Fully automated analysis of space-based lidar data: an overview of the CALIPSO retrieval algorithms and data products

Mark Vaughan [a], Stuart Young [b], David Winker [c], Kathleen Powell [a], Ali Omar [c], Zhaoyan Liu [d], Yongxiang Hu [c], and Chris Hostetler [c]

[a] SAIC, Mail Stop 435, NASA Langley Research Center, Hampton VA, USA 23681
[b] CSIRO Atmospheric Research, Private Mail Bag 1, Aspendale, VIC, 3195, Australia
[c] NASA, Mail Stop 435, NASA Langley Research Center, Hampton VA, USA 23681
[d] Hampton University, Mail Stop 435, NASA Langley Research Center, Hampton VA, USA 23681

ABSTRACT

The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations[1] (CALIPSO) satellite will be launched in April of 2005, and will make continuous measurements of the Earth's atmosphere for the following three years. Retrieving the spatial and optical properties of clouds and aerosols from the CALIPSO lidar backscatter data will be confronted by a number of difficulties that are not faced in the analysis of ground-based data. Among these are the very large distance from the target, the high speed at which the satellite traverses the ground track, and the ensuing low signal-to-noise ratios that result from the mass and power restrictions imposed on space-based platforms. In this work we describe an integrated analysis scheme that employs a nested, multi-grid averaging technique designed to optimize tradeoffs between spatial resolution and signal-to-noise ratio. We present an overview of the three fundamental retrieval algorithms (boundary location, feature classification, and optical properties analysis), and illustrate their interconnections using data product examples that include feature top and base altitudes, feature type (i.e., cloud or aerosol), and layer optical depths.

1. INTRODUCTION

Conceptually, the automated analysis of lidar backscatter data is a relatively straightforward process. The analyst must first locate all layer boundaries, and then must identify each layer as being either cloud or aerosol. Knowing the 'where' and 'what' of each layer, the range-resolved profiles of optical properties can then be derived. In the final step of the process, the analyst must composite all of the measured and derived spatial, morphological, and optical characteristics of each layer into a sensible data format that lends itself to easy use by researchers from a broad cross-section of scientific disciplines. However, as is the case with so many complex tasks, the preceding concise enumeration glosses over a number of critical points. Particularly with respect to computer automation, the lidar data analysis devil does indeed lie deep in the details. In this paper we will expose those critical levels of detail on which the foundations of the CALIPSO analysis architecture are constructed. We will provide the heritage and briefly describe the mechanics of our algorithms for identifying layer boundaries within a single lidar profile (section 2); for discriminating between cloud layers and aerosol layers and determining layer subtype (section 3); and for retrieving profiles of volume backscatter and extinction coefficients from the lidar measurements (section 4). We will also describe the content and structure of the CALIPSO Level II lidar data products, and trace through the one-to-one matching between the retrieval algorithm(s) and each major class of data products (section 6).

The fundamental algorithms in the CALIPSO analysis framework can be considered as *profile processes*. These routines are solidly rooted in traditional lidar data analysis techniques, and take as inputs either a single profile of lidar measurements, or, in the case of the layer classification routines, 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., volume extinction coefficients), the integral of such a profile (e.g., optical depth), or some additional information about the content and/or structure of the input profile (e.g., as provided by base and top altitudes). In every case, the uncertainties associated with the outputs of these profile processes are a strong function of the signal-to-noise ratio (SNR) of the input data. Therefore, perhaps the most critical and innovative steps in the CALIPSO analysis framework occur during the *scene processes* that run immediately prior to the initiation of any of the individual profile

processes. These scene processes are essentially sophisticated, multi-level averaging schemes whose function are to identify and extract high SNR profile data from an extended, contiguous sequence of profile measurements; that is, from a lidar "scene". The averaged profiles thus obtained are then delivered for analysis to the appropriate profile processes. A thorough discussion of the rationale motivating the CALIPSO scene processes, and of the techniques employed, is given in section 5.

The CALIPSO satellite carries a polarization-sensitive lidar that makes backscatter measurements at 532 nm and 1064 nm. A complete instrument description is given in a companion paper published in this volume[2]. In this work we assume that the CALIPSO profile data has been properly geolocated and calibrated during the Level I data processing[3]. Therefore the primary measurement quantity used in the Level II data processing will be range-resolved profiles of attenuated backscatter coefficients, $\beta'_\lambda(r)$, as follows:

$$\beta'_\lambda(r) = \frac{r^2 P_\lambda(r)}{C_\lambda} = \left(\beta_{\lambda,m}(r) + \beta_{\lambda,p}(r)\right) T_{\lambda,m}^2(r) \, T_{\lambda,O_3}^2(r) \, T_{\lambda,p}^2(r). \tag{1}$$

The system constant at each wavelength is given by C_λ, and the background-subtracted raw data by $P_\lambda(r)$. Backscatter coefficients are represented by $\beta(r)$; the two-way transmittance due to any scattering (or absorbing) species is given by T^2, and the subscripts m, O_3, and p specify, respectively, molecular, ozone, and particulate contributions to the signal. Here we understand the term "particulates" to represent either clouds or aerosols, as appropriate. The two-way transmittance for any constituent, x, is

$$T_x^2(r) = \exp\left(-2\,\tau_x(r)\right) = \exp\left(-2 \int_0^r \sigma_x(r')dr'\right). \tag{2}$$

where $\tau_x(r)$ specifies the optical depth and $\sigma_x(r)$ is the volume extinction coefficient.

## 2. TECHNIQUES FOR FINDING CLOUD AND AEROSOL LAYER BOUNDARIES

Space-based lidar measurements can be used to detect the vertical location of several different classes of geophysical objects. Specifically, individual lidar returns can contain information on the base and top altitudes of clouds and aerosol layers, and on the altitude of the Earth's surface. However, before we can determine the specific class of geophysical object that has been detected, we first must determine its spatial extent. That is, before we can say *what* it is, we first must know *where* it is…and, with the possible exception of surface spikes, we cannot use specifics of the "*what*" to help us in our search for the "*where*". It is only after determining the spatial boundaries that we can begin to quantify the optical properties that will allow us to classify an object as either cloud or aerosol. Therefore, within the context of our boundary detection scheme, we employ the generic term "feature" to describe any extended and contiguous region of enhanced backscatter signal that rises significantly above the expected molecular value. Clearly this definition encompasses all of our targets of interest; that is, clouds, aerosol layers, and surface returns. The sole task of the feature finding algorithm is thus to separate the genuine features from the pseudo-features (i.e., noise excursions). The subsequent task of identifying feature type – that is, classifying each feature as either a cloud or an aerosol – is deferred until later in the analysis (i.e., section 3).

The CALIPSO feature-finding algorithm is driven by a robust profile-scanning engine (i.e., a profile process) originally developed for application to ground-based observations[4] and later adapted for space-based analyses using LITE data[5]. For each profile analyzed, a range-dependent threshold is constructed based on estimates of both random and signal-induced noise levels[6]. Because these estimates are derived from the profile being analyzed, the magnitudes of the threshold values scale automatically with increased signal averaging. Scanning begins at the top of the profile and data points are examined sequentially, with features being identified as those regions where the profile data exceed their corresponding threshold values over an altitude range greater than some predetermined minimum feature thickness. The scanning procedure is applied to the 532 nm attenuated scattering ratios, given by

$$R'(r) = \frac{\beta'(r)}{\beta_m(r)\, T_m^2(r)\, T_{O_3}^2(r)} = \left(1 + \frac{\beta_p(r)}{\beta_m(r)}\right) T_p^2(r) \quad . \tag{3}$$

The attenuated scattering ratio profiles are used in order to exploit a very specific geometric structure. In a hypothetical noise-free profile, the slope of R′ with respect to altitude would be exactly zero. Figure 1 shows a profile of R′ measured during orbit 14 of the LITE mission[7]. As in the hypothetical profile, this slope of R′ with respect to altitude is indeed essentially zero in the clear air regions of the measured data from 22-km to 17.5-km and from ~12-km to ~3-km. The adaptive thresholding scheme incorporated into the CALIPSO profile scanner uses local topographical information to identify regions of clear air at feature base, and to adjust the threshold array to compensate for the signal attenuation imparted by the newly located feature. This sequence is illustrated by application to the LITE data shown in Figure 1. Figure 1b shows the range-dependent threshold computed at the beginning of the scan. The cirrus between 17.4-km and 13.2-km lies comfortably above the initial threshold, and hence its detection is assured. However, the aerosol layer between 1.6-km and the surface lies entirely beneath this threshold, and therefore would not be detected using a static threshold array. Figure 1c shows the revised threshold (in red) that is automatically computed immediately upon exiting the base of the cirrus layer. For comparison, the initial threshold is shown in blue. Regions identified as features are drawn in green; "non-feature" regions are drawn in pale gray. This automated, in-line scaling of the threshold array to account for signal attenuation due to detected features makes possible the subsequent detection of underlying features that, as seen in Figure 1b, would otherwise be missed. The scaling factor, computed using the mean attenuated scattering ratio in the clear air region beneath the feature, also provides a useful estimate of the two-way transmittance of all over-lying features (i.e., as in equation (1) for $\beta_p(r) = 0$).

Figure 1: Adaptive threshold technique applied to attenuated scattering ratios measured during LITE.
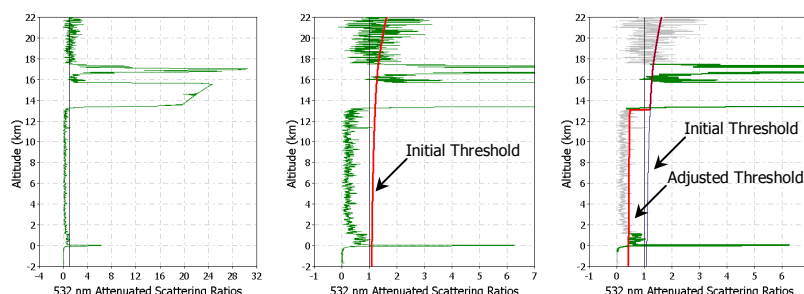
## 3.   SCENE CLASSIFICATION ALGORITHMS AND THE IDENTIFICATION OF LAYER TYPE

After the scanning module identifies all features within a profile, the task of ascertaining feature type falls to the cloud-aerosol discrimination (CAD) algorithm. Once a feature has been successfully classified as being either cloud or aerosol, additional scene classification algorithms (SCA) are applied to the data to establish either a cloud lidar ratio and cloud ice-water phase or an aerosol lidar ratio and aerosol subtype.

3.1.   Discriminating clouds from aerosols

Intrinsic scattering properties are those that depend only on the type of particulate being measured, and are independent of the amount or concentration. *Extrinsic* properties on the other hand depend directly on particulate amount within the scattering volume. To differentiate between clouds and aerosols, the CALIPSO CAD algorithm relies on the statistical differences in both the intrinsic and extrinsic scattering characteristics exhibited by the two types of features[8]. The nature of these differences is illustrated in Figure 2a, where modeled values of the particulate backscatter color ratio, $\chi = \beta_{1064} / \beta_{532}$, are plotted as a function of particulate backscatter coefficients at 532 nm, β_{532}, for several feature types defined by the Optical Properties of Aerosols and Clouds (OPAC) software package[9]. Because clouds generally consist of particles that are quite large with respect to the CALIPSO wavelengths, the backscatter and extinction coefficients are not expected to show any substantial spectral variation. Therefore both the backscatter and extinction color ratios should be approximately one. Conversely, the particle sizes for most aerosols (with the possible exception of some dust layers) are much smaller, and thus should exhibit a spectrally dependent scattering efficiency. The expectation is that aerosol backscatter coefficients will be smaller at 1064 nm than at 532 nm, so that in general $\chi_{aerosol} < 1$. The OPAC

model predictions confirm these expectations with respect to the intrinsic property χ, and further show that a similar situation exists for the extrinsic property β_{532}.

Figure 2: Scattering properties of clouds and aerosols; (a) OPAC model predictions of χ and β_{532} (b) LITE measurements of χ' (c) LITE-derived distributions of β'_{532} for clouds and aerosols found between 0 and 10 km
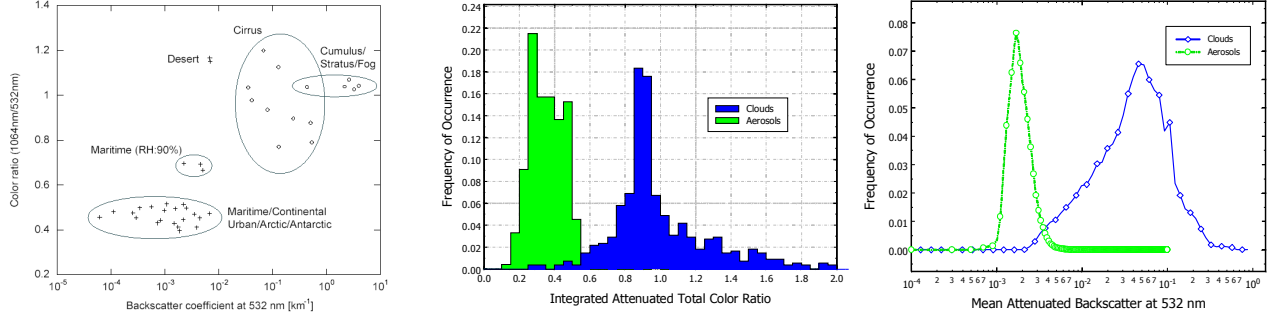
Unfortunately, the OPAC-type separation between clouds and aerosol cannot be immediately discerned in the CALIPSO backscatter data. To extract the particulate backscatter coefficients at any altitude, the data at each wavelength must first be corrected for the attenuation of the signal by molecules and particles at and above this altitude. This in turn requires the selection of an appropriate lidar ratio. We are thus left in a logical quandary: to specify the correct lidar ratio, we must first discriminate between clouds and aerosols; however, to properly discriminate between clouds and aerosols, we first need the feature's lidar ratio. To resolve this dilemma, the discrimination between clouds and aerosols is accomplished by computing the value of a confidence function which takes as its inputs three immediately measurable feature attributes: the mid-feature altitude, $Z_{mid} = \frac{1}{2}\left(Z_{top} + Z_{base}\right)$; the feature mean attenuated backscatter at 532 nm,

$$\overline{\beta}'_{532} = \frac{1}{Z_{top} - Z_{base}} \int_{Z_{top}}^{Z_{base}} \beta'_{532}(r)\,dr\,; \qquad \text{and the feature integrated attenuated total color ratio,}$$

$$\chi' = \int_{Z_{top}}^{Z_{base}} \beta'_{1064}(r)\,dr \Big/ \int_{Z_{top}}^{Z_{base}} \beta'_{532}(r)\,dr\,.$$ The latter two, $\overline{\beta}'_{532}$ and χ', are integrated, attenuated quantities that serve as

proxies for their counterparts, β_{532} and χ, respectively. That these represent reasonable (albeit imperfect) approximations is demonstrated in Figure 3b and Figure 3c, which show distributions compiled from the LITE data.

The confidence function is given by

$$F\left(z, \chi', \overline{\beta}'_{532}\right) = \frac{P_{cloud}\left(z, \chi', \overline{\beta}'_{532}\right) - K(z)\,P_{aerosol}\left(z, \chi', \overline{\beta}'_{532}\right)}{P_{cloud}\left(z, \chi', \overline{\beta}'_{532}\right) + K(z)\,P_{aerosol}\left(z, \chi', \overline{\beta}'_{532}\right)}\,. \tag{4}$$

where $P_{feature}\left(z, \chi', \overline{\beta}'_{532}\right)$ is the probability distribution function (PDF) for a given feature type and K is a scaling factor that quantifies the relative frequency of occurrence of the two feature types. Values of F range between -1 and 1. The sign of the return value determines the feature type: negative values are aerosols and positive values are clouds. The magnitude of F assigns a confidence score to the classification[8]. The PDFs that will be used initially for $\overline{\beta}'_{532}$ and χ' have been derived from measurements made by LITE and the Cloud Physics Lidar[10] (CPL). The algorithm architecture allows individual PDFs to be updated at arbitrary intervals.

3.2. Determining cloud ice-water phase and lidar ratio

The CALIPSO scene classification algorithm suite contains a module dedicated to partitioning those features identified as clouds into three distinct groups: ice clouds, water clouds, and mixed phase clouds. The ice-water phase algorithm uses the layer-integrated volume depolarization ratio, δ_v, the uncertainty in δ_v, $\Delta\delta_v$, and the mean attenuated backscatter coefficient to compute the probability that a cloud is ice or water according to the following functions:

$$P_{ice}\left(\beta', \delta_v, \Delta\delta_v\right) = \left(1 + \exp\left(-\Delta\delta_v\left(\delta_v - Y_1\left(\beta'\right)\right)\Big/ Y_1\left(\beta'\right)\right)\right)^{-1} \tag{5a}$$

and
$$P_{water}\left(\beta', \delta_v, \Delta\delta_v\right) = 1 - \left(1 + \exp\left(-\Delta\delta_v\left(\delta_v - Y_2\left(\beta'\right)\right)\Big/ Y_2\left(\beta'\right)\right)\right)^{-1} \tag{5b}$$

where Y_1 and Y_2 are empirically developed sigmoidal (S-curve) functions defined respectively as $Y_1\left(x\right) = 0.25\left(1 + x^2\right)$ and $Y_2\left(x\right) = 0.05 + 0.2\,x^{1.25} + 0.4\,x^4 + 0.5x^8$. If the depolarization ratio calculation fails to generate a high confidence classification (i.e., either $P_{ice} > 0.5$ or $P_{water} > 0.5$), a second probability test is performed based on the temperature at cloud top. For this test $P_{water}\left(T\right) = 1\big/\left(1 - W\left(T\right)\right)$ and $P_{ice}\left(T\right) = 1 - P_{water}\left(T\right)$ where $W\left(T\right) = \exp\left(-mT + b\right)$ for empirically determined constant coefficients m and b.

After the ice-water phase has been determined, a phase-appropriate cloud lidar ratio is selected. For lofted layers, cloud lidar ratios can frequently be estimated directly from the profile data by using the two-way transmittance and the integrated attenuated particulate backscatter[11,12]. However, in those cases where direct retrieval is impossible or inadvisable, a lidar ratio must still be selected in order to initiate the extinction analysis. In these cases, water clouds are assigned lidar ratios of $S_{water} = 18$ sr based on Mie calculations[13]. The lidar ratio for ice clouds is currently chosen using an empirically derived linear function of mid-cloud temperature developed by Platt[14,15].

3.3. Aerosol subtyping and lidar ratio selection

Perhaps the single most difficult task among the scene classification algorithms is determining the appropriate lidar ratio to be used in the optical analyses of aerosol layers. While the lofted layer technique can be applied to aerosols as well as clouds, the vast majority of the aerosol layers that will be encountered by CALIPSO are likely to be either planetary boundary layers (PBL) or other surface-attached aerosol layers (SAL). Exacerbating the problem is the fact that the range spanned by aerosol lidar ratios, S_a, is much greater than that for clouds. Field measurements of aerosol lidar ratios in the SAL vary from 15 sr to in excess of 120 sr[16] and, absent certain kinds of ancillary instrumentation (e.g., a sun photometer), cannot be made by a simple backscatter lidar. Lidar-only measurements of S_a require a Raman lidar or a high spectral resolution lidar (HSRL). CALIPSO therefore determines a value for S_a using a model-matching scheme: the optical (e.g., $\overline{\beta}'_{532}$, δ_v and χ'), geophysical (e.g., latitude, longitude), and temporal (season) characteristics of the layer are used as decision points to navigate a flow chart that ultimately selects a "most likely" aerosol model for each aerosol layer[17]. There are six aerosol types from which to choose. Four of these (polluted continental, biomass burning, desert dust, and polluted dust) are derived from a comprehensive cluster analysis applied to AERONET data gathered from numerous sites from around the globe[18]. The remaining two – clean continental and marine – were synthesized using measurements of, respectively, long-range continental transport[19] and sea salt acquired using a backscatter nephalometer[20].

4. RETRIEVING PROFILES OF OPTICAL PROPERTIES

4.1. Multiple scattering concerns

In the analysis of space-based lidar data, multiple scattering effects cannot be ignored. Even for very narrow fields of view (130 μrad for CALIPSO), the large distance between the lidar and the Earth generates a relatively large spot size at the target altitudes. As multiple scattering effects can thus be significant within the CALIPSO measurements, a range-dependent correction factor, $\eta(r)$, is applied to the particulate optical depth term in the lidar equation.[21] This correction factor is computed based on the CALIPSO instrument geometry and modeled cloud and aerosol scattering properties. Based on scattering phase functions derived from the aerosol models described in section 3.3 and on a range of representative cirrus cloud phase functions[22,23], a family of $\eta(r)$ curves has been developed for application to features of each specified type. Thus for any given feature, once feature type has been determined by the scene classification routines, the corresponding multiple scattering correction is assigned to that feature for all subsequent analyses. The lone exception is dense water clouds, for which multiple scattering corrections will be derived from the measured profiles of 532 nm depolarization ratios.

4.2. Computing particulate extinction and backscatter profiles

Within the lidar community there exist a number of effective methods for deriving particulate extinction and backscatter coefficients from the calibrated, range-corrected lidar signal. Among these, the most widely used are the Klett method[24], the Fernald method[25], and the so-called linear iterative method first introduced in the late 1960s by Elterman[26] and subsequently used extensively by Platt[14,27]. Klett and Fernald offer closed form, analytic solutions, while the linear iterative technique provides a strictly numerical solution. The Klett and Fernald algorithms were both originally developed in the context of single scattering. In later years, both algorithms were adapted for use in multiple scattering analyses via the introduction of a correction factor applied not to the accumulated optical depth, as in the CALIPSO technique, but instead to the range-resolved extinction coefficients (e.g., as in [12]). Solutions for this formulation are typically derived by replacing the range-dependent value $\eta(r)$ with a layer-averaged constant, $\bar{\eta}$ [27], or by assuming that the multiple scattering correction remains constant through the layer. The fact that $\bar{\eta}$ is generally considered as an unknown quantity, to be derived as part of the analysis, is the prime motivation for replacing $\eta(r)$ with $\bar{\eta}$. However, in the CALIPSO framework $\eta(r)$ is considered to be known: a species-specific multiple scattering profile is assigned to each feature by the cloud and aerosol subtyping routines. When $\eta(r)$ is assumed to be known, a Fernald-type solution can be readily derived if multiple scattering is modeled as being a local property modifying the volume extinction coefficient. Conversely, when multiple scattering is modeled as layer property that modifies the optical depth term, the mathematical approach employed in developing the Fernald solution breaks down, and, to our knowledge, no alternative analytical solution has been proposed. The linear iterative method, however, can be easily adapted to incorporate the CALIPSO multiple scattering parameterization, and we have therefore selected this algorithm for use in deriving profiles of volume backscatter and extinction coefficients from both clouds and aerosols[28,29].

## 5. MULTI-GRID AVERAGING STRATEGIES

The algorithms discussed thus far are profile processes, designed for the analysis of a single lidar profile or profile segment. However, due to the random noise inherent in the signal, the quality of information contained in a single, full resolution backscatter profile can vary enormously, and depends strongly on the scattering intensity of the features being measured. The backscatter intensities within clouds and aerosol layers can range over several orders of magnitude. The detection of the strongest of these features is straightforward, and can be done reliably on a single shot basis. On the other hand, the detection and analysis of the faintest features often requires considerable horizontal and vertical averaging, both to enhance our ability to detect features and to ensure the accurate retrieval of feature optical properties. There are, however, several caveats that apply to the signal averaging process. While averaging will increase our retrieval accuracies when applied to a homogeneous region, in a heterogeneous scene it can smear or eliminate important structural variability within features, and can irretrievably intermingle the signals from both aerosols and clouds into a single profile, thus degrading or even destroying the information content of the measurements. Therefore, to make efficient use of our profile processes, we also require a set of fully automated "scene processes" designed to intelligently extract suitably averaged data from an arbitrary collection of contiguous lidar profiles. The sections below describe the scene processes developed for layer detection and the subsequent optical analyses.

5.1. SIBYL: A Selective, Iterated Boundary Locator

The scene process used for finding features is a multi-level averaging scheme known as the selective iterated boundary locator (SIBYL). The necessity for this kind of variable averaging scheme can be readily appreciated by examining Figure 3. This figure shows an image of a representative segment of lidar backscatter data acquired during the LITE mission (3a), together with single-shot profile data from three selected locations (3b, 3d, and 3e) and a horizontally averaged profile (3c) from one of these three locations. In 3a, note in particular the aerosol mass at approximately 2.5 km on the far left edge, and the faint cirrus at approximately 14 km in the left center. Both features are plainly visible in the image plot, yet due to random noise in the backscatter signal neither shows prominently in the corresponding profile plots (3b and 3e). This lack of visibility is in stark contrast to the profile of cirrus overlying dense stratus shown in 3d; these features are plainly visible in both the image and the profile views. The feature boundaries in Figure 3d are easily retrieved from the single-shot data. However, to locate boundaries in the regions indicated by Figures 3b and 3e requires additional horizontal averaging, and perhaps some vertical smoothing as well. The effects of averaging over even a relatively short distance can be quite dramatic. This is illustrated by Figure 3c, which shows data from the same

region as Figure 3b, except that the profile has been averaged over a horizontal distance of 20 km. After averaging, the backscatter intensities of both the cirrus at 14 km and the aerosol layer at 1.5 km stand out noticeably against the ambient molecular scattering.

Figure 3: (a) Heterogeneous scattering scene measured by LITE; (b) single-shot profile showing thin cirrus and aerosol layers; (c) as in b, but averaged to a 20-km horizontal resolution; (d) single-shot profile showing strong cirrus overlying stratus; (e) single-shot profile showing faint cirrus overlying PBL aerosols.
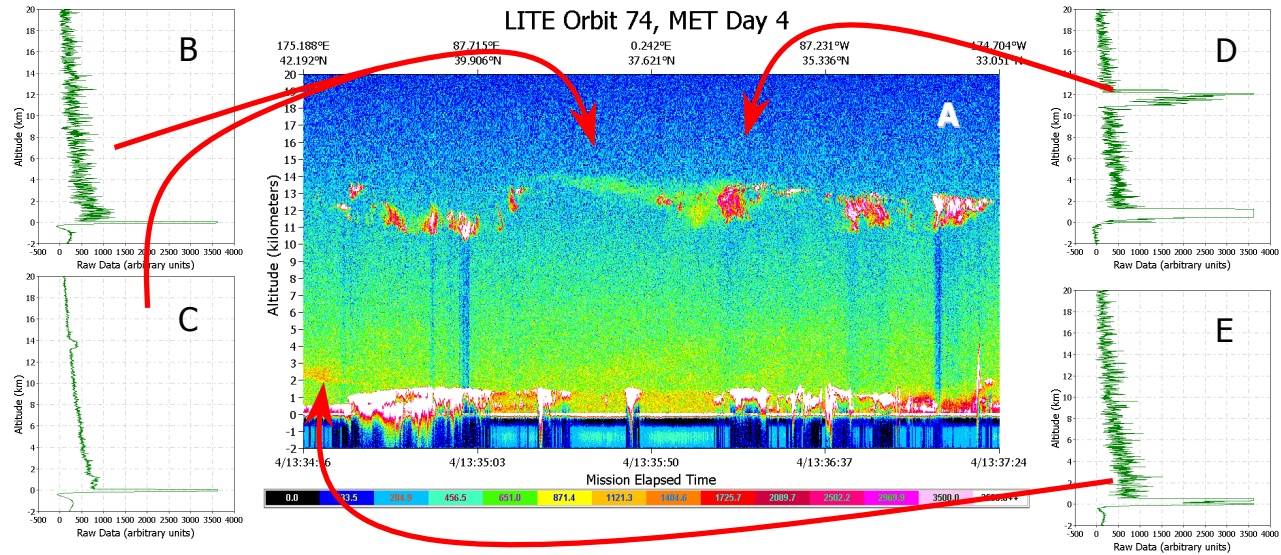
As can be seen in the lower altitudes in the center portion of Figure 3a (and to a lesser extent in the cirrus layer), regions of extremely strong scattering can be embedded in otherwise very faint features. Moreover, because these regions of strong scattering and weak scattering are intermingled, seemingly at random, results returned by SIBYL will necessarily be inhomogeneous with respect to spatial resolution. That is, intense features can be found and hence reported on a much finer spatial scale than weak, diffuse features.

The ability to find features at successively greater horizontal averaging levels, corresponding to successively coarser spatial resolutions depends heavily on a technique known as feature clearing. For each feature found within a profile, the feature clearing technique (1) estimates the feature two-way transmittance; (2) *removes* the feature backscatter data from the profile and replaces it with data derived from a molecular model; and (3) corrects all backscatter data beneath the removed feature for the attenuation ascribed to that feature. Figure 4 presents a graphical depiction of this process. The goal of the feature clearing procedure is to generate a modified profile that represents the backscatter that would have been measured had the feature(s) not been present[6]. As illustrated in Figure 4d, a contiguous sequence of feature cleared profiles can then be further averaged to expose here-to-fore obscured features. The attenuation corrections that are applied beneath each extracted feature ensure that the profiles used in subsequent averages are properly weighted throughout with respect to the varying optical depths of the features removed in previous steps. Without the application of the feature clearing step, successive averages would likely fail to enhance the contrast of (apparently) weaker features with respect to the ambient background, and the profile scanner would thus fail.

The SIBYL algorithm is designed to be fully configurable via runtime script files. The current configuration repeats the "scan, clear, reaverage, scan again" sequence for three horizontal averaging intervals: 5-km, 20-km, and 80-km. The 5-km interval is dictated by on-board data averaging considerations. (See [2]; CALIPSO achieves a degree of data compression by applying a range-dependent onboard averaging scheme over each 5-km horizontal segment.) The 20-km and 80-km increments were selected to achieve a factor-of-two increase in the profile SNR at each successive averaging interval.

In addition to the feature clearing technique, the SIBYL algorithm also employs a 'full resolution search mode' (FRSM) specifically designed to separate clouds from aerosols within the SAL and/or PBL. Immediately following the initial scan of the 5-km averaged data, the FRSM procedure builds a modified threshold array for use in applying the scanning engine to each of the single shot profiles that were used to construct the 5-km average. Because the features identified

during the FRSM scan will be confined to regions of exceptionally strong backscatter, we can define them a priori as being clouds (e.g., refer to Figure 2). The cloud regions thus identified are excluded from all subsequent averaging operations. To avoid false positives and conserve computational resources, the search of the single shot profiles is confined to those regions where features were previously identified in the initial scan of the 5-km averaged data.

Figure 4: SIBYL's feature clearing technique panel (a) shows a 5-km averaged profile with a moderate cirrus layer overlying a weaker absorbing aerosol layer; panels (b) and (c) show, respectively, the removal of the cirrus layer and the correction of the underlying data for the estimated cirrus attenuation. Averaging four consecutive feature cleared 5-km profiles results in the 20-km horizontal average depicted in panel (d). This 20-km average reveals the previously obscured aerosol layer. The dotted lines in panels (a) and (d) represent the thresholds used by the profile-scanning engine.
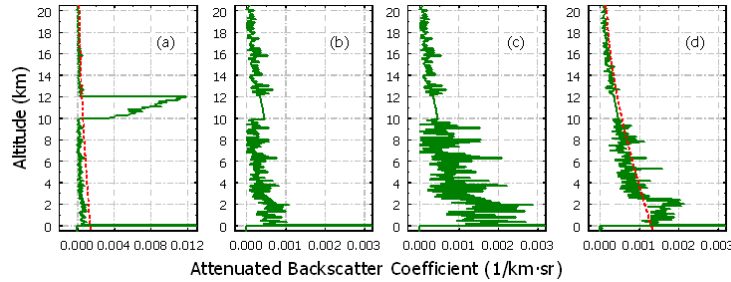
5.2.   HERA: The <u>H</u>ybrid <u>E</u>xtinction <u>R</u>etrieval <u>A</u>lgorithm

The averaging scheme incorporated into the SIBYL algorithm partitions each 80-km segment of data according to the backscatter intensity measured at 532 nm. At the conclusion of the SIBYL process, those regions containing features will have been positively identified. However, the type of feature represented by each region has not yet been determined. Furthermore, although the feature clearing procedure makes initial estimates of layer optical depths, which are then used to correct the profile data for the signal attenuation imparted by removed features, the SIBYL algorithm does not attempt to retrieve layer optical properties. SIBYL's contributions are limited to performing a spatial analysis of each scene. The subsequent tasks of determining feature type, computing final estimates of optical depth, and extracting profiles of particulate backscatter and extinction coefficients are all allocated to the hybrid extinction retrieval algorithm (HERA).

Strictly speaking, HERA is not one distinct, separately identifiable algorithm. It is instead 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 averages varying numbers of profiles on the SIBYL spatial grid to produce a single attenuated backscatter profile that is subsequently used by the layer classification algorithms and the extinction retrieval algorithms. Whereas SIBYL estimates the total transmittance loss through a feature and, when appropriate, corrects underlying signals for this attenuation, HERA calculates and corrects for attenuation within all features in order to produce profiles of extinction and backscatter. HERA is also required to correct 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, or in all 1064-nm features).

Consider the images shown in Figure 5. The left-most panel shows a segment of LITE data that has been resampled using the CALIPSO lidar simulation software[30] to generate a sequence of synthetic backscatter profiles. The center panel shows the results returned by the SIBYL algorithm. Red boxes indicate features found at the 5-km resolution. Similarly, yellow and green boxes represent 20-km features and 80-km features, respectively. A surface detection algorithm is applied to the base of the lowest feature detected in each 5-km column. Those portions of the feature attributed to the surface echo are shown in gray. In this particular example, there were no features identified in the full resolution scans. The right-most panel of Figure 5 makes the separation by region more obvious. Regions of clear air are shown in cyan and subsurface regions in pale gray.

Figure 5c represents the scene description provided to HERA. The scene is partitioned into regions, and the regions are initially classified as being features, clear air, surface returns, or subsurface data. To initiate the optical processing, HERA first retrieves copies of the attenuated backscatter coefficient profiles for the scene that have been averaged to a 5-km horizontal resolution. Each atmospheric region is then processed in turn, starting with the uppermost region and

working from top to bottom and left to right. Processing terminates either at the base of the lowest feature, where the signal becomes totally attenuated, or upon reaching the surface. The following steps are performed for each region analyzed:

Figure 5: (a) synthetic CALIPSO data – attenuated backscatter coefficients derived from LITE measurements; (b) feature boundaries as found by SIBYL (red boxes = 5-km resolution; yellow = 20-km; green = 80-km; gray = surface); (c) HERA analysis regions (cyan = clear air; dark gray = surface; pale gray = subsurface; gray regions are not analyzed)
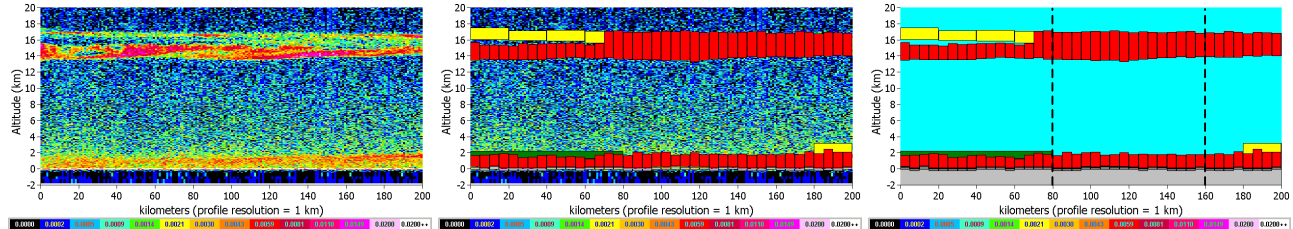
a. Using the map of region boundaries provided by SIBYL, all data within the region is averaged horizontally to form a single profile. Note that different regions can have different horizontal extents, depending on the amount of averaging applied by SIBYL its search for features.

b. For features, the layer classification algorithms are invoked to determine the feature type and to obtain initial estimates of the lidar ratios and type-specific profiles of multiple scattering corrections for both the 532 nm and 1064 nm channels. Within the clear air regions, default values of lidar ratio representative of background aerosols are used and $\eta(r)$ is set uniformly to 1.

c. Using the supplied values of S and $\eta(r)$, extinction and backscatter coefficients are retrieved at both wavelengths using the linear iterative method. On those occasions when SIBYL can provide a reliable estimate, HERA will use the feature optical depth to constrain the extinction retrieval by adjusting the lidar ratio. When a constraint is not available, HERA will still adjust the lidar ratio as necessary to ensure that the solutions produced are both physically meaningful and consistent with the measured attenuated backscatter profile.

d. Once an acceptable set of solutions has been obtained, all 532 nm and 1064 nm attenuated backscatter data lying beneath the current region are corrected for attenuation effects by dividing through by the value of particulate two-way transmittance, T_p^2, that was used or calculated in the retrieval of the region's extinction profile.

Perhaps the most significant aspect of the HERA architecture is the fact that the layer classification algorithms are always operating on the most up-to-date estimates of the layer optical properties. As noted in section 3.1, classification of feature type depends strongly on both backscatter intensity and the integrated attenuated color ratio. By correcting the backscatter data beneath each solved region, $\overline{\beta}'_{532}$ and χ' can be recomputed immediately prior to launching the extinction retrieval. This interleaving of the layer classification algorithms with the extinction solver generates better estimates of the lidar ratios and multiple scattering factors needed in the analyses, and thus improves the quality and performance of all facets of the HERA retrieval.

A more in-depth description of the HERA algorithm is presented in recent publications by Young et al. [28,29,31]

6. FROM ALGORITHMS TO DATA PRODUCTS

Even the most cursory examination of LITE or CPL data shows that aerosols and clouds commingle both spatially and temporally, seemingly at will. This commingling is reflected in the underlying mechanics of the CALIPSO analysis routines. SIBYL locates features without regard to their status as clouds or aerosols. Furthermore, to compute the extinction profile for any feature, HERA must account for the attenuation effects from all overlying features, irrespective of type. (Feature classification is immaterial to the operation of the linear iterative algorithm.) It is only after the completion of both the SIBYL and HERA algorithms that the cloud and aerosol results can be separately extracted. The mapping of algorithms to data products is diagrammed in Figure 6. During the operation of SIBYL and HERA, results are accumulated as internal data storage in a set of intermediate data files. At the conclusion of the processing, the final data products are extracted from these intermediate files.

The CALIPSO data products are generated in three standard formats: a vertical feature mask (VFM), which provides concise information on the spatial and morphological distribution of features; a suite of cloud and aerosol layer products that provide statistical descriptions of all features detected, and a set of profile products that map the vertical distributions of backscatter and extinction coefficients separately for both clouds and aerosols. An overview of each product class is given in the sections below.

Figure 6: Algorithmic origination for each data product type
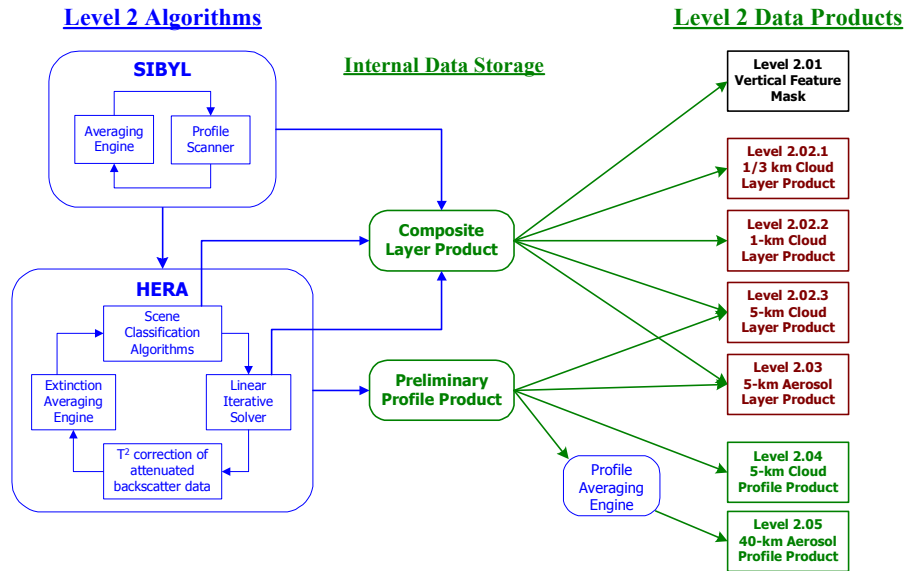
6.1. Vertical feature mask

The vertical feature mask data product is designed to provide researchers simple "where" and "what" information about any region within an along-track swath of atmosphere. The VFM data consists of an array of bit-mapped integers provided at the full downlinked resolution of the CALIPSO lidar, with one integer representing each range resolution element. Each integer value contains a set of scene classification flags that characterize the corresponding spatial region in terms of feature presence and feature type. The interpretation of the individual bits and groups of bits is provided in Table 1. A pictorial example of the VFM is shown in Figure 7. The left-hand panel of the figure shows simulated CALIPSO data derived from measurements acquired by the Cloud Physics Lidar during the CRYSTAL-FACE campaign[32]. The right-hand panel shows a color-coded plot of region type as determined by the feature finder and layer identification algorithms and reported in the VFM data product.

Figure 7: (a) CALIPSO simulations based on CPL data acquired during the CRYSTAL-FACE campaign; (b) vertical feature mask of the data. The textural change in (a) at an altitude of 20-km is due to differences in the onboard averaging resolution of the downlinked data.
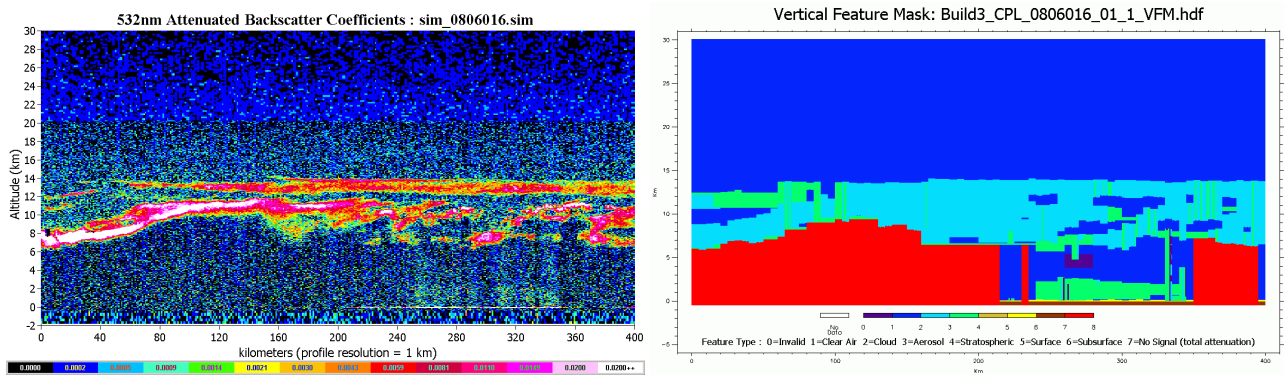
Table 1: Bit interpretation for the CALIPSO Vertical Feature Mask

Bit(s)	Field Description	Bit Interpretation	
1-3	Region type (there are $2^3 = 8$ possible classifications)	0 = invalid (bad or missing data) 1 = "clear air" 2 = cloud 3 = aerosol	4 = stratospheric feature 5 = surface 6 = subsurface 7 = no signal (totally attenuated)
4-5	Region type QA (there are $2^2 = 4$ choices)	0 = none 1 = low	2 = medium 3 = high
6-7	Ice/water phase (clouds only)	0 = unknown/not determined 1 = ice	2 = water 3 = mixed phase
8-9	Ice/water phase QA	0 = none 1 = low	2 = medium 3 = high
10-12	Cloud type (if region type = 2)	0 = not determined 1 = thin low overcast 2 = thick low overcast 3 = broken, low 4 = altocumulus	5 = transparent cirrus 6 = altostratus, nimbostratus, or cumulonimbus 7 = other
	Aerosol type (if region type = 3)	0 = not determined 1 = clean marine 2 = dust 3 = polluted continental	4 = clean continental 5 = polluted marine 6 = smoke 7 = other
	Stratospheric feature type (if region type = 4)	0 = not determined 1 = non-depolarizing PSC 2 = depolarizing PSC	3 = non-depolarizing aerosol 4 = depolarizing aerosol
13-14	Feature type QA	0 = none 1 = low	2 = medium 3 = high
15-17	Horizontal averaging required for detection	0 = not applicable 1 = 1/3-km 2 = 1-km	3 = 5-km 4 = 20-km 5 = 80-km
18-32	spare		

6.2.   Layer products

The VFM is essentially a boolean map of the atmosphere from which one can determine the spatial distribution of clouds and aerosols in the Earth's atmosphere. However, because the VFM is compiled using data derived only from SIBYL and the SCA, it does not provide information on the optical properties of the detected features. This next level of detail is provided by the CALIPSO layer products, which provide a tabulated listing of all features detected within a specified horizontal column. In addition to the spatial and morphological information provided by the VFM, the layer products also contain layer-integrated optical properties for each layer provided by HERA. The fundamental element of the CALIPSO layer products consists of a single set of *column descriptors* associated with a variable number of *layer descriptors*. The column descriptors specify the temporal and geophysical location of the column of the atmosphere through which a given lidar profile penetrated. Also included in the column descriptors are indicators of the surface lighting conditions at the laser footprint (i.e., day or night), information about the surface elevation and type (e.g, land vs. water) derived from external databases, and the number of features identified within each the specific column. Each feature within a column is further characterized by a set of layer descriptors. These layer descriptors describe what might be called the bulk properties of the layer; that is, they provide information about spatial and optical properties that can be ascribed to or describe an entire feature. Typical layer descriptors include such things as base and top altitudes, integrated attenuated backscatter, lidar ratio, and optical depth.

The CALIPSO layer products are separated by feature type. Both cloud and aerosol layer products are provided at a 5-km horizontal resolution. Note that in reporting the layer products, those features found at horizontal averaging resolutions greater than 5-km will be replicated as necessary to span the entire averaging interval required for their

detection. Additional cloud layer products are provided at horizontal resolutions of 1/3-km and 1-km. For these higher resolution results, only those clouds that were strong enough to be detected at the corresponding resolution are reported. Listings of key column descriptors and layer descriptors are provided in Table 2 and Table 3.

Table 2: CALIPSO column descriptors

| Column parameter	Description
Profile ID	Unique integer identifier associated with each laser pulse
Latitude & longitude	Latitude & longitude of the laser footprint
Profile time	Laser firing time and date
Day/night flag	Lighting conditions at the surface
Solar zenith and azimuth angle	
Sensor nadir angle	
532 column reflectance	Reflectances derived from 532 nm parallel and perpendicular channel background monitors and associated uncertainty
Tropopause height & temperature	Obtained from NASA's Global Modeling and Assimilation Office[33]
IGBP surface type	Surface type derived from the International Geosphere-Biosphere Programme (IGBP) database[34]
DEM surface elevation	Height of surface at the footprint location derived from a digital elevation map
Lidar surface elevation	Height of surface determined from lidar return (if found)
Number of layers found	From 0 to a maximum of 15

Table 3: CALIPSO layer descriptors

Layer parameter	Description
Base & top altitudes	Layer base and top heights, in kilometers
Opacity flag	Is feature transmissive or opaque?
Horizontal averaging	SIBYL horizontal averaging level at which the layer was detected
Total attenuated backscatter statistics at 532 nm	Integrated attenuated backscatter and associated uncertainty; also, the minimum, maximum, mean, standard deviation, centroid, and skewness of the 532 nm attenuated total backscatter computed between feature base and feature top
Total attenuated backscatter statistics at 1064 nm	Integrated attenuated backscatter and associated uncertainty; also, the minimum, maximum, mean, standard deviation, centroid, and skewness of the 1064 nm attenuated total backscatter computed between feature base and feature top
Volume depolarization ratio statistics	Layer-integrated volume depolarization ratio $\left(\Sigma\beta'_\perp / \Sigma\beta'_\parallel \right)$ and associated uncertainty; also, the minimum, maximum, mean, standard deviation, centroid, and skewness of the volume depolarization ratios computed between feature base and feature top
Attenuated total color ratio statistics	Layer-integrated attenuated total color ratio $\left(\Sigma\beta'_{1064} / \Sigma\beta'_{532} \right)$ and associated uncertainty; also, the minimum, maximum, mean, standard deviation, centroid, and skewness of the attenuated total color ratios between feature base and feature top
Feature classification flags	Includes cloud/aerosol classification and ice/water phase discrimination; see Table 1 for a complete description of all fields
Measured two-way transmittance at 532 nm (T^2, ΔT^2)	For transmissive features, the mean attenuated scattering ratio beneath the feature and its associated uncertainty
Two-way transmittance measure-ment region	For transmissive features, the base and top altitudes of the region beneath the feature where the two-way transmittance was measured
Feature optical depth at 532 nm	Calculated 532 nm optical depth and associated uncertainty
532 nm lidar ratio	Final 532 nm lidar ratio as reported by HERA
Feature optical depth at 1064 nm	Calculated 1064 nm optical depth and associated uncertainty
1064 nm lidar ratio	Final 1064 nm lidar ratio as reported by HERA

6.3. Profile products

The most detailed analysis of feature optical properties is provided by the range-resolved cloud and aerosol profile products. The cloud products are reported on a regularly spaced grid at a resolution of 5-km horizontally and 60-m vertically. Included in the standard output are profiles of total, parallel, and perpendicular backscatter at 532 nm, extinction coefficients and particulate depolarization ratios at 532 nm, and particulate backscatter color ratios. The ancillary scalar-valued information reported with each set of profiles includes time and position data (latitude, longitude), surface elevation statistics, and a set of quality assurance flags. Meteorological context is provided by accompanying profiles of temperature and atmospheric pressure and number density, obtained from the Global Modeling and Assimilation Office (GMAO)[33] and interpolated to the CALIPSO orbit track. The aerosol profile products are similar in format. However, because the spatial scale of aerosol variability is considerably greater than that for clouds[35], and because the weaker scattering from aerosols requires more averaging in order to obtain accurate results, the aerosol profile products are reported on a coarser spatial grid of 40-km horizontally by 120-meters vertically. Also, because aerosol scattering is expected to show much greater spectral dependence than the scattering from clouds, the aerosol profile products include volume backscatter and extinction coefficients derived from the 1064 nm channel.

As indicated in Figure 6, compiling the aerosol profile products requires an additional averaging step. Recall that within HERA the backscatter and extinction profiles for aerosol layers are generated at 5-km, 20-km, and 80-km horizontal resolutions, and are intermixed according to the spatial resolution required for layer boundary identification by SIBYL. As a result, as illustrated by the center panel of Figure 8, the initial outputs from HERA are spatially inhomogeneous. The post-HERA averaging engine resamples this mixed resolution output, averaging and/or replicating as necessary to generate the uniform grid of optical properties shown in the right-most panel of Figure 8.

Figure 8: (a) simulated aerosol backscatter data derived from CPL measurements; (b) spatially inhomogeneous profile products that are the intermediate output of the SIBYL+HERA processing; (c) level 2 aerosol profile products regridded to a uniform vertical and horizontal resolution
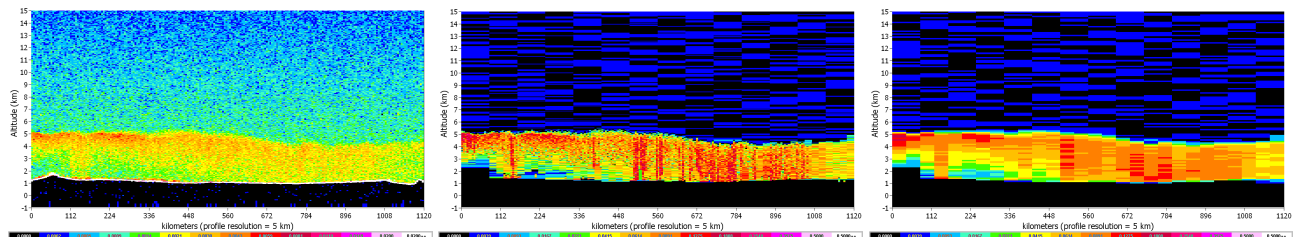
REFERENCES

1   D. M. Winker, J. R. Pelon, and M. P. McCormick, "The CALIPSO mission: spaceborne lidar for observation of aerosols and clouds", *Proc. of SPIE*, **4893**, pp.1-11, 2003.
2   D. M. Winker and C. A. Hostetler, "Status and performance of the CALIOP lidar", current proceedings.
3   J. A. Reagan, X. Wang, and M. J. Osborn, "Spaceborne lidar calibration from cirrus and molecular backscatter returns", *IEEE Trans. Geosci. Remote Sens.*, **40**, pp.2285-2290, 2002.
4   D. M. Winker and M. A. Vaughan, "Vertical distribution of clouds over Hampton, Virginia observed by lidar under the ECLIPS and FIRE ETO programs", *Atmospheric Research*, **34**, pp. 117-133, 1994.
5   C. M. R. Platt, D. M. Winker, M. A. Vaughan, and S. D. Miller, "Backscatter-to-extinction ratios in the top layer of tropical mesoscale convective systems and in isolated cirrus from LITE observations", *Journal of Applied Meteorology*, **38**, pp. 1330-1345, 1999.
6   M. A. Vaughan, D. M. Winker, and C. A. Hostetler, "SIBYL: a Selective Iterated Boundary Location Algorithm for Finding Cloud and Aerosol Layers in CALIPSO Lidar Data", in *Lidar Remote Sensing in Atmospheric and Earth Sciences*, L. R. Bissonnette, G. Roy and G. Vallée (editors), Defence R&D Canada - Valcartier, Québec, Canada, pp. 791-794, 2002.
7   D. M. Winker, R. Couch, and M. P. McCormick, An overview of LITE: NASA's Lidar In-space Technology Experiment, *Proc. IEEE*, **84**, pp.164-180, 1996.

8   Z. Liu, M. A. Vaughan, D. M. Winker, C. A. Hostetler, L. R. Poole, D. L. Hlavka, W. D. Hart and M. J. McGill, "Use of Probability Distribution Functions for Discriminating Between Cloud and Aerosol in Lidar Backscatter Data", *Journal of Geophysical Research*, **109**, doi:10.1029/2004JD004732, 2004.

9   M. Hess , P. Koepke, and I. Schult, "Optical Properties of Aerosols and clouds: The software package OPAC", *Bull. Am. Met. Soc.*, **79**, pp. 831-844, 1998.

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

11 J. M. Alvarez and M. A. Vaughan, "Numerical calculation of cloud optical extinction from lidar", OSA Proceedings on the Inaugural Forum for the Research Center for Optical Physics, 1993, **19**, pp. 90-95, Arlene Maclin, Ed.; Optical Society of America, Washington, D.C., 1994.

12 S. A. Young, "Lidar analysis of lidar backscatter profiles in optically thin clouds" *Applied Optics,* **34,** 7019-7031, 1995.

13 R. G. Pinnick, S. G. Jennings, P. Chylek, C. Ham, and W. T. Grandy, Jr., "Backscatter and Extinction in Water Clouds", *Journal of Geophysical Research*, **88**, pp. 6787-6796, 1983.

14 Platt, C. M. R., S. A. Young, P. J. Manson, G. R. Patterson, S. C. Marsden, R. T. Austin, and J. H. Churnside, "The Optical Properties of Equatorial Cirrus from Observations in the ARM Pilot Radiation Observation Experiment", *Journal of the Atmospheric Sciences*, **55**, pp. 1977-1996, 1998.

15 C. M. R. Platt, S. A. Young, R. T. Austin, G. R. Patterson, D. L. Mitchell, and S. D. Miller, "LIRAD Observations of Tropical Cirrus Clouds in MCTEX. Part I: Optical Properties and Detection of Small Particles in Cold Cirrus", *Journal of the Atmospheric Sciences*, **59**, pp. 3145-3162, 2002.

16 S. J. Masonis, *An empirical study of the lidar ratio and its variability, with implications for determining climate forcing by satellite-borne lidar*, Ph.D. dissertation, University of Washington, 257 pp., 2001. [Available from Bell and Howell Information and Learning, 300 North Zeeb Road., P.O. Box 1346, Ann Arbor, MI 48106-1346]

17 A. H. Omar, D. M. Winker, J-G. Won, M. A. Vaughan, C. A. Hostetler, and J. A. Reagan, "Selection Algorithm for the CALIPSO Lidar Aerosol Extinction-to-Backscatter Ratio", IEEE Geoscience and Remote Sensing Symposium 2003: Proceedings, IEEE International, **3**, pp. 1526-1530, 2003.

18 A. H. Omar, D. M. Winker, and J-G. Won, "Aerosol Models for the CALIPSO Lidar Inversion Algorithms", in *Laser Radar Technology for Remote Sensing, Proc. SPIE*, **5240**, Christian Werner, Editor, Deutsches Zentrum fuer Luft- und Raumfahrt e.V. (Germany), pp. 104-115, 2004.

19 T. L. Anderson, S. J. Masonis, D. S. Covert, and R. J. Charlson, "In situ measurements of the aerosol extinction-to-backscatter ratio at a polluted continental site", *Journal of Geophysical Research*, **105**, pp. 26,907-26,915 2000.

20 S. J. Masonis, T. L. Anderson, D. S. Covert, V. Kapustin, A. D. Clarke, S. Howell, and K. Moore, "A study of the extinction-to-backscatter ratio of marine aerosol during the Shoreline Environmental Aerosol Study (SEAS)", *Journal of Atmospheric and Oceanic Technology*, **20**, pp. 1338-1402, 2002.

21 D. M. Winker, "Accounting for Multiple Scattering in Retrievals from Space Lidar", in *Proc. SPIE: 12th International Workshop on Lidar Multiple Scattering Experiments,* Christian Werner, Ulrich G. Oppel, and Tom Rother (editors), **5059**, pp. 128-139, 2003.

22 Y. Takano and K. N. Liou, "Solar radiative transfer in cirrus clouds. Part I: Single-scattering and optical properties of hexagonal ice crystals," *Journal of the Atmospheric Sciences*, **46**, pp. 3-18, 1989.

23 G. M. McFarquhar, A. J. Heymsfield, A. Macke, J. Iaquinta, S. M. Aulenbach, "Use of observed ice crystal sizes and shapes to calculate mean-scattering properties and multispectral radiances: CEPEX April 4, 1993, case study," *Journal of Geophysical Research*, **104**, pp. 31 763-31 779, 1999.

24 J. D. Klett, "Lidar inversion with variable backscatter/extinction ratios", *Applied Optics*, **24**, pp. 1638-1643, 1985.

25 F. G. Fernald, B. J. Herman, and J. A. Reagan, "Determination of Aerosol Height Distributions by Lidar" , *Journal of Applied Meteorology*, **11**, pp. 482-489, 1972.

26 L. Elterman, "Aerosol measurements in the troposphere and stratosphere", *Applied Optics*, **5**, pp. 1769-1776, 1966.

27 C. M. R. Platt, "Lidar and Radiometric Observations of Cirrus Clouds", *Journal of the Atmospheric Sciences*, **30**, pp. 1191-1204, 1973.

28 S. A. Young, M. A. Vaughan, and D. M. Winker, Adaptive Methods for Retrieving Extinction Profiles from Space Applied to CALIPSO Lidar Data, in *Lidar Remote Sensing in Atmospheric and Earth Sciences: Reviewed and*

Revised Papers Presented at the 21st International Laser Radar Conference(ILRC 2002), L. R. Bissonnette, G. Roy and G. Vallée (editors), Defence R&D Canada - Valcartier, Québec, Canada, pp. 743-746, 2002.

[29] S. A. Young, M. A. Vaughan, and D. M. Winker, "Adaptive Algorithms for the Fully-Automated Retrieval of Cloud and Aerosol Extinction Profiles from CALIPSO Lidar Data", IEEE Geoscience and Remote Sensing Symposium 2003: Proceedings, IEEE International, **3**, pp. 1517-1519, 2003.

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

[31] S. A. Young, D. M. Winker, M. A. Vaughan, and K. A. Powell, "The Retrieval Of Extinction Profiles In Optically-Inhomogeneous Cloud And Aerosol Layers Detected By The CALIPSO Lidar", in *Reviewed and Revised Papers Presented at the 22nd International Laser Radar Conference(ILRC 2004)*, G. Pappalardo and A. Amodeo (editors), ESA SP-561, pp. 999-1002, 2004.

[32] E. Jensen, D. Starr, and O. B. Toon, "Mission investigates tropical cirrus clouds", *Eos*, **84**, pp. 45, 2004.

[33] File Specification for GEOS-DAS Gridded Output, Version 5.2; see http://gmao.gsfc.nasa.gov/operations/DAO-1001v5.2.pdf

[34] Available for download at http://edcftp.cr.usgs.gov/pub/data/glcc/ap/lambert/apigbp1_2l.img.gz

[35] T. L. Anderson, R. J. Charlson, D. M. Winker, J. A. Ogren, and Kim Holmén, "Mesoscale Variations of Tropospheric Aerosols", *Journal of the Atmospheric Sciences*, **60**, pp. 119-136, 2003.