratios are adjusted in order of the decreasing integrated attenuated backscatter of the feature to which they belong. This process is illustrated in the block at the top centre of the figure (steps 45 to 55).

Finally, once all the features have been analyzed and the lidar ratios adjusted as required, the underlying attenuated backscatter profiles in each column are rescaled by the overlying

attenuation in the complex feature in that column. This process is illustrated at the bottom right of the figure in steps 56 to 62. Next, the `done` flag is set to `TRUE` for each feature that has been processed (steps 63 to 65) before returning to the feature loop in Figure 2.10.

B. Begin Complex feature Analysis branch.

30. Set the complex feature index to the value specified in the layer descriptor for the current feature. The array `mcfn` specifies the complex feature to which the current feature belongs. Initially set the logical variable `iteratec` to `FALSE`
31. Calculate the optical thickness `taucfm` (and its uncertainty, `dtaucfm`) of the complex feature by using the transmittance method and the clear regions above and below the complex feature. In each column, convert the  $T^2$  to  $\tau$  and average the  $\tau$ s.
32. Test the magnitude of the average optical thickness `taucfm` measured in step 31.
33. If `taucfm` is greater than zero, set the logical variable `iterateC` to `TRUE`. This decision, when combined with calculations of the complex feature optical thickness and its uncertainty in step 46, will permit a decision to be made in step 47 to force an iteration to attempt a retrieval that produces an effective optical thickness that is consistent with the value measured in step 31. Once the decision has been made to iterate the retrieval or not, solve the whole complex feature (steps 34 to 62) at 532 nm, until a consistent `tau` is achieved, then repeat at 1064 nm.
34. Get the number of features, `mcf`, in the current complex feature, `jcf`. Initialize the counter for the number of tries at a consistent solution to unity. Initialize the index of the feature, `mf`, whose lidar ratio is to be adjusted to the one with the largest integrated attenuated backscatter.
  - 34(b) Reset all values of the temporary array `stemp(row,col)` to zero.
35. Begin looping through the features in the complex feature, varying index `kf` from unity to the number of features in this complex feature, `mcf`.
36. Get the index `lf` for the next feature to be analyzed in the current complex feature.
37. Check if the feature `lf` has already been analyzed or `lf` is in the embedded list of feature `jcf`.
38. Check the adjacency flag of the feature `lf`. If the adjacency flag has a value greater than 1 and the feature spans more than one column, then the feature has embedded features, so proceed to route C to D that is detailed in Figure 2.12.
39. If the adjacency flag has a value not greater than 1, then the feature has no embedded features. Correct the attenuated backscatter array for regions containing the current feature, for the attenuation effects of features that are wholly above, or intrude into, its top using values in the array `stemp`.
40. Correct the attenuated backscatter array for regions containing the current feature, for the attenuation effects of features that are totally included within its boundaries, carrying out the calculation from the top of the current feature so as not to duplicate the effects of step 39.

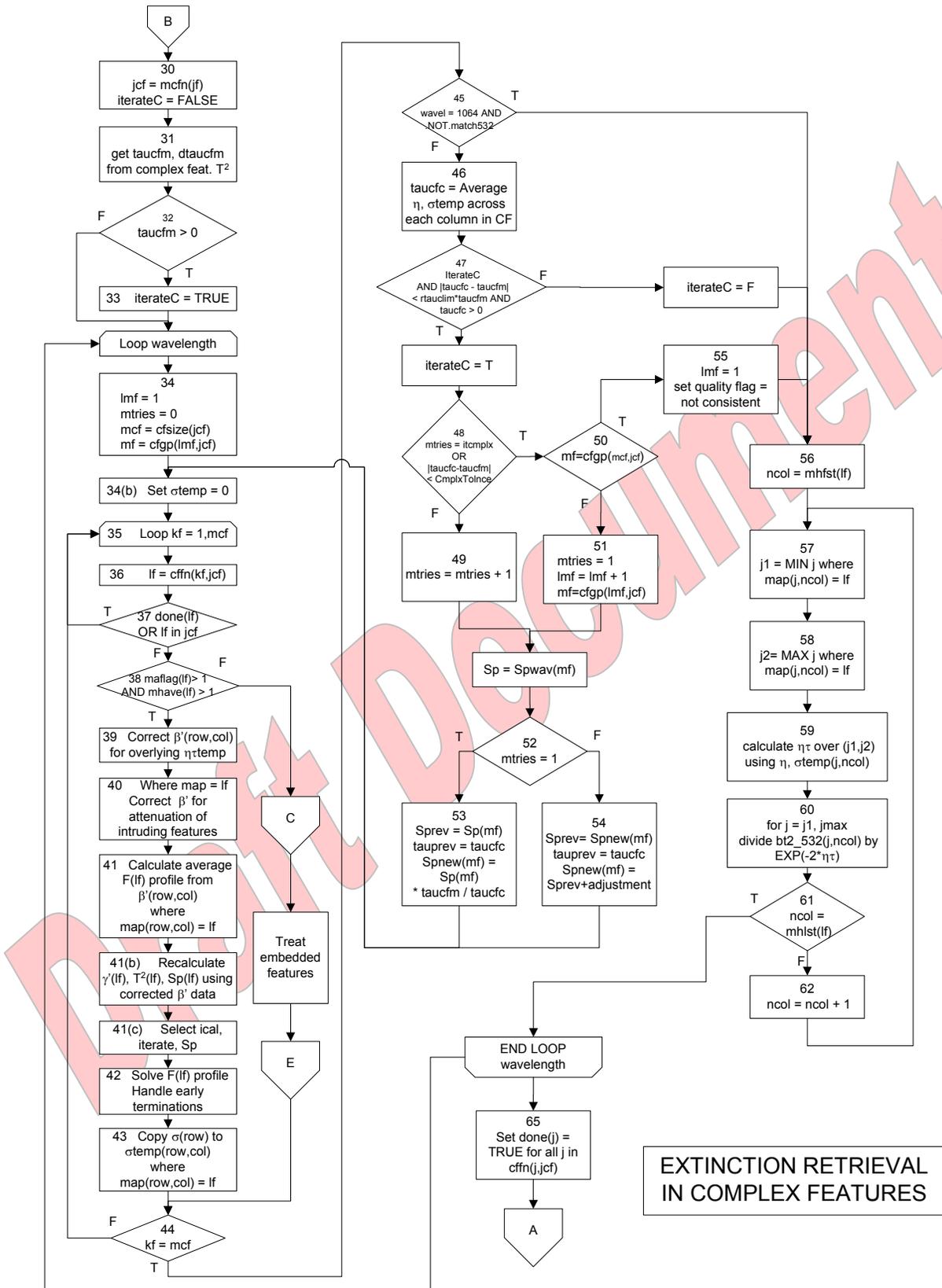


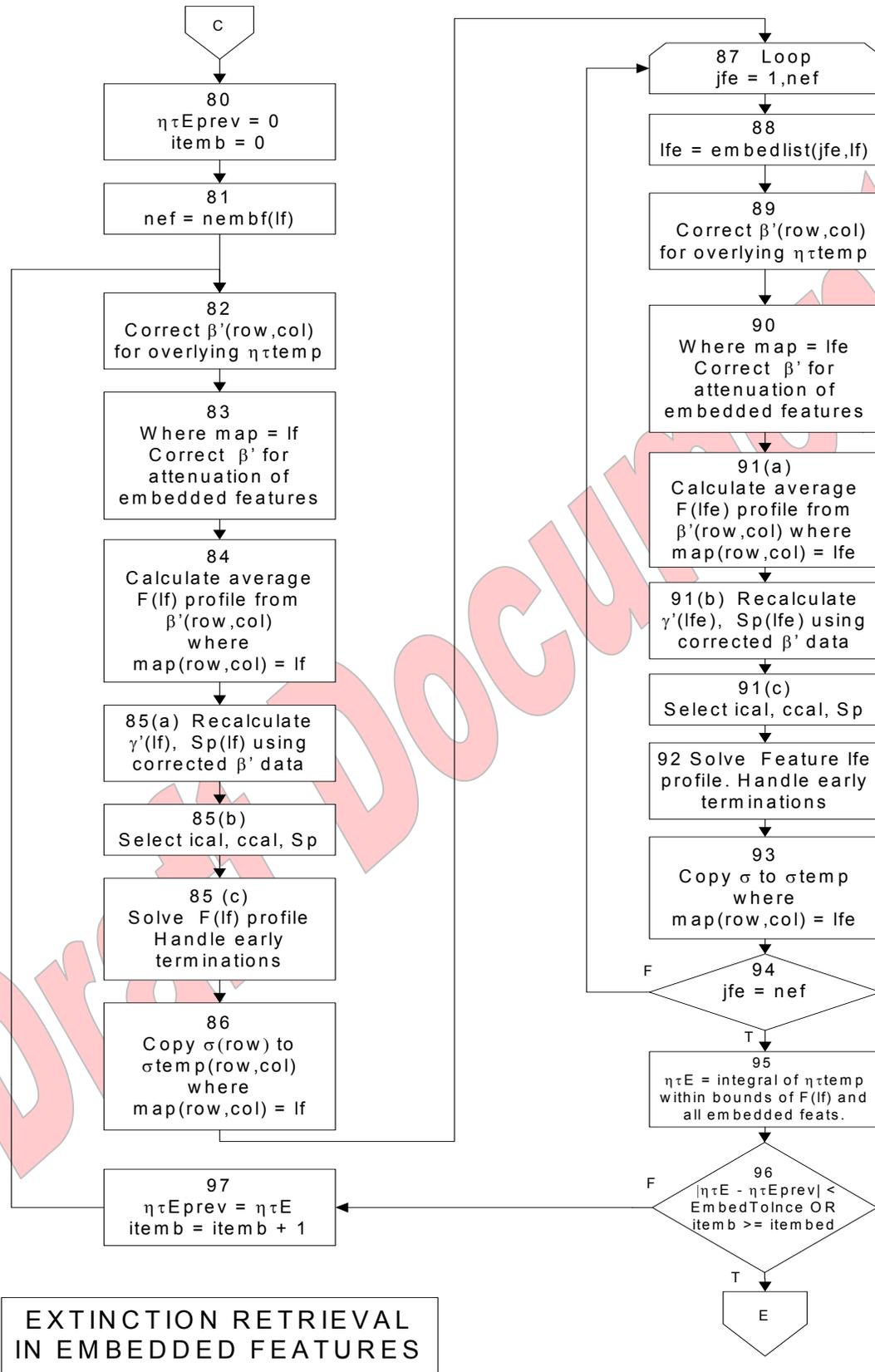
Figure 2.11: HERA retrievals for complex features.

41. In each column, average those points where  $\text{map}(\text{row}, \text{col}) = \text{lf}$  to form profiles of attenuated backscatter and uncertainty.
  - 41(b) Recalculate the layer descriptors  $\gamma'(\text{lf})$ ,  $T^2(\text{lf})$  and their uncertainties for the current feature (lf) from the corrected attenuated backscatter data where this is necessary.
  - 41(c) Select the index of the calibration (normalization) range and select the normalization factor according to the total overlying transmittance. (See step 17 above.) Get the value of  $S_p$  and uncertainty, after considering  $T^2$ ,  $\gamma'$  and the overlying attenuation.
42. Call the profile extinction solver with the appropriate parameters to calculate profiles of backscatter, extinction and effective optical thickness ( $\eta\tau$ ), and their uncertainties. If the solution attempt terminates before base reached, fill values are assigned beyond the termination range. All underlying map data are set to -1 to indicate to averaging routines that data in this region are not to be used. Also set `iterateC` FALSE as consistency check no longer valid.
43. Copy the calculated profile of  $\sigma$  to the *temporary* array of values where indicated by the current index. Note that a temporary array is used because the solution may be iterated.
44. If all the features in the complex feature have been processed, proceed to the consistency checks, otherwise continue working through the features in the list `cffn`.
45. If all the features in the complex feature have been processed, then test to see whether an iteration is required (`iterateC` is TRUE). First, check the wavelength. If the wavelength is 1064 nm and the `match532` switch (Table 2.1) is FALSE, then bypass the iteration and proceed to the block (beginning at step 56) where underlying data are rescaled before returning to the end of the loop at A in Figure 2.10.
46. If iteration is required, calculate the average effective optical thickness (`taucfc`) and its uncertainty, in the complex feature just processed. This is the mean across all the columns occupied by the current complex feature, of the effective optical thicknesses  $\eta\tau$  calculated by integrating `sigmatemp(row,col)` with respect to `row` in each column. Note that the integration is actually with respect to range, so the variation of the range increment with range needs to be included. For the situation of two adjacent features, F1 between  $r_1$  and  $r_2$ , and F2 between  $r_2$  and  $r_3$ , the effective optical thickness between  $r_1$  and  $r_3$  is calculated as
$$\eta\tau(r_1, r_3) = \eta_1(r_2)\tau_1(r_1, r_2) + \eta_2(r_3)\tau_2(r_2, r_3).$$
The optical thickness of one feature is scaled by the value of its multiple scattering function at its base. This is added to the same quantity for the second feature. Note that work on multiple scattering in complex features is still in progress and this definition may be updated in a later release of the document.
47. Test whether the calculated complex feature optical thickness (`taucfc`) agrees with the measured complex feature optical thickness (`taucfm`) within the relative tolerance specified in the run-time file multiplied by `taucfm` AND if `taucfc` is greater than zero AND if the `iterateC` logical variable has previously been set TRUE. If these tests fail, exit to step 65.
48. If the tests in 47 are successful, test to see if the maximum number of tries has been reached OR if changes in `taucfc` are less than the tolerance specified in the run-time file.
49. If the tests in 48 are unsuccessful, then increment the number of tries.

50. If the tests in 48 are successful, test to see if the index of the feature whose  $S_p$  has been chosen for iteration is the final feature (in order of decreasing gamma-primed) in the complex feature. If so, reset the index  $l_{mf}$  to unity and proceed to step 55.
51. If the feature is not the one with the smallest gamma-primed, change the index to the feature with the next smallest gamma-primed, and reset the number of tries to unity. Note that the iteration will proceed to a maximum of  $it_{cmplx}$  for every feature in the complex feature.
52. Get the  $S_p$  for the new feature. Test if this is the first try in the current feature.
53. For the first try, adjust the (new) lidar ratio for the appropriate feature by the ratio of the measured and calculated optical thicknesses. Set the previous value of lidar ratio (used in calculating the factor in step 54) to the original value. Return to the start of the loop at step 34(b) and reprocess the complex feature using the updated lidar ratio for the selected feature. Also set  $\tau_{auprev} = \tau_{aucfm}$  to allow scaling in subsequent tries.
54. For subsequent tries, adjust the lidar ratio using the “secant” method:  
$$\text{Slope} = (S_{pnew}(mf) - S_{prev}) / (\tau_{aucf} - \tau_{auprev})$$

Then redefine  $S_{prev} = S_{pnew}(mf)$   
 $\tau_{auprev} = \tau_{aucf}$  for the next calculation.

$$S_{pnew}(mf) = S_{prev} + (\tau_{aucfm} - \tau_{auprev}) * \text{slope}.$$
55. If the adjusted lidar ratio is for the feature with the smallest gamma-primed (i.e. all the features in the complex feature have been adjusted and the solution is still not consistent) then set the quality flag for the retrieval to indicate an inconsistent result and return to **A** via the rescaling block (steps 56 to 62).
56. Start procedure for rescaling underlying attenuated backscatter profiles. Note that this must be done column by column because the attenuation of the complex feature varies column by column. Set the current column number to the first column of feature  $l_f$ .
57. Steps 57 – 62 form a loop in which the underlying features in each column of feature  $l_f$  are corrected. Get the first range index in the current column for feature  $l_f$ .
58. Get the final range in the current column for feature  $l_f$ .
59. The attenuation is calculated from  $\sigma_{matemp}(row,col)$  between the range indices ( $j_1, j_2$ ) defined for the current column in steps 57 and 58, correctly scaling for the  $\eta$  of the features as described in 46 above.
60. Divide the attenuated backscatter profile, in the region below the base of feature  $l_f$  down to the last range, by the attenuation due to feature  $l_f$  in current column calculated in the previous step.
61. Check if the current column is the last column in feature  $l_f$ .
62. If the last column has not been reached, increment the column index and return to step 57.
65. If all columns have been corrected, set the **done** flag to **TRUE** for all the features in the complex feature as listed in  $c_{ffn}$ , then return to processing loop in Figure 2.10 via **A**.



**EXTINCTION RETRIEVAL  
IN EMBEDDED FEATURES**

Figure 2.12: HERA retrievals in embedded features

#### 2.4.4.3. Complex embedded features

The additional complexity associated with embedded features is that neither the embedded feature nor the larger feature can be solved directly, as each of the features causes some attenuation in the other. The solution of this problem is achieved by solving the outer feature first, then correcting the attenuation down to the top of the inner, embedded, feature using the solution of the outer feature. The embedded feature is then solved and its calculated attenuation is then used to correct the lower region of the outer feature. The solution of the whole complex feature is iterated using the updated attenuation corrections at each stage until the calculated optical thickness of the whole complex feature converges. The required number of passes through a feature containing embedded features is a function of the number of embedded features in any common column of the outer feature and the *degree* of embeddedness. (It is possible for an 80 km feature, occupying sixteen columns, to have an embedded 20 km feature, occupying four columns, that contains one or more embed features of 5 km resolution each occupying one column.) The procedure is shown in the flowchart depicted in Figure 2.12.

C. This is the point at which the embedded feature branch leaves the main loop in Figure 2.11.

80. The overall attenuation (effective optical thickness  $\eta\tau_{Eprev}$ ) of the complex feature is initialized to zero.
81. Between steps 81 and 96 the solutions of the outer feature and its inner, embedded features are iterated until the total, embedded effective optical thickness converges. The outer feature is solved in steps 81 to 86. First the number of embedded features  $n_{ef}$  in feature  $lf$  is read from the array  $n_{emb}$ .
82. Correct the attenuated backscatter array for regions containing the current feature,  $lf$ , for the attenuation effects of features that are either wholly above its top, or intrude into the feature from above, using values stored in the array  $\sigma_{matemp}$ .
83. Correct the attenuated backscatter array for regions containing the current feature, for the attenuation effects of features that are embedded in the feature, using values stored in the array  $\sigma_{matemp}$ , carrying out the calculation so as not to duplicate the effects of step 82. These corrections will be zero on the first iterative loop.
84. Calculate the average profile for  $lf$  by averaging the corrected attenuated backscatter data in columns and rows where  $map(row,col) = lf$ .
85. (a) Recalculate the layer descriptors  $\gamma'(lf)$ ,  $T^2(lf)$  and their uncertainties for the current feature ( $lf$ ) from the corrected attenuated backscatter data where this is necessary.
  - 85(b) Select the index of the calibration (normalization) range and select the normalization factor according to the total overlying transmittance. (See step 17 above.) Get value of  $S_p$  and uncertainty after considering  $T^2$ ,  $\gamma'$  and overlying attenuation.
  - 85(c) Call the profile extinction solver using the averaged profile as an input. If the solution attempt terminates before base reached, fill values are assigned beyond termination range. All underlying  $map$  data are set to -1 to indicate to averaging routines that data in this region are not to be used. Also set `iterateC` FALSE and `itemb = itembed` as consistency check no longer valid.

86. Copy the values of extinction to the temporary array `sigmatemp(row,col)` in all the columns where `map(row,col) = lf`.
87. Steps 87 to 94 form a loop in which each of the `nfe` features embedded in feature `lf` is solved.
88. Read the index of the embedded feature `lfe` corresponding to position number `jfe` in the list of embedded features `embedlist(jfe,lf)`.
89. Correct the attenuated backscatter array for regions containing the current feature, `lfe`, for the attenuation effects of features that are either wholly above its top, or intrude into the feature from above, using values stored in the array `sigmatemp`.
90. Correct the attenuated backscatter array for regions containing the current feature, for the attenuation effects of features that are embedded in the feature, carrying out the calculation so as not to duplicate the effects of step 89.
91. (a) Calculate the average profile for `lfe` by averaging the corrected attenuated backscatter data in columns and rows where `map(row,col) = lfe`.
  - 91(b) Recalculate the layer descriptors  $\gamma'(lfe)$ ,  $T^2(lfe)$  and their uncertainties for the current feature (`lfe`) from the corrected attenuated backscatter data where this is necessary.
  - 91(c) Select the index of the calibration (normalization) range and select the normalization factor according to the total overlying transmittance. (See step 17 above.) Get value of `Sp` and uncertainty after considering  $T^2$ ,  $\gamma'$  and overlying attenuation.
92. Call the profile extinction solver using the averaged profile as an input. If the solution attempt terminates before base reached, fill values are assigned beyond the termination range. All underlying `map` data set to -1 to indicate to averaging routines that data in this region are not to be used. Also set `iterateC FALSE` and `itemb = itvar` as consistency check no longer valid.
93. Copy the values of extinction to the temporary array `sigmatemp(row,col)` in all the columns where `map(row,col) = lfe`.
94. If all the embedded features have not been analyzed, continue looping.
95. If all the embedded features have been processed, calculate the effective optical thickness,  $\eta\tau E$ , for the complex, embedded feature by integrating the temporary array `sigmatemp` scaled appropriately by the multiple scattering functions of the features in the columns, over the vertical extent of feature `lf` and all its embedded features, and averaging over the horizontal extent of feature `lf`.
96. If the current value of  $\eta\tau E$  agrees with the previous value,  $\eta\tau E_{prev}$ , within a tolerance that is the greater of that calculated from the computer precision or that specified in the runtime file (`EmbedTolnce`) calculated from the computer precision, or if the number of iterations, `itemb`, has exceeded the limit, `itembed`, then return to the main processing loop in Figure 2.11 via `E`.
97. If the current value of  $\eta\tau E$  does not agree with the previous value,  $\eta\tau E_{prev}$ , set the value of  $\eta\tau E_{prev}$  to the new value,  $\eta\tau E$ , and return to step 82 and continue iterating the solution.

### 3. Accounting for Multiple Scattering in Extinction Retrievals

#### 3.1. Introduction

The form of the extinction solution adopted for CALIOP is based on a solution of the classic lidar equation, which is a single-scatter approximation. However, all lidar measurements contain multiply-scattered components in addition to the primary, single-scattered signal (Winker, 2003). Multiple scattering can alter the apparent extinction and transmittance of the medium, can produce depolarization of the return signal, and can produce “stretching” of the return pulse (Chepfer et al., 1999; Hu, 2007; Miller and Stephens, 1999). For most lidar systems the magnitude of the multiply-scattered signal is so small these effects are insignificant and can be ignored without introducing significant errors. Multiple scattering effects can be very significant for satellite lidar systems, however, due to the large footprints typical of space geometries. The diameter of the CALIOP laser footprint is two orders of magnitude larger than that of the typical ground-based or airborne lidar, allowing a much greater fraction of the multiply-scattered light to contribute to the return signal. Because of this, the extinction retrieval algorithm must explicitly account for multiple scattering effects on the return signal in at least some situations.

The nature of multiple scattering is fundamentally dependent on the scattering phase function of the scattering particles and the sensing geometry of the lidar. The extinction cross-section and homogeneity of the scattering medium also play a role. When the scattering medium becomes highly turbid and the scattering approaches the diffusion regime, multiple scattering overwhelms the single-scattered component of the return signal and pulse stretching becomes significant. To fully understand the multiple scattering issue, a variety of scenarios have been considered.

The multiple scattering analyses shown below were performed using a Monte Carlo simulation code developed to investigate lidar multiple scattering issues (Winker and Poole, 1995). The Monte Carlo approach allows the full physics to be incorporated and can accommodate any sensing geometry. Lidar multiple scattering effects can vary dramatically in nature and magnitude depending on the lidar sensing geometry and the characteristics of the ‘target’ being sensed. General principles will be discussed within the context of retrieving aerosol extinction. Following that, issues specific to cirrus and other ice clouds will be discussed. Some general considerations are:

*Aerosol* – aerosols are tenuous,  $\sigma < 1 \text{ km}^{-1}$ , except perhaps very near sources, and have broad scattering phase functions.

*Ice Clouds* – In ice clouds, the ratio of multiple scattering to single scattering is relatively small because the scattering phase function is sharply peaked in the forward direction. Multiple scattering effects, for CALIOP, tend to be much more significant than in aerosols. The tops of deep convective clouds are composed of ice crystals and are usually weakly attenuating. The lidar signal often penetrates to an optical depth of 3 or more, which is often several kilometers below cloud top.

*Water Clouds* – In water clouds the signal is rapidly attenuated, often within 100 meters or less. Multiple scattering builds up rapidly in dense boundary layer clouds and is dependent on both range and extinction coefficient. The current multiple scattering parameterization described in this document is not well suited to water clouds. The CALIOP extinction retrieval is applied to

water clouds, as well as to ice clouds and aerosol layers, but as the assumptions of the retrieval are violated the results for water clouds are invalid and should not be used.

More discussion on the effects of multiple scattering in the CALIOP measurements can be found in Winker (2003).

## 3.2. Approach

The general approach to parameterizing and analyzing multiple scattering effects will be discussed in the context of scattering in homogeneous aerosol layers, as well as issues specific to aerosols. This approach will then be compared and contrasted to extinction retrievals of ice clouds. Generally speaking, pulse stretching due to multiple scattering is insignificant in aerosol and cirrus layers, due to the nature of the scattering phase function of aerosols and ice crystals and the fact that the mean free path is much greater than the lidar footprint diameter (100 meters). When pulse-stretching is not significant, multiple scattering effects can often be effectively parameterized in terms of a modified extinction cross-section.

### 3.2.1. Specifics

#### Aerosol

The solution for the aerosol backscatter coefficient,  $\beta_a(r)$ , is parameterized in terms of  $S^* = \eta S_a$  so that we only need to know  $S^*$  to perform the inversion. To obtain the extinction profiles,  $\sigma_a(r)$ , however, we need  $S_a$  to convert backscatter to extinction.

We have developed aerosol models (discussed in Part 3) which allow the scattering phase function and the aerosol absorption to be computed. From these, we compute  $S_a$  and the scattering phase function,  $\mathbf{P}(\theta)$ , which is necessary to compute multiple scattering effects. Selection of the aerosol model is performed by the CALIOP Scene Classification Algorithm. Choosing the aerosol model, then, specifies  $S_a$  and  $\mathbf{P}(\theta)$ . The multiple scattering factor,  $\eta(r)$ , depends on a number of parameters. Many of them are effectively fixed however: wavelength, distance to the layer, field of view, and laser beam divergence. For aerosols (and ice crystals)  $\eta(r)$  is only weakly dependent on the magnitude of the extinction coefficient, so that only dependencies on penetration depth into the layer and the aerosol type need to be considered. Multiply-scattered 532 nm return signals were computed for a variety of aerosol models using a Monte Carlo code allowing 40 orders of scattering. As illustrated in Figure 3.1, the total signal strength and the magnitude of the multiply-scattered components of the signal are a function of the aerosol extinction coefficient.

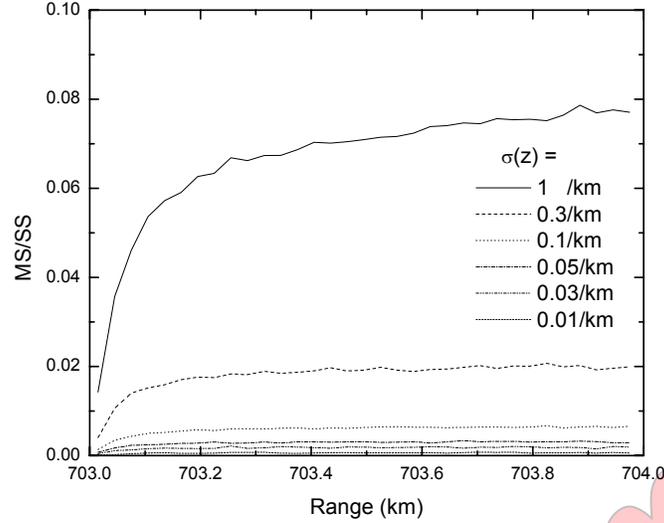


Figure 3.1: Multiple scattering for a uniform aerosol layer located between 703 km and 704 km range.

The total return signal  $P_{TS}(r)$  - which is the sum of singly- and multiply-scattered signal components – is written in terms of  $\eta(r)$ :

$$P_{TS}(r) = \frac{C}{r^2} \beta(r) \exp(-2\eta(r)\tau(r)) \quad (3.1)$$

This expression reduces to the expression for the single-scatter lidar return,  $P_{SS}(r)$ , when  $\eta(r) = 1$ . An advantage of this parameterization is that  $\eta(r)$  can easily be derived from Monte Carlo calculations of the total and single-scatter ( $P_{SS}$ ) return signals:

$$\eta(r) = 1 - \frac{\ln[P_{TS}(r)/P_{SS}(r)]}{2\tau(r_0, r)} \quad (3.2)$$

where  $\tau(r_0, r)$  is the accumulated optical depth between the layer top and range  $r$ . It turns out that  $\eta$  is nearly independent of extinction coefficient, although it is still dependent on range within the layer (see Figure 3.2, which shows results from layers with constant  $\sigma$  ranging from  $0.01 \text{ km}^{-1}$  to  $1 \text{ km}^{-1}$ ).  $TS(r)$  and  $SS(r)$  are the profiles of the total lidar backscattered return signal and the singly-scattered return signal, respectively.

For deep aerosol layers,  $\eta(r)$  approaches unity, so multiple scattering has little effect on the optical depth retrieved. In this case the primary effect of multiple scattering is to reduce the apparent extinction at the top of the layer and alter the shape of the extinction profile.

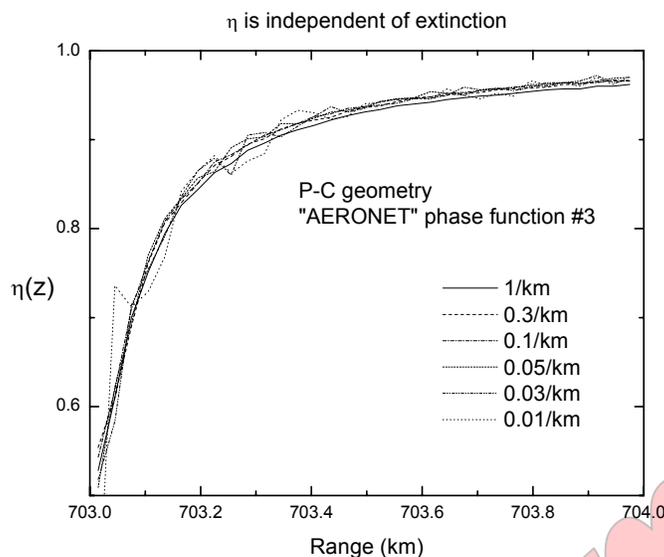


Figure 3.2: Multiple scattering parameter for a uniform aerosol layer located between 703 km and 704 km range, parameterized in terms of  $\eta(r)$ .

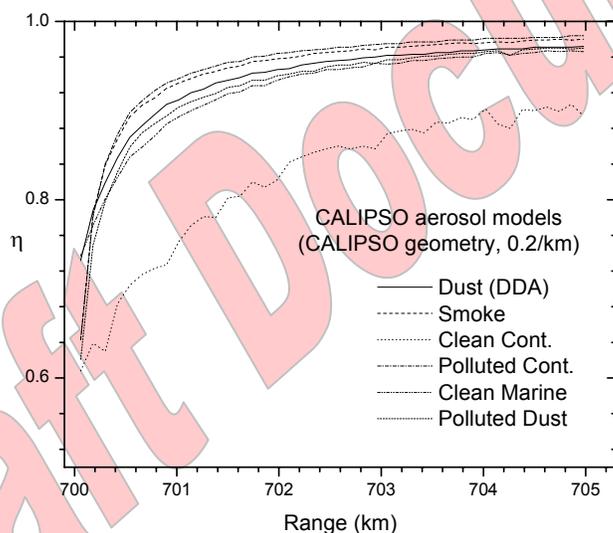


Figure 3.3: Multiple scattering parameter for a uniform aerosol layer located between 700 km and 705 km range, for the 6 aerosol types derived from AERONET measurements.

Figure 3.3 shows  $\eta(r)$  for the current set of 6 CALIOP aerosol models (see CALIOP ATBD Part 3), all for an extinction cross-section of  $0.2 \text{ km}^{-1}$ . The figure shows little variation in  $\eta$  between models, except for the background “clean continental” model. The definition of this model is uncertain, so it is not clear if this multiple scattering function is realistic. In any case, the aerosol optical depth (AOD) in clean continental conditions is low, so any errors would tend to be low in terms of absolute AOD.

In the interest of simplicity, for the initial release of the CALIOP extinction data products,  $\eta$  has been set to unity for all aerosol layers. From Figs. 3.2 and 3.3, it can be seen that for layers thicker than a few hundred meters the effect of multiple scattering on the layer AOD is less than 10% and for very thick layers the effect on AOD becomes insignificant.

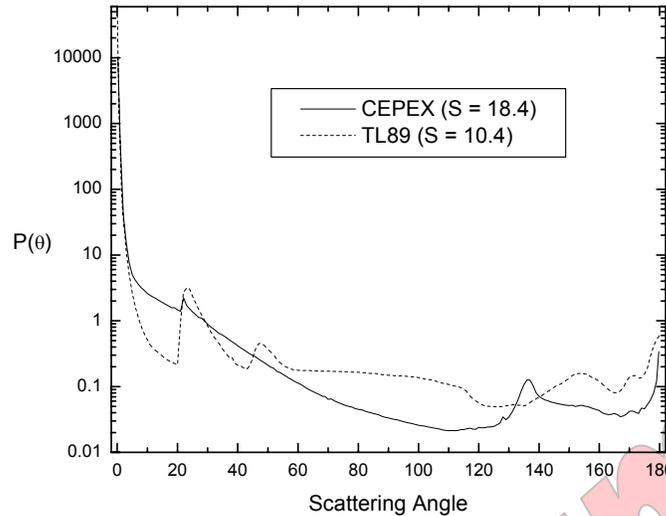


Figure 3.4: Cirrus phase functions for tropical convective cirrus (CEPEX) and polar cirrus (TL89).

### Ice Clouds

Cirrus clouds are composed of ice crystals having a wide variety of shapes and sizes. To investigate the potential variability of  $\eta(r)$  in ice clouds, two extreme cases are considered: tropical convective cirrus and polar cirrus. Phase functions derived from *in situ* measurements of vertical profiles of crystal size and shape within a tropical cirrus anvil are available from the CEPEX experiment. Phase functions were computed for several altitudes within the cloud using size-dependent combinations of irregular aggregates, bullet rosettes, hexagonal columns, and spheres (McFarquhar, et al., 1999). The dependence of the phase functions on height within the cloud was found to be small in the CEPEX measurements. A phase function computed from measurements within the cloud is shown in Figure 3.4. Also shown in Figure 3.4 is a phase function representative of cold cirrus in the polar regions taken from Takano and Liou (1989), who based their calculation on regular hexagonal columns and a size distribution measured in cirrostratus. The crystal aspect ratio was varied with size according to observed behavior. It can be seen that the anvil cirrus, which is dominated by irregular particles, has a higher lidar ratio, weaker halo features, and significantly different side scattering compared to the regular hexagonal columns of the polar cirrus.

Measurements of cirrus lidar ratio (Sassen and Comstock, 2001; Eloranta, et al., 2001) show a range of values, with a mean value of about 25 sr and the majority of the values falling between 12 sr and 50 sr. The lidar ratios of the phase functions shown in Figure 3.4 are low relative to these measured values. Ray tracing calculations show that  $S_c$  increases significantly for hollow crystals (Takano and Liou, 1989) and for irregular crystals (Macke, et al., 1996) relative to solid, regular crystals due to suppression of the backscatter. We can hypothesize that this is responsible for the higher values observed experimentally versus modeled values. Figure 3.5 shows the multiple scattering functions,  $\eta(r)$ , corresponding to several phase functions: the Takano and Liou phase function, CEPEX phase functions for two different altitudes within a cloud, and a CEPEX phase function with suppressed backscatter to provide a large lidar ratio ( $S_c = 50$  sr). In spite of the significant differences between the four phase functions, the  $\eta(r)$  are

seen to be quite similar and nearly independent of range. The variation in  $\eta(r)$  between different cirrus models is fairly small and suggests that errors in the retrieved extinction due to incorrect selection of the cirrus model will also be small. Thus it appears that, for cirrus, a single multiple scattering factor,  $\eta_c$ , can be used rather than a range-dependent function.

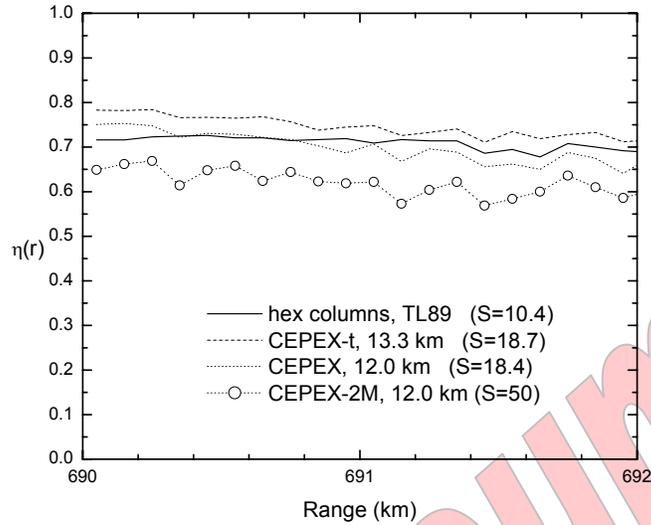


Figure 3.5: Range-dependent multiple scattering functions for cirrus.

Draft Document

## 4. References

- Barrett, E. W., and O. Ben-Dov 1967: “Application of the lidar to air pollution measurements”, *J. Appl. Meteorol.*, **6**, 500.
- Chepfer, H., J. Pelon, G. Brogniez, C. Flamant, V. Trouillet, and P. Flamant, 1999: “Impact of cirrus cloud ice crystal shape and size on multiple scattering effects: application to spaceborne and airborne backscatter lidar measurements during LITE mission and E LITE campaign”, *Geophys. Res. Lett.*, **26**, 2203-2206.
- Davis, P.A., 1969: “The analysis of lidar signatures of cirrus clouds”, *Appl. Opt.*, **10**, 2099-2102.
- Elterman, L., 1966: “Aerosol measurements in the troposphere and stratosphere”, *Appl. Opt.*, **5**, 1769-1776.
- Eloranta, E. W., R. E. Kuehn, and R. E. Holz, 2001: “Measurements of Backscatter Phase Function and Depolarization in Cirrus Clouds made with the University of Wisconsin High Spectral Resolution Lidar”, in *Advances in Laser Remote Sensing: Selected Papers Presented at the 20th International Laser Radar Conference*, Vichy, France, July 2001, A. Dabas, C. Loth, J. Pelon, eds., pp. 255-258 (Editions de l’Ecole Polytechnique, Palaiseau, France).
- Fernald, F. G., 1984: “Analysis of atmospheric lidar observations: some comments”, *Appl. Opt.*, **23**, 652-653.
- Fernald, F.G., B. M. Herman and J. A. Reagan, 1972: “Determination of aerosol height distributions with lidar”, *J. Appl. Meteorol.*, **11**, 482-489.
- Gambling, D. and K. Bartusek, 1972a: “Lidar observations of tropospheric aerosols”, *Atmos. Environ.*, **6**, 181-190.
- Gambling, D. and K. Bartusek 1972b: “Discussion: Lidar observations of tropospheric aerosols” *Atmos Environ.*, **6**, 869-870.
- Gambling, D. G., K. Bartusek and W. G. Elford, 1971: “A 12-month study of aerosols below 60 km”, *J. Atmos. Terr. Phys.*, **33**, 1403-1413.
- Hitschfeld, W., and J. Bordan, 1954: “Errors inherent in the radar measurements of rainfall at attenuating wavelengths”, *J. Meteorol.*, **11**, 58-67.
- Hu, Y., 2007: “Depolarization ratio–effective lidar ratio relation: Theoretical basis for space lidar cloud phase discrimination”, *Geophys. Res. Lett.*, **34**, L11812, doi:10.1029/2007GL029584.
- Hu, Y., D. Winker, P. Yang, B. Baum, L. Poole, L. Vann, 2001: “Identification of cloud phase from PICASSO-CENA lidar depolarization : a multiple scattering sensitivity study”, *J. Quant. Spectro. Rad. Trans.*, **70**, 569-579.
- Klett, J. D., 1981: “Stable analytical inversion for processing lidar returns”, *Appl. Opt.*, **20**, 211-220.
- Klett, J. D., 1985: “Lidar inversion with variable backscattering/extinction ratios”, *Appl. Opt.*, **24**, 6451-6456.
- Liu, Z., M. Vaughan, D. Winker, C. Hostetler, L. Poole, D. Hlavka, W. Hart, and M. McGill, Use of probability distribution functions for discriminating between cloud and aerosol in lidar backscatter data, *J. Geophys. Res.*, **109**, D15202, doi:10.1029/2004JD004732, 2004.

- Macke, A., J. Mueller, and E. Raschke, 1996: “Single scattering properties of atmospheric ice crystals”, *J. Atmos. Sci.* **53**, 2813-2825.
- Miller, S. D. and G. L. Stephens, 1999: “Multiple scattering effects in the lidar pulse stretching problem”, *J. Geophys. Res.*, **104**, D18, 22205-22220.
- Pinnick, R. G., S. G. Jennings, P. Chylek, C. Ham, and W. T. Grandy, 1983: “Backscatter and Extinction in Water Clouds”, *J. Geophys. Res.*, **88**, 6787-6796.
- Platt, C. M. R., 1973: “Lidar and radiometer observations of cirrus clouds”, *J. Atmos. Sci.*, **30**, 1191-1204.
- Platt, C. M. R., 1979: “Remote sounding of high clouds: I. Calculation of visible and infrared optical properties for lidar and radiometer measurements”, *J. App. Meteor.*, **18**, 1130 – 1143.
- Platt, C. M. R., 1981: “Remote sounding of high cirrus clouds III: Monte Carlo calculations of multiple-scattered lidar returns”, *J. Atmos. Sci.*, **38**, 156-167.
- Sassen, K., and B. S. Cho, 1992: “Sub-visual thin cirrus lidar dataset for satellite verification and climatological research”, *J. Appl. Meteor.*, **31**, 1275-1285.
- Sassen, K. and J. M. Comstock, 2001: “A midlatitude cirrus cloud climatology from the facility for atmospheric remote sensing. Part III: Radiative properties”, *J. Atmos. Sci.*, **58**, 2113-2127.
- Stephens, G. L., et al. (2002), “The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation”, *Bull. Am. Meteorol. Soc.*, **83**, 1771–1790.
- Takano, Y. and K. N. Liou, 1989: “Solar radiative transfer in cirrus clouds. Part I: Single-scattering and optical properties of hexagonal ice crystals”, *J. Atmos. Sci.*, **46**, 3-18.
- Vaughan, M. A., D. M. Winker, and C. A. Hostetler, 2002: “SIBYL: A Selective Iterated Boundary Location Algorithm for Finding Cloud and Aerosol Layers in CALIPSO Data”, in: *Lidar Remote Sensing in Atmospheric and Earth Sciences: reviewed and revised papers presented at the twenty-first International Laser Radar Conference (ILRC21)*, Québec, Canada, L. R. Bissonnette, G. Roy and G. Vallée, Eds, Québec, Defence R&D Canada, Valcartier, 791–794.
- Vieze, W., E. E. Uthe and R. T. H. Collis, 1969: “Lidar observations of airfield approach conditions: An exploratory study”, *J. App. Meteor.*, **8**, 274-283.
- Winker, D. M. and L. R. Poole, 1995: “Monte-Carlo calculations of cloud returns for ground-based and space-based lidars”, *Appl. Phys. B*, **60**, 341-344.
- Winker, D. M., R. H. Couch, and M. P. McCormick, 1996: “An overview of LITE: NASA's Lidar In-space Technology Experiment”, *Proc. IEEE*, **84**, 164-180.
- Winker, D. M. 2003: Accounting for Multiple Scattering in Retrievals from Space Lidar, Proc. SPIE vol. 5059, 128-139.
- Winker, D. M. W. H. Hunt, and M. J. McGill, 2007: “Initial performance assessment of CALIOP”, *Geophys. Res. Lett.*, **34**, L19803.
- Young, S. A., 1995: “Lidar analysis of lidar backscatter profiles in optically thin clouds”, *Appl. Opt.*, **34**, 7019-7031.

Young, S. A., M. A. Vaughan, and , D. M. Winker, 2003: “Adaptive Algorithms for the Fully-Automated Retrieval of Cloud and Aerosol Extinction Profiles from CALIPSO Lidar Data”, *In: IGARSS 2003: Proceedings: IEEE 2003 International Geoscience and Remote Sensing Symposium: Learning from Earth's Shapes & Colors*, Centre de Congrès Pierre Baudis, Toulouse, France. Piscataway, NJ: Institute of Electrical and Electronics Engineers, Vol. 3, 1517 – 1519.

Draft Document