An Optimal-Estimation based hybrid Rayleigh-Mie Extinction and Backscatter retrieval Method for HSRL lidars

D.P. Donovan[1] **and Gerd-Jan van Zadelhoff**[2]

[1]*KNMI, PO Box 201, 3730 AE, the Netherlands, donovan@knmi.nl*
[2]*KNMI, PO Box 201, 3730 AE, the Netherlands, zadelhof@knmi.nl*

ABSTRACT

The Rayleigh signal from a High-Spectral Resolution Lidar (HSRL) can be used to estimate the extinction profile in a rather direct manner by estimating the derivative of the range-corrected logarithmic signal. However, the applicability of this method is limited due to the high SNR required. Thus, it can be said that accurate but low-precision extinction information is, in general, provided by the Rayleigh signal. In contrast, extinction information can also be extracted from the Mie signal channel which, in general, may be viewed as less accurate (since factors such as the extinction-to-backscatter ratio must be assumed) but more precise (since the sensitivity to the SNR ratio of the input data is much lower). Noting these two observations, it is useful to investigate retrieval procedures which simultaneously uses both the Rayleigh and Mie signals in order to optimally combine the Rayleigh channel derived information with the less accurate but more precise information derived using the Mie channel. To this end, an optimal-estimation based variational approach has been developed for the retrieval of lidar extinction and backscatter from HSRL lidar data. In this paper, the retrieval algorithm is introduced and applied to synthetic data from the EarthCARE simulator.

1. INTRODUCTION

Unlike an elastic backscatter lidar, the high spectral resolution capability of the EarthCARE lidar (ATLID) will provide the unprecedented opportunity to directly measure aerosol and cloud extinction profiles without invoking critical assumptions (such as accurate a-priori knowledge of the backscatter-to-extinction ratio (S) [1]). Considering the cross-talk corrected and background subtracted ATLID signals, the lidar equations for the Rayleigh and Mie signals can be written as

$$
\begin{aligned}
p_R(z) &= \frac{C_R}{(z_{lid}-z)^2}\beta_R(z) Ms_R^{-1}(z) \\
&\times \; e^{-2\int_0^z \left(\alpha_M(z')+\alpha_R(z')\right)dz'}
\end{aligned} \tag{1}
$$

and

$$
\begin{aligned}
p_M(z) &= \frac{C_M}{(z_{lid}-z)^2}\beta_M(z) Ms_M^{-1}(z) \\
&\times \; e^{-2\int_0^z \left(\alpha_M(z')+\alpha_R(z')\right)dz'}
\end{aligned} \tag{2}
$$

where C_M is the lidar Mie signal calibration constant, C_R is the lidar Rayleigh signal calibration constant, β_M is the Rayleigh backscatter coefficient, β_R is the Rayleigh backscatter coefficient, α_M and α_R are the respective extinction coefficients, z is the altitude coordinate, z_{lid} is the lidar altitude and $Ms(z)$ is the altitude dependent multiple scattering factor (which may, in general, be different for both the Rayleigh and Mie signals). Both the Rayleigh backscatter and extinction are well-understood functions of the atmospheric density profile and the calibration coefficients are assumed to be known to within some associated uncertainty

Taking the logarithmic derivative of Eqn.(1) and rearranging gives

$$
\begin{aligned}
\alpha_M(z) &= -\alpha_R(z) \\
&- \frac{1}{2}\frac{d}{dz}\log\left[\frac{p_R(z)(z_{lid}-z)^2}{\beta_R(z)} Ms_M(z)\right]
\end{aligned} \tag{3}
$$

Thus, assuming that the multiple scattering factors can be estimated (for example, in an iterative fashion with the aid of an analytical model such as that due to [2] or [3]) both the extinction and backscatter can be estimated in a simple direct fashion. However, the procedure just sketched out is only viable when high SNR measurements are being dealt with. In order to process further, we must also consider how to estimate extinction using the Mie signal.

1.0.1. *Klett-Like Inversion*

If we define:

$$
P_M(z) \equiv Ms_M(z)(z_{lid}-z)^2 p_M(z) S(z) e^{2\int_0^2 \alpha_R(z')dz'} \tag{4}
$$

and

$$
P_R(z) \equiv Ms_R(z)(z_{lid}-z)^2 p_R(z) e^{2\int_0^2 \alpha_R(z')dz'} \tag{5}
$$

where $S(z) = \frac{\alpha(z)}{\beta(z)}$ then Eqns. (1) and (2) become, respectively,

$$
P_M(z) = C_M \alpha(z) \exp[-2\tau(z_{lid}, z)] \tag{6}
$$

and

$$
P_R(z) = C_R \beta_R(z) \exp[-2\tau(z_{lid}, z)] \tag{7}
$$

where $\tau(z_1, z) = \int\limits_{z_1}^{z} \alpha_M(z')dz$.

© Proceedings of the 8th International Symposium on Tropospheric Profiling, ISBN 978-90-6960-233-2
Delft, The Netherlands, October 2009. Editors, A. Apituley, H.W.J. Russchenberg, W.A.A. Monna

Noting that $\frac{d\tau}{dz} = \alpha(z)$, Eqn.(6) can be written as

$$P_M(z) = C_M \frac{d\tau(z)}{dz} \exp[-2\tau(z_{lid}, z)] \quad (8)$$

which is a differential equation whose solution in terms of α can be written as

$$\alpha_M(z) = -\frac{1}{2}\left[\frac{P_M(z)}{\frac{P_M(z_o)}{\alpha_o} - 2\int\limits_{z_o}^{z} P_M(z')dz'}\right] \quad (9)$$

Where α_o is the lidar extinction coefficient at some boundary range z_o. α_o is related to the Mie signal calibration factor as

$$C_M = \frac{P_M(z_o)}{\alpha_o \exp[-2\int_{z_{lid}}^{z} \alpha_M(z')dz']} \quad (10)$$

Eqn.(9) is similar to the well-know Klett-Fernald inversion equation [5], except here the range dependent extinction-to-backscatter ratio and multiple scattering factor has been absorbed into the definition of $P_M(z)$ (see Eqn.(5)).

If the Lidar ratio ($S(z)$) can be specified, along with $Ms_M(z)$, α_o and z_o, then the extinction profile can be estimated using Eqn.(9) . Further, Eqn.(10) can then be used to estimate C_M and C_R[1] which, along with the estimated extinction profile, can be used to predict the Rayleigh signal (Eqn.(4)). This process is necessarily iterative as the multiple-scattering factor is itself a function of the extinction.

2. OPTIMAL ESTIMATION PROCEDURE

Eqn.(9) forms the basis of the inversion procedure. However, a feasible approach to specifying $S(z)$ along with $Ms_M(z)$ and α_o must be employed. The basic idea used in this work is to retrieve the extinction profile from the Mie channel using the S profile that is most consistent with observed Rayleigh signal. To do this, we cast the problem in an optimal estimation (OE) framework [4]. Accordingly, we seek to find the solution vector $\mathbf{x}$ which minimizes the cost-function

$$\begin{aligned} C_f &= [\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \mathbf{S_e^{-1}} [\mathbf{y} - \mathbf{F}(\mathbf{x})] \\ &+ [\mathbf{x} - \mathbf{x_a}]^{\mathbf{T}} \mathbf{S_a^{-1}} [\mathbf{x} - \mathbf{x_a}] \end{aligned} \quad (11)$$

where:

$\mathbf{y} = (C_{M,o}, P_{R,o}(z_1), P_{R,o}(z_2)...P_{R,o}(z_{nz}))^T$ is the observation vector where the 'o' subscript denotes an observed quantity. In particular, $P_{R,o} = p_{R,o,i} Ms_{r,i}$.

$\mathbf{x} = (\ln(\alpha_o), \ln(S(z_1)), \ln(S(z_2)) .... \ln(S(z_{nz})))^T$ is the state vector. The logarithm of the variables is used so that the retrieved quantities will be positive.

$\mathbf{F}(\mathbf{x})$ is the forward model. In our particular case, $\mathbf{F}(\mathbf{x}) = (C_M, P_R(z_1), P_R(z_2).....P_R(z_{nz}))^T$.

[1]The relationship between the Rayleigh and Mie channel calibration factors is assumed to be well-know on the basis of instrument considerations and/or in flight calibration procedures.

$\mathbf{S_e}$ is the combined forward model and observation error covariance matrix and

$\mathbf{S_a}$ is the a priori error covariance matrix and

$\mathbf{x}_a = (\ln(\alpha_{o,a}), \ln(S_a(z_1)), \ln(S_a(z_2)) .... \ln(S_a(z_{nz})))^T$ is the a priori state vector

The minimization of the cost-function (Eqn.(11)) is accomplished using a procedure build around a general implementation of the Levenberg-Marquardt algorithm (See Section 5.7 in [4])

2.0.2. *Forward model*

In this work the forward model for matching the first term in the observation vector corresponds to a discretized version of Eqn.(10), i.e.

$$C_M = \frac{P_{M,i_o}}{\alpha_o \exp[-2\alpha_j \Delta z^j|_{i_o \le j \le i_{lid}}]} \quad (12)$$

where the subscripts denote the range-gate and we are utilizing the Einstein summation convention (where repeated upper and lower indices imply summation over that index), $\mathbf{\Delta z}$ is the range gate width vector, i_o is the height index corresponding to the boundary range and i_{lid} is the height index corresponding to the lidar altitude.

For the remaining elements of the forward model vector we have (from Eqn.(7))

$$P_{R,i} = C_R \beta_{R,i} \exp[-2\alpha_j \Delta z^j|_{i \le j \le i_o}] \quad (13)$$

In the above the elements of α are given by a discrete form of Eqn.(9) i.e.

$$\alpha_i = -\frac{1}{2}\left[\frac{P_{M,i}}{\frac{P_{M,i_o}}{\alpha_o} - 2P_{M,j}\Delta z^j|_{i_o \le j \le i}}\right] \quad (14)$$

where

$$P_{M,i} = Ms_{M,i}(z_{lid} - z_i')^2 p_{M,i} S_i \exp[-2\tau_i] \quad (15)$$

The multiple scattering correction factors (the elements of $\mathbf{Ms}_M$ and $\mathbf{Ms}_R$) are calculated using the model of [3] which has been validated against detailed Monte-Carlo lidar simulations as well as other analytical approaches.

2.0.3. *A priori Information*

In practice, z_o is chosen to be a point furthest away from the lidar where the SNR is greater than around 2. α_o is then crudely estimated from the observed signal itself using Eqn.(12) together with the expected value of C_{lid} assuming that there is no aerosol or cloud extinction and using an assumed value for S_o and setting $Ms_{M,o}$ to 1. This generally yields a low value for α_o with a large error range but it is sufficient for our purposes. The corresponding entry in the a priori error covariance matrix $\mathbf{S_a}[1,1]$ is set to a large value (generally corresponding to 100 times the estimate of α_o).

The remaining values of the a priori error covariance matrix $\mathbf{S_a}$ and state vector $\mathbf{x_a}$ must be set according to the target in question. In particular, the range dependent lidar-ratio values ($S_{a,i}$) and the corresponding entries in $\mathbf{S_a}$ will vary depending on whether the target present at a given range is ice cloud, water cloud or a particular aerosol type. For this reason, the OE procedure uses a internal classification procedure that decides between ice cloud, water cloud and aerosol. The a priori terms are then specified based on the classification. The classification procedure is also needed to decide upon the effective particle size profile used in the multiple scattering correction step.

3. EXAMPLE APPLICATION

The OE procedure has been integrated into the EarthCARE Simulator (ECSIM) environment. ECSIM is an "end-to-end" simulation environment focused on (but not limited to) the EarthCARE platform. The simulator contains a 3-D lidar Monte-Carlo forward model for the detailed accurate modeling of general lidar signals. As well, a separate instrument model is used to model the instrument response using the output of the forward model as input. For more information on ECSIM please see the ECSIM Models and Algorithms Document (`ECSIM_MAD.pdf`) which can be downloaded from `ftp:bbc.knmi.nl,user:simguest, Password:S139st, Directory:ECSIMV1P3`.

An example ECSIM extinction field is shown in the Top–Left panel of Fig.1. Here the extinction field corresponds to a "fractal" cirrus cloud with extinction values at 353 nm ranging between minimum values on the order of 0.1 km^{-1} to maximum values on the order of 2-3 km^{-1} The abrupt cirrus cloud top and bottom boundaries are an artifact of the method used to generate the fractal field [6].

Simulated ATLID signals generated by ECSIM are shown in the middle panels of Figs. 1 and (for completeness the simulated radar reflectivity is shown in the Top-Right panel of Fig.1). Using the simulated lidar signals the extinction was derived using the OE procedure as well as by a more conventional Rayleigh-only method based on Eqn.(3). These results are shown in the bottom panels. By comparing the respective results to the "truth" (Top–Left panel of Fig.1) it can be seen that the OE results are more accurate and less noisy than the corresponding Rayleigh only results.

4. SUMMARY

A new approach for the retrieval of extinction and backscatter from space-borne HSRL lidar has been developed. The procedure attempts to combine the best features of both direct Rayleigh-only approaches with the best features of Klett-type inversions. The core of the retrieval procedure involves solving a semi-analytical modified form of the well-known Klett equation which includes provisions for an unknown range dependent cloud/aerosol backscatter-to-extinction ratio (S). This equation is used to invert the Mie channel signal (including a procedure for correcting for the effects of multiple scattering) and the results are then used to predict the signal observed in the Rayleigh channel. The procedure uses Optimal-Estimation (OE) techniques in order to solve the for the unknowns in the Klett-like inversion so that the best-feasible-match with the observed Rayleigh channel signal is obtained

In preliminary testing within ECSIM, the new approach shows promise and may be considered a viable candidate for future application to actual ATLID data. Further development of the approach outlined here is planned to be carried out using both simulated and actual data.

ACKNOWLEDGMENTS

This work was carried out as part of the ESA sponsored CASPER study.

REFERENCES

[1] Piironen, P. and E. W. Eloranta, 1994: Demonstration of a high-spectral-resolution lidar based on a iodine absorption filter. *Opts. Lett.*, **19**, 234–236.

[2] Eloranta, E., 1998: Practical model for the calculation of multiply scattered lidar returns. *Appl. Opts.*, **37**, 2464–2472.

[3] Hogan, R., 2006: Fast approximate calculation of multiply scattered lidar returns. *Appl. Opts.*, **45**, 5984–5992.

[4] Rodgers, C., 2000: *Inverse methods for Atmospheric Sounding; Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics*, volume 2. World Scientific, 1 edition, 169 pp.

[5] Klett, J., 1985: Lidar inversion with variable backscatter/extinction ratios. *Appl. Opts.*, **24**, 1638–1643.

[6] Hogan, R. and S. Kew, 2005: A 3d stochastic cloud model for investigating the radiative properties of inhomogeneous cirrus clouds. *Q. J. R. Meteorol. Soc.*, **131**, 2585–2608.

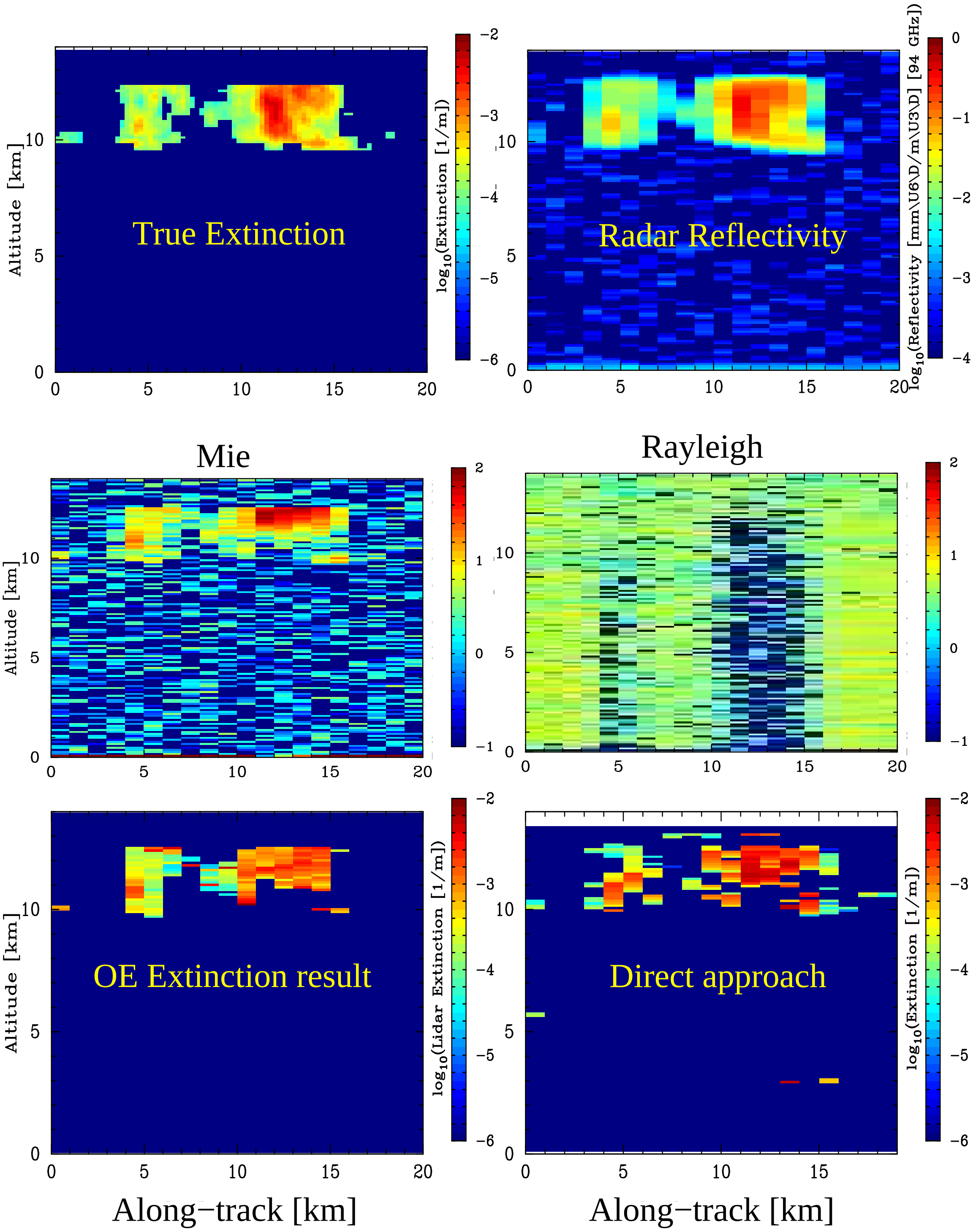

Figure 1. Ideal extinction (Top–Left), Radar reflectivity at 1-km horizontal resolution (Top–Right), Cross-talk and background corrected lidar Mie channel signals at 1-km horizontal resolution (Middle-Left), Cross-talk and background corrected lidar Rayleigh channel signals at 1-km horizontal resolution (Middle-Right), Extinction field retrieved using the OE procedure (Bottom-Left) and Extinction field retrieved using a direct approach (via a form of Eqn.(3) using the Rayleigh signal only (Bottom-Right).