



# Retrieval of aerosol backscatter, extinction, and lidar ratio from Raman lidar with optimal estimation

A. C. Povey<sup>1</sup>, R. G. Grainger<sup>1</sup>, D. M. Peters<sup>1</sup>, and J. L. Agnew<sup>2</sup>

<sup>1</sup>Atmospheric, Oceanic, and Planetary Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK

<sup>2</sup>STFC Rutherford-Appleton Laboratory, Harwell, Oxford OX11 0QX, UK

Correspondence to: A. C. Povey (povey@atm.ox.ac.uk)

Received: 29 August 2013 – Published in Atmos. Meas. Tech. Discuss.: 30 October 2013

Revised: 21 January 2014 – Accepted: 6 February 2014 – Published: 13 March 2014

**Abstract.** Optimal estimation retrieval is a form of nonlinear regression which determines the most probable circumstances that produced a given observation, weighted against any prior knowledge of the system. This paper applies the technique to the estimation of aerosol backscatter and extinction (or lidar ratio) from two-channel Raman lidar observations. It produces results from simulated and real data consistent with existing Raman lidar analyses and additionally returns a more rigorous estimate of its uncertainties while automatically selecting an appropriate resolution without the imposition of artificial constraints. Backscatter is retrieved at the instrument's native resolution with an uncertainty between 2 and 20%. Extinction is less well constrained, retrieved at a resolution of 0.1–1 km depending on the quality of the data. The uncertainty in extinction is > 15%, in part due to the consideration of short 1 min integrations, but is comparable to fair estimates of the error when using the standard Raman lidar technique.

The retrieval is then applied to several hours of observation on 19 April 2010 of ash from the Eyjafjallajökull eruption. A depolarising ash layer is found with a lidar ratio of 20–30 sr, much lower values than observed by previous studies. This potentially indicates a growth of the particles after 12–24 h within the planetary boundary layer. A lower concentration of ash within a residual layer exhibited a backscatter of  $10 \text{ Mm}^{-1} \text{ sr}^{-1}$  and lidar ratio of 40 sr.

## 1 Introduction

Aerosols impact the Earth's radiation budget both directly, by reflecting solar radiation back into space (Haywood and Shine, 1995), and indirectly, by altering the properties and distribution of clouds (Lohmann and Feichter, 2005) or reacting with other species (Colbeck, 1998). The lack of knowledge about the global distribution and composition of aerosols is currently the single greatest source of uncertainty in estimates of net radiative forcing and therefore is a factor in the ability to predict the impacts of climate change (IPCC, 2007).

Lidar (light detection and ranging) is an active remote sensing technique for observing the distribution of molecules and particles in the atmosphere as a function of height by means of the light they backscatter from a laser beam. The name intentionally emulates radar as both techniques use the time of flight of a pulsed source to deduce the distance to the scatterer (Fugii and Fukuchi, 2005). Despite its exceptionally high spatial and temporal resolution, lidar is not as widely applied as other techniques in the study of aerosol's effect on climate as a single lidar samples only one location. With the launch of a space-based lidar (Winker et al., 2009) and the development of networks across northern America (Welton et al., 2000), Europe (Pappalardo et al., 2005), and Asia (Sugimoto et al., 2008), the importance and availability of lidar data increases.

The energy observed by a lidar is a function of the extinction and backscattering coefficients – the cross section per unit volume to either attenuate the beam or to scatter light directly back towards the instrument. These coefficients are functions of the microphysical properties of the aerosol

present, such as refractive index and size distribution. Deriving such properties directly is possible, but the problem is very poorly constrained. Its solution either requires a greater number of measurements, such as a multi-wavelength system (Müller et al., 1999), or further assumptions about the scatterers. This more complex problem is disregarded here in favour of the better-constrained estimation of extinction and backscatter.

Optimal estimation retrieval is a form of nonlinear regression which determines the most probable circumstances that produced a given observation, weighted against any prior knowledge of the system. For several decades, it has been successfully applied to the analysis of satellite (e.g. Marks and Rodgers, 1993; Li et al., 2008; Watts et al., 2011), radar (Grant et al., 2004), and ground-based radiometer observations (e.g. Guldner and Spankuch, 2001) but has not seen substantial use within the lidar community. This paper applies the technique to the estimation of aerosol extinction and backscatter from two-channel Raman lidar observations. The retrieval processes the entire profile simultaneously, making optimal use of the information available and choosing the most appropriate vertical resolution for the result while fully characterising the covariant error due to measurement noise, model error, and other assumptions. This widely recognised retrieval algorithm, which is less dependant on ad hoc corrections and assumptions while providing rigorous error estimation, provides an alternative to the traditional Raman lidar technique that can make optimal use of the available information.

Section 2 outlines the retrieval algorithm and existing analysis methods. Section 3 evaluates the retrieval's performance against existing techniques with simulated data and considers the error budget. Section 4 applies the algorithm to observations, while Sect. 5 provides some conclusions.

## 2 Methods

### 2.1 Optimal estimation retrieval

As outlined in Rodgers (2000), optimal estimation solves the inverse problem

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\epsilon}, \quad (1)$$

where  $\mathbf{y}$  is a column vector describing the measurements;  $\boldsymbol{\epsilon}$  gives the noise on that measurement; and the forward model  $\mathbf{F}(\mathbf{x}, \mathbf{b})$  translates a state of the instrument and atmosphere, summarised by unknown parameters  $\mathbf{x}$  and known parameters  $\mathbf{b}$ , into a simulated measurement.

Approximating the probability density function (PDF) for all quantities as Gaussian and using Bayes' theorem, the probability that the system has a state  $\mathbf{x}$  given the

measurement  $\mathbf{y}$  can be written as

$$-2 \ln P(\mathbf{x}|\mathbf{y}) = [\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{b})]^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{b})] + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a], \quad (2)$$

where the covariance matrix  $\mathbf{S}_\epsilon$  describes the random experimental error and  $\mathbf{x}_a$  is the a priori, the state expected before the measurement is made. The uncertainty in that expectation is described by the a priori covariance  $\mathbf{S}_a$ . The quantity  $-2 \ln P(\mathbf{x}|\mathbf{y})$  is hereafter referred to as the cost as it measures the goodness of fit for a solution. Good models should have a cost approximately equal to the number of measurements. Hence, the cost will herein be quoted normalised by the length of  $\mathbf{y}$ .

It can be shown that the iteration

$$\mathbf{x}_{i+1} = \mathbf{x}_i + [(1 + \Gamma_i) \mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i]^{-1} \{ \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}_i, \mathbf{b})] - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a) \} \quad (3)$$

converges to the most probable state  $\hat{\mathbf{x}}$ , where the Jacobian  $\mathbf{K}_i = \partial \mathbf{F} / \partial \mathbf{x}_i$  and  $\Gamma_i$  is a scaling constant. General practice, outlined in Fig. 1, is that if the cost increases after an iteration,  $\Gamma_i$  is increased by a factor of 10. Otherwise, it is reduced by a factor of 2. Evaluation ceases after

- the cost function decreases by much less than the number of measurements;
- the cost decreased and the change in the state is much less than the predicted error,
 
$$(\mathbf{x}_{i+1} - \mathbf{x}_i)_j^2 \ll (\mathbf{S})_{jj} \quad \forall j,$$
 where the error covariance matrix of the solution  $\mathbf{S} = (\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i + \mathbf{S}_a^{-1})^{-1}$ ;
- the step in state space,
 
$$\{ \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}_i, \mathbf{b})] - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a) \} (\mathbf{x}_{i+1} - \mathbf{x}_i)^T,$$
 is much less than the length of the state vector;
- 30 iterations, which is considered a failure to converge.

The averaging kernel,

$$\mathbf{A} = \hat{\mathbf{S}} \hat{\mathbf{K}}^T \mathbf{S}_\epsilon^{-1} \hat{\mathbf{K}}, \quad (4)$$

describes the extent to which the true and a priori states each contribute to the solution as it can be shown that

$$\hat{\mathbf{x}} = \mathbf{A} \mathbf{x} + (\mathbf{I} - \mathbf{A}) \mathbf{x}_a + \hat{\mathbf{S}} \hat{\mathbf{K}}^T \mathbf{S}_\epsilon^{-1} \boldsymbol{\epsilon}, \quad (5)$$

where a hat indicates the value after convergence. An ideal retrieval would have a kernel equal to the identity. In practice, the rows of  $\mathbf{A}$  will be peaked functions showing how the information in one retrieved bin is derived from an average of the true values around it. The width of that peak is therefore a measure of the resolution of the retrieval.

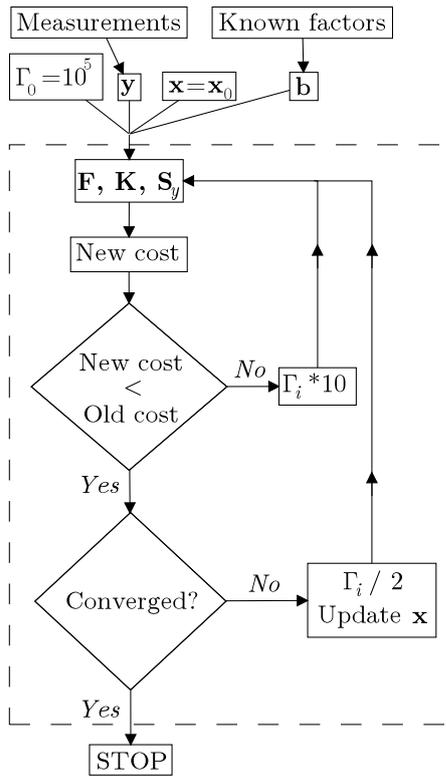


Fig. 1. Schematic of the optimal estimation retrieval algorithm.

### 2.2 Existing lidar analyses

The number of photons observed from a height  $R$  is expressed by the lidar equation (Measures, 1992)

$$E(R, \lambda) = \frac{E_L C(R, \lambda)}{R^2} \beta(R, \lambda) \exp \left[ - \int_0^R \alpha(R', \lambda_L) + \alpha(R', \lambda) dR' \right], \quad (6)$$

where  $\beta(R, \lambda)$  is the coefficient for incident laser light, wavelength  $\lambda_L$ , to be backscattered at a wavelength  $\lambda$ ;  $\alpha(R, \lambda)$  is the extinction coefficient;  $E_L$  is the number of photons emitted by the laser pulse; and  $C(R)$ , known as the calibration function, describes the alignment and efficiency of the detection system. As both the extinction and backscatter are unknown, a single profile presents an underconstrained measurement.

The atmosphere is assumed to contain only two components such that

$$\beta = \beta^{(p)} + \beta^{(m)} \quad (7)$$

$$\alpha = \alpha^{(p)} + \alpha^{(m)}, \quad (8)$$

where  $\beta^{(p)}$ ,  $\alpha^{(p)}$  denote backscattering and extinction by aerosols and  $\beta^{(m)}$ ,  $\alpha^{(m)}$  denote scattering by molecules, which is well modelled by Rayleigh scattering.

The dominant return for any lidar is the elastic profile (where  $\lambda = \lambda_L$ ), from which the backscatter is commonly derived by a technique known as onion peeling or the Fernald–Klett method (Klett, 1981; Fernald, 1984). A Raman lidar monitors a second channel containing the Raman scattering from a single species in the atmosphere, such that  $\beta$  becomes a known function of number density. Ansmann et al. (1992) outlined a means to invert Eq. (6) in such circumstances to derive the extinction and backscatter separately.

A few applications of nonlinear regression to lidar have been published, though these retrieve the microphysical properties of aerosol rather than the optical properties outlined above or later in this paper. A retrieval of ice water path and effective radius in cirrus clouds from coincident, space-borne lidar and radar measurements was developed in Delanoe and Hogan (2008), though its results were found to be highly dependant on the microphysical assumptions. Pounder et al. (2012) derived high-quality extinction retrievals from three simultaneous observations with different fields of view using a linearised model of the lidar equation that included multiple scattering while applying Twomey–Tikhonov smoothing rather than an a priori. Marchant et al. (2010) presented an original, if limited, linearised scheme that decomposed scattering over a basis of precomputed aerosols. This was expanded to a retrieval of effective radius in multiwavelength studies via a Kalman filter in Marchant et al. (2012).

A related method known as regularisation has also been used to derive extinction and backscatter from Raman lidar profiles. The introduction of Veselovskii et al. (2002) provides a good review of early attempts and the methodology. Shcherbakov (2007) and Pornsawad et al. (2012) demonstrated that such methods return solely positive extinction and can produce more accurate products than the Ansmann method but use Tikhonov smoothing within the retrieval, which generates significant errors where there are substantial gradients. Though the formulation of that technique resembles that of optimal estimation, they differ in their interpretation. As the Tikhonov matrix  $\mathbf{H}$  is singular, the smoothing term  $\mathbf{H}^T \mathbf{H}$  has no inverse that would correspond to a  $\mathbf{S}_a$ . Hence, the desired smoothing is introduced in a manner that has no physical analogue within the measurement system. Evaluation also requires a set of basis functions to be defined, artificially imposing a structure onto the system.

It is preferable to impose the basis for smoothing solutions through an a priori covariance matrix derived from actual data and the physical processes driving the system, as facilitated by optimal estimation retrieval. The impact of these assumptions can be assessed through the averaging kernel, such that it is clear where the data are the dominant influence on the solution. Other techniques do not provide such a balance and, in fact, rarely discuss the choice of basis functions or their impact on the solution.

### 2.3 Forward model

Lidars frequently use photomultiplier tubes (PMTs) as their detector, which produce a cascade of electrons when struck by a photon. If the rate of photons is less than two per bin, noise can be very effectively removed by applying a threshold to the output to return a count of the number of photons per bin. As count rates increase, multiple pulses are more likely to overlap and be counted only once, such that this mode becomes increasingly nonlinear (Müller, 1973). The most frequent correction for this (Whiteman et al., 1992) assumes that after any count, the detector will be unable to detect another for a “dead time”  $\tau_d$  and that that time is independent of the count rate (known as the non-paralysable correction). The observed profile can then be expressed as

$$\varphi_i = \frac{ME_i}{1 + \tau_d \tau_b^{-1} E_i}, \quad (9)$$

where subscript is used to denote a function of range,  $f(R_i) \equiv f_i$ ;  $M$  is the number of laser shots  $E$  is averaged over; and  $\tau_b$  is the duration of a bin (such that  $R_i = \frac{1}{2} i c \tau_b$ ). This will apply to both channels, though each may have a distinct value of  $\tau_d$ .

For large count rates, it is also possible to operate the PMT in an analogue mode, which simply averages the output current during each range bin. This is linear over a large dynamic range but suffers additional noise from thermal excitations, variations in pulse height, electrical interference, and other effects. In such circumstances,

$$\varphi_i = aE_i + b, \quad (10)$$

where  $a$  and  $b$  are constants and  $\varphi$  is measured in volts (whereas it is photon counts in Eq. 9).

The correction of Eq. (9) is increasingly unreliable for count rates greater than 10 MHz. When both detection modes are operated simultaneously, it is common practice to “glue” the two signals together using analogue observations near the instrument, where the signal is largest, and photon counting elsewhere (e.g. Newsom et al., 2009). The gluing procedure in essence finds appropriate values for the constants in Eq. (10). Such techniques will not be used in this paper as the two signals will be considered separately. A simple extension of this work would be to include both Eqs. (9) and (10) in the forward model and retrieve using both signals simultaneously.

It is not necessary to retrieve the extinction and backscatter at the native resolution of the instrument. The state vector can be defined on any arbitrary grid,  $r$ , and then interpolated onto the instrument’s range axis  $R$ . For example, a grid with spacing that increases as a function of height could be used such that all retrieved values have uncertainties of a similar magnitude. The use of a coarser grid will also reduce the computational expense of the retrieval. For simplicity, only a regular 33 m grid is considered in this paper, though the general expressions are presented to most accurately represent

the model used in the calculations. No significant difference has been found between interpolated solutions and those at full resolution.

Neglecting multiple scattering and assuming  $\alpha^{(p)} \propto \lambda^{-1}$  (Ansmann et al., 1992), the number of photons observed from range bin  $R_i$  will be

$$E_i^{(L)} = E_L \frac{C_i^{(L)}}{R_i^2} \left[ \frac{\sigma_R^{(L)}}{\mathcal{B}} N_i + \text{spline}_{r \rightarrow R_i} [\tilde{\beta}^{(p)}] \right] \exp \left[ -2 \left( \sigma_R^{(L)} \mathcal{N}_i + \text{spline}_{r \rightarrow R_i} [\tilde{\chi}] \right) \right] + E_B^{(L)}, \quad (11)$$

$$E_i^{(ra)} = E_L \frac{C_i^{(ra)}}{R_i^2} N_i \exp \left[ - \left( \sigma_R^{(L)} + \sigma_R^{(ra)} \right) \mathcal{N}_i - \left( 1 + \frac{\lambda_L}{\lambda_{ra}} \right) \text{spline}_{r \rightarrow R_i} [\tilde{\chi}] \right] + E_B^{(ra)}, \quad (12)$$

where superscript denotes functions of wavelength,  $f(\lambda_X) \equiv f^{(X)}$ ;  $\sigma_R$  is the cross section for Rayleigh scattering, which has lidar ratio  $\mathcal{B} = 8\pi/3$ ;  $N$  is the atmospheric number density;  $\mathcal{N}_i = \int_0^{R_i} N(R') dR'$ ; and  $E_B$  is the background count rate, which is estimated from observations as  $R \rightarrow \infty$ . The aerosol optical thickness  $\chi = \int_0^R \alpha^{(p)}(\lambda_L, R') dR'$  and  $\beta$  are evaluated at  $\lambda_L$ , though this dependence is dropped for brevity. The calibration function  $C$  is assumed known and is input as a parameter.

A tilde is used to represent variables on the retrieved grid  $r$ , which are interpolated using the cubic spline method of Press et al. (1992) onto the measured grid  $R$ . The aerosol optical thickness is evaluated on grid  $r$  with the trapezium rule,

$$\tilde{\chi}_j = \tilde{\alpha}_0^{(p)} r_0 + \frac{1}{2} \sum_{k=1}^j [\tilde{\alpha}_k^{(p)} + \tilde{\alpha}_{k-1}^{(p)}] [r_k - r_{k-1}], \quad (13)$$

and then interpolated onto  $R$ . Note that the extinction is assumed constant through the first bin, such that it acts as a boundary term rather than a physically meaningful value. This avoids various difficulties with observation very near the instrument.

Though this problem could be linearised, there will be some error involved in that approximation. Solving the non-linear problem presented is not an overly intensive calculation and so there is no need to simplify the problem.

Extinction and backscatter are both functions of  $N$  and so will be correlated. This should be identified within  $\mathbf{S}_a$  but cannot be easily estimated. Further, the use of correlated variables will emphasise degenerate states of the forward model, which can slow the retrieval’s convergence. This can be averted by retrieving the lidar ratio  $B = \alpha^{(p)}/\beta^{(p)}$  instead (Shcherbakov, 2007; R. Hogan, personal communication, 2012), which is independent of  $N$ ,

$$\tilde{\chi}_j = \tilde{\beta}_0^{(p)} \tilde{B}_0 r_0 + \frac{1}{2} \sum_{k=1}^j [\tilde{\beta}_k^{(p)} \tilde{B}_k + \tilde{\beta}_{k-1}^{(p)} \tilde{B}_{k-1}] [r_k - r_{k-1}]. \quad (14)$$

All elements of  $\mathbf{x}$  should be positive or there will be a degeneracy in the impact of backscatter and optical depth on the elastic channel, which can impede the retrieval. For the retrieval of  $\beta^{(p)}$  and  $\alpha^{(p)}$ , this will be prevented by setting all negative values to zero after evaluating Eq. (3). Though an unsophisticated solution, it is found that values that are zeroed in one iteration generally increase in the next. It is unusual that a pixel retrieves a negative value in the final iteration without also having a large error. For the retrieval of  $\beta^{(p)}$  and  $B$ , it was found preferable to instead retrieve  $\ln \beta^{(p)}$  while enforcing a lower limit of unity on  $B$ . Retrieving the logarithm eliminates negative values from the solution but can contort state space, retarding convergence.

The measurement and state vectors are then

$$\mathbf{y} = \begin{pmatrix} \varphi_0^{(L)} \\ \varphi_1^{(L)} \\ \vdots \\ \varphi_{m-1}^{(L)} \\ \varphi_0^{(ra)} \\ \varphi_1^{(ra)} \\ \vdots \\ \varphi_{m-1}^{(ra)} \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} \tilde{\beta}_0^{(p)} \\ \tilde{\beta}_1^{(p)} \\ \vdots \\ \tilde{\beta}_{n-1}^{(p)} \\ \tilde{\alpha}_0^{(p)} \\ \tilde{\alpha}_1^{(p)} \\ \vdots \\ \tilde{\alpha}_{n-1}^{(p)} \end{pmatrix} \quad \parallel \quad \begin{pmatrix} \ln \tilde{\beta}_0^{(p)} \\ \ln \tilde{\beta}_1^{(p)} \\ \vdots \\ \ln \tilde{\beta}_{n-1}^{(p)} \\ \tilde{B}_0 \\ \tilde{B}_1 \\ \vdots \\ \tilde{B}_{n-1} \end{pmatrix}.$$

The first guess for  $\mathbf{x}$  in the iteration Eq. (3) is taken as  $\beta^{(p)} = 10^{-5} \text{ Mm}^{-1} \text{ sr}^{-1}$  and  $B = 58 \text{ sr}$ . These values were chosen as they tend to reduce the number of iterations. Their value does not affect the final result (if the retrieval converges).

One final note must be made of the treatment of measurement error (which is assumed uncorrelated),  $\mathbf{S}_\epsilon$ . The observed photon counts should be Poisson distributed, such that their variance is equal to their mean. This is widely used to justify approximating the variance of a lidar measurement with the measurement itself. This is not strictly valid as the measurement is only a single sample of a distribution. However, a lidar sums profiles over several seconds or minutes of laser shots during data collection, giving no further measure of their variance.

The optimal estimation scheme requires an unbiased estimate of the variance. Using the measurement itself causes the retrieval to favour observations that coincidentally suffer large, positive noise as they appear to be more precise. This effect is most pronounced at low signal levels and introduces a high bias into the retrieval. To alleviate this, the variance will be estimated by the application of a five-bin, sliding-window average through the data. The impact of this will not be explored in detail, though preliminary studies found that even minimal smoothing of the variance vastly reduced biases. A more involved statistical analysis could be applied, considering the correlation of adjacent bins, but is not deemed necessary.

For further details, derivations, and justification of the forward model, please consult Chapter 2 of Povey (2013).

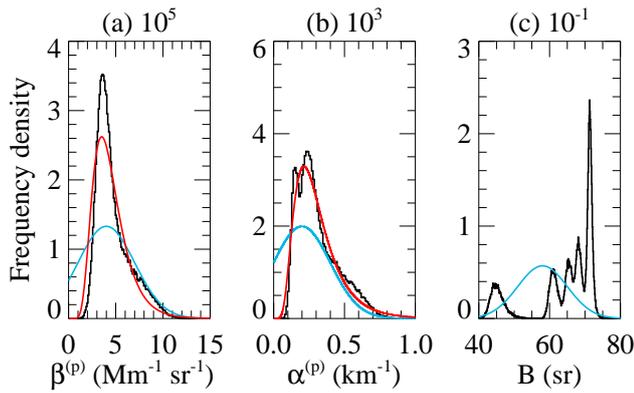
## 2.4 A priori

Arguably the most important component of an optimal estimation scheme is its a priori. Ideally, the a priori would not greatly affect the retrieval, but in practice it constrains which states are deemed to be both physically possible and likely. In this problem, solutions should be reasonably smooth as aerosols are often well mixed through the planetary boundary layer (PBL), but gradients should not be completely excluded as layering does occur.

The exact composition and optical properties of aerosol are highly variable such that a detailed a priori is unlikely to be representative of the broad range of potential states (Lopatin et al., 2013). Climatologies, from which an a priori would be derived in most applications, rarely exist and most lidar data are derived from elastic instruments and so will be influenced by the assumed lidar ratios. Therefore, the most appropriate a priori would be an order-of-magnitude estimate to weakly constrain the magnitude of the state vector while prescribing vertical and inter-variable correlations.

Some generalised descriptions of representative aerosol types have been explored in the literature. The OPAC model (Optical Properties of Aerosol and Cloud; Hess et al., 1998) defines size distributions, refractive indices, and number densities for a variety of cloud and aerosol particles. These provide the necessary inputs for Mie codes (Grainger et al., 2004) to calculate the extinction and backscatter. Combinations of these based on expected and observed compositions then produce characteristic aerosol mixtures, such as marine or urban.

For the data to be considered, the continental type should be appropriate – comprising soot with soluble and insoluble aerosols. An ensemble of scattering properties was constructed by randomising the abundance of these components, using the OPAC model values as the mean of a Gaussian distribution with width estimated by 10 % of that mean (the exact value assumed was found to be unimportant). A simple treatment of aspherical particles using the T-matrix code of Dubovik et al. (2006) produces an effectively identical distribution. The resulting distributions of  $\beta^{(p)}$ ,  $\alpha^{(p)}$ , and  $B$  are shown in Fig. 2. The a priori is based on qualitative fits to these, shown in blue when retrieving linearly and in red for a logarithmic retrieval, where the values are given in Table 1. Though the logarithmic retrievals appear to give a better fit to the distributions, the retrieval of  $\ln \beta^{(p)}$  and  $\ln \alpha^{(p)}$  was found to be overly constrained. Though the lidar ratio is theoretically a better description of the state, its distribution is not symmetric and so not necessarily well suited to optimal estimation. A relatively broad a priori distribution has been selected to compensate. These distributions demonstrate an approximately 95 % correlation between  $\beta^{(p)}$  and  $\alpha^{(p)}$ , which is included in  $\mathbf{S}_a$  for the linear retrieval. Though its exact value appears to be unimportant, it would be desirable to obtain a more rigorous estimate.



**Fig. 2.** A priori aerosol distributions of (a) backscatter, (b) extinction, and (c) lidar ratio. Simulations using the parameters of the OPAC model are shown in black. The blue curve represents the linear retrieval of  $x$  and the red  $\ln x$ . Fit values summarised in Table 1. The title of each plot gives a scale factor for its vertical axis.

**Table 1.** A priori values and uncertainties, as shown in Fig. 2. Mean denotes the value of  $x_a$  at  $R = 0$  and SD denotes the square root of the diagonal elements of  $S_a$ .

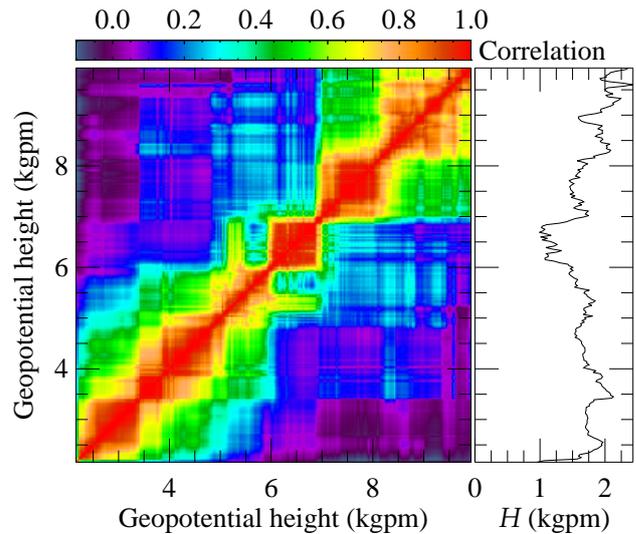
		$\beta^{(p)}$ ( $\text{m}^{-1} \text{sr}^{-1}$ )	$\alpha^{(p)}$ ( $\text{m}^{-1}$ )	$B$ (sr)
Linear	Mean	$4 \times 10^{-6}$	$2 \times 10^{-4}$	58
	SD	$3 \times 10^{-6}$	$2 \times 10^{-4}$	7
Log	Mean	-12.4	-8.2	–
	SD	0.4	0.5	–

The OPAC model states that the density of non-dust aerosols decreases exponentially with a scale height of 2 km. The prescribed values will therefore decrease similarly in  $x_a$ . Further, there will almost certainly be some vertical correlation of the measurements due to vertical mixing. The simple model of a Markov process proposed by (2.83) of Rodgers (2000) shall be used with correlations decaying exponentially with separation,

$$(S_a)_{ij} = \sqrt{(S_a)_{ii}(S_a)_{jj}} \exp\left(-\frac{|r_i - r_j|}{H}\right), \quad (15)$$

where  $H$  is a scale height.

The suitability of this scale height can be assessed by investigating the covariance of some measure of aerosol scattering. A convenient option is backscatter sondes (NDACC, 1989–2000), which measure the light backscattered from a xenon flashlamp approximately every 30 m during a balloon ascent (Rosen and Kjome, 1991; Rosen et al., 2000). Profiles over 10 years of observations at three sites have been used in Fig. 3 to produce a correlation matrix of backscatter ratio with height. Its rows decay roughly exponentially with height, which when fitted to Eq. (15) give  $H = 1$ –2 km in the free troposphere, consistent with the model.



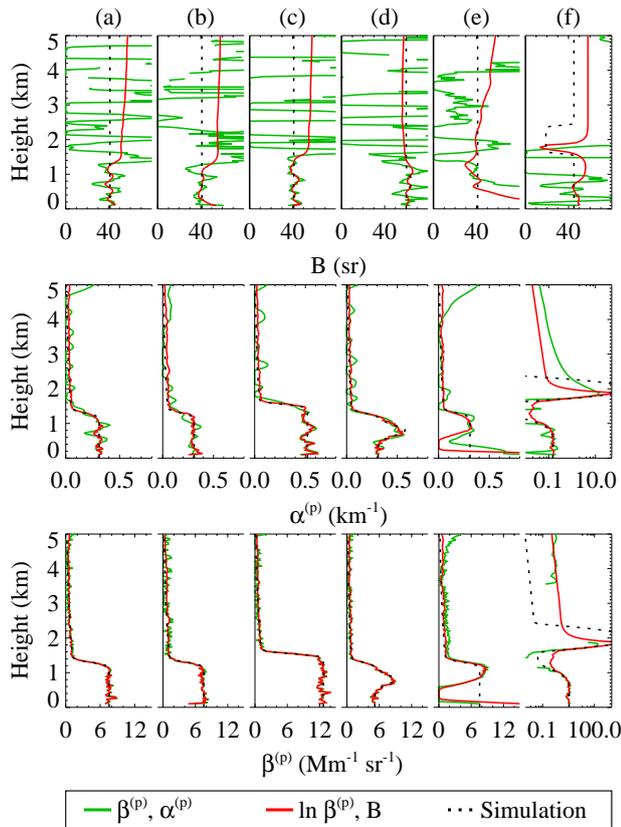
**Fig. 3.** The observed vertical correlation of backscatter. (Left) Auto-correlation of backscatter with height, derived from 198 backscatter sonde profiles collected between 1989 and 2000 at Laramie, WY, USA; Lauder, New Zealand; and Thule, Greenland. Box-like features are produced by layers of unusually large aerosol concentration during a single launch. (Right) Least-squares fit of Eq. (15) to each row of that matrix for  $H$ .

This will not necessarily apply within the PBL, which is only weakly coupled to the free troposphere (Oke, 1987). Several studies of the vertical distribution of aerosol within the PBL have been performed with tethered balloons, though the data could not be readily accessed. These generally find that aerosol concentrations are constant with height (Figs. 10–12, 4, and 2 of Ewell et al. (1989), Greenberg et al. (2009), and Ferrero et al. (2010), respectively), but occasionally observe fine structure (Fig. 6 of Ferrero et al., 2011). Further, a myriad of literature covers observations of aerosol layers tens to hundreds of metres thick (e.g. Di Girolamo et al., 1999; Dacre et al., 2011) or variations within lofted aerosol features (e.g. Althausen et al., 2000; Huang et al., 2010).

A rigorous a priori covariance matrix would represent both the general tendency for aerosols to be well mixed throughout the troposphere and the fine-scale structure that occasionally occurs. The average position of the top of the PBL would be expressed by a significant decrease in correlation between areas above and below it. At the moment, there is insufficient information to quantify these effects with any degree of certainty. As such, a conservative estimate of  $H = 100$  m is used here, which will not make the best use of the available information but does not overconstrain the solution.

### 3 Simulations

Simulated data can be easily produced with the forward model, using the NOAA (1976) standard atmosphere. The

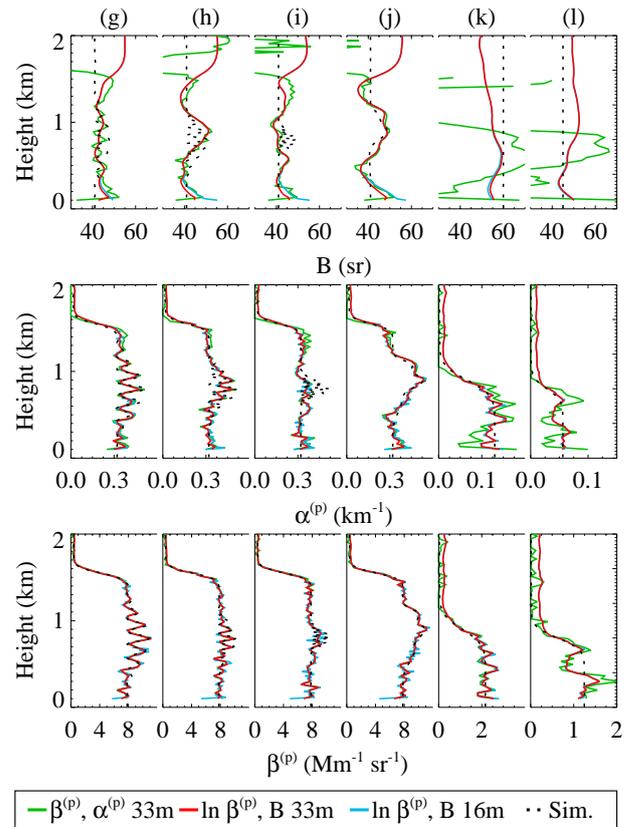


**Fig. 4.** Performance of the retrieval with simulated data for the linear (green) and logarithmic (red) retrieval modes. **(a)** An idealised, well-mixed PBL with  $\chi = 0.50$ . **(b)** As **(a)**, but observed at 20% of the previous laser energy. **(c)** Similar to **(a)**, but with  $\chi = 0.89$ . **(d)** As **(a)**, but with a larger  $B$  and the addition of an aerosol layer at 800 m. **(e)** As **(a)**, but applying an incorrect nonlinear correction. **(f)** Observation of a cloud, shown on a log scale.

PBL extinction profile is modelled by an error function (Steyn et al., 1999) multiplied by an exponential decay above the PBL. Aerosol and cloud layers are modelled by Gaussian peaks (G. Biavati, personal communication, 2011). An analytic model outlined in Povey et al. (2012) is used to generate the calibration function and detector nonlinearity. Once simulated, Poisson noise is added to the profiles.

### 3.1 Sensitivity

The retrieval from six simulated profiles by both proposed configurations is shown in Fig. 4. The two configurations give equivalent results and successfully retrieve the simulated profile in cases a–d. Cases e and f return large costs, such that it is obvious they have failed. In case e, a different nonlinear correction was used in the simulation and causes underestimation of the state where the observed profile has maximal energy. The observation of a cloud in case f is reasonable within the PBL but fails above that. The large scattering within the cloud is outside of the range prescribed by the

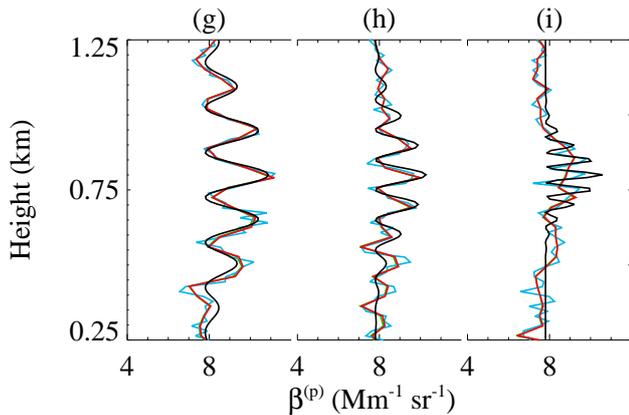


**Fig. 5.** As Fig. 4 but highlighting the sensitivity to fine-scale fluctuations. The blue curve shows the logarithmic retrieval at twice the previous resolution. **(g)** As **(a)**, but with the addition of sinusoidal “aerosol layers” of width 300 m. **(h)** As **(g)**, but width 200 m. **(i)** As **(g)**, but width 100 m. **(j)** As **(a)**, but including three overlapping layers. **(k)** As **(a)**, but with  $\chi = 0.12$ . **(l)** As **(a)**, but with  $\chi = 0.04$ .

a priori and, though it obtains a decent fit to the visible region of the cloud, vertical correlations cause incorrect retrieval beneath it. Successfully fitting cloud and aerosol observations simultaneously requires a forward model and a priori designed for the several orders of magnitude spanned by the state vector.

The lidar ratio profiles indicate that there is a decrease in the information content of the measurement above the PBL, where scattering (and therefore the magnitude of the return) is lower. The two configurations react differently to this. The lidar ratio configuration returns a smooth  $B$  profile that tends towards its a priori value above the PBL, as would be expected, while the extinction configuration gives a much noisier profile, indicating it is less influenced by the a priori.

A further six simulations containing small-scale fluctuations are presented in Fig. 5. The two configurations behave as before, with the lidar ratio mode returning a smoother profile but losing sensitivity above the PBL. The “layers” of cases g and h are correctly positioned by both modes, if slightly underestimated in magnitude. In case i, the layers are



**Fig. 6.** A closer examination of cases g–i of Fig. 5, highlighting how the coarse retrievals fail to capture the smallest features and that all resolutions tend to smooth the magnitude of the peaks.

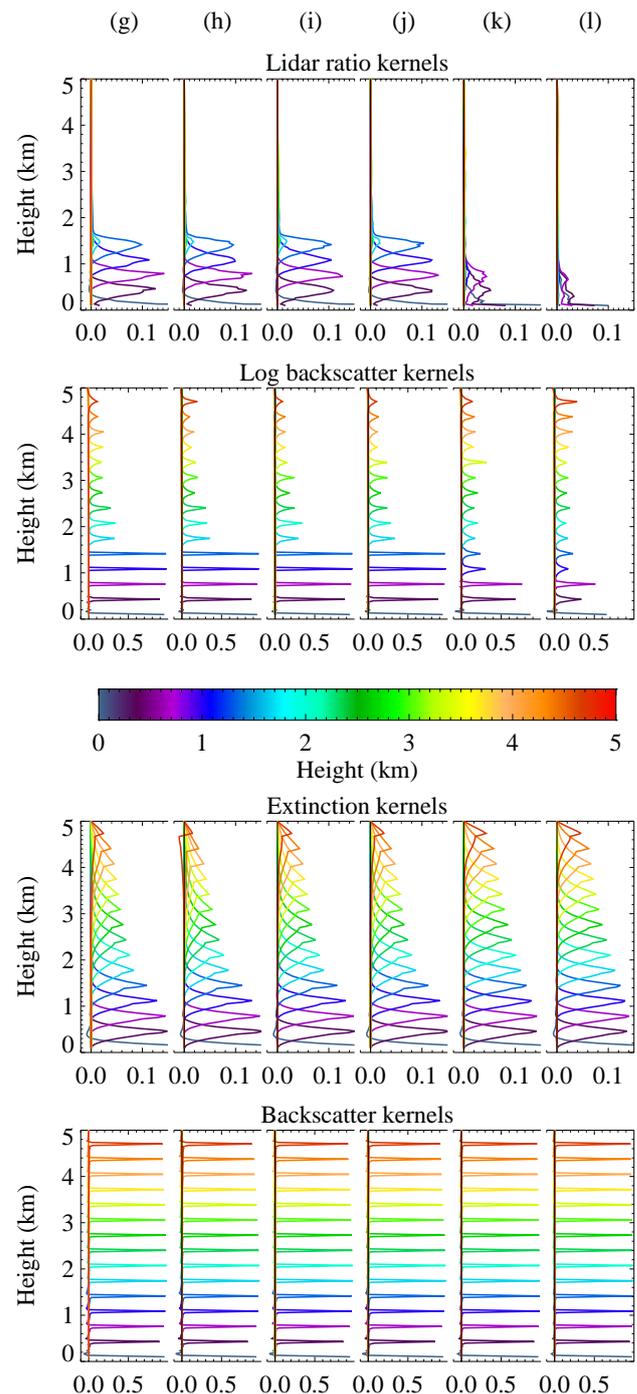
not resolved (see Fig. 6) as the features occupied only one retrieval bin. Doubling the resolution gives equivalent performance to cases g and h but with slightly increased noise and significantly increased processing time. Cases k and l are more difficult retrievals as they present lower signal-to-noise ratio (SNR). They are still consistent with the true profile but with greater errors.

Both configurations give a respectable fit to the extinction profile, though they do increasingly underestimate the magnitude of peaks as they become narrower. This decreased sensitivity is clear within the averaging kernels (Fig. 7). Though the backscatter kernels are virtually delta functions in the PBL, the extinction and lidar ratio kernels have widths of 300–1000 m (the effective resolution of those products, which increases with height). The kernels also illustrate the loss of sensitivity in the lidar ratio configuration above the PBL, with the magnitude of both kernels decreasing significantly. In cases k and l, the sensitivity is also lower due to the reduced SNR. The extinction configuration maintains sensitivity throughout the profile, though its extinction kernels are skewed about their centre (which may derive from  $E^{(ra)}$  measuring the integral of  $\alpha$ , such that bins beneath a level contribute greater information content). Overall, the kernels indicate that the smoother profiles returned by the lidar ratio mode are due to a greater reliance on the a priori.

### 3.2 Error analysis

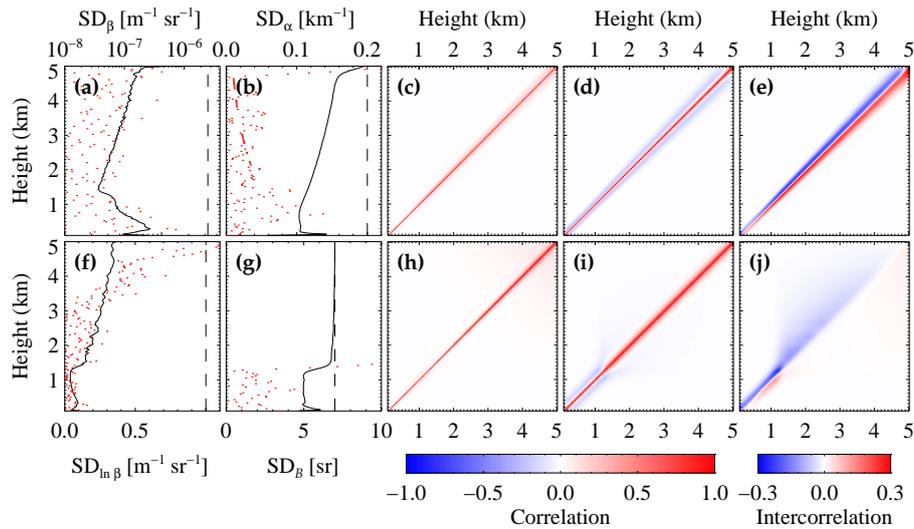
#### 3.2.1 Retrieval error

The error covariance matrices for case a in both configurations are shown in Fig. 8. They confirm that the linear configuration makes the best use of the available information as that mode behaves identically within and above the PBL, whilst the logarithmic configuration reverts to the a priori covariance in the free troposphere. Where there is information, the



**Fig. 7.** Selected rows of the averaging kernels for Fig. 5, denoting the relative contribution of the true state at each height (y axis) to the value retrieved at a height denoted by the colour. Above colour bar – lidar ratio configuration. Below – extinction configuration.

form of the covariance matrix has changed significantly from the a priori in both cases. Autocorrelation in the backscatter has decreased with regions near the surface being only 10 % correlated to adjacent bins. The extinction matrix, plots d and



**Fig. 8.** Covariance matrix for the retrieval of case a by the linear and log modes (top and bottom rows, respectively). (a) Backscatter standard deviation, being the diagonal of the covariance matrix. The absolute difference between the simulated profile and retrieval is shown in red. (b) Extinction standard deviation. (c) Backscatter autocorrelation. (d) Extinction autocorrelation. (e)  $\beta^{(p)}$  vs.  $\alpha^{(p)}$  error intercorrelation. (f) Log backscatter standard deviation. (g) Lidar ratio standard deviation. (h) Log backscatter autocorrelation. (i) Lidar ratio autocorrelation. (j)  $\ln \beta^{(p)}$  vs.  $B$  error intercorrelation.

i, show adjacent bins are correlated to  $\sim 60\%$ , whilst other nearby bins are slightly anticorrelated. The intercorrelation of the variables has also evolved, with points above a level being anticorrelated and those below positively correlated. It will need to be confirmed whether real data give similar results.

The diagonals of the covariances, plots a and f, can be used to approximate the error on the products. These give the bounds of Figs. 9 and 10. The lidar ratio a priori uncertainty is too small as it is not consistent with the simulated profile. Since that error is simply the a priori variance, this indicates that the a priori is overly constrictive. The extinction retrieval is better, though it underestimates the error in the PBL. Both fail to appreciate the error caused by assuming a non-paralysable dead-time correction when a different form (Donovan et al., 1993) was simulated.

The above-mentioned figures also compare the retrieval to the Ansmann method. For a fair comparison, the derivative is averaged over 300 m to be equivalent to the effective resolution of the retrieval. They are in good agreement in the PBL and the retrievals exhibit a lesser spread and error than the Ansmann solutions in the free troposphere. The Fernald–Klett method gives equivalent answers when given the correct lidar ratio.

### 3.2.2 Parameter error

In real retrievals, there will be some error in the model parameters  $\mathbf{b}$ . This additional uncertainty that can be included in the retrieval by extending the measurement uncertainty to

cover all sources of error,

$$\boldsymbol{\epsilon}_y = \boldsymbol{\epsilon} + \mathbf{K}_b(\mathbf{b} - \hat{\mathbf{b}}) + \Delta \mathbf{f}, \quad (16)$$

where  $\mathbf{K}_b = \partial F / \partial \mathbf{b}$ ,  $\hat{\mathbf{b}}$  is the best estimate of the true parameters  $\mathbf{b}$ , and the last term describes any inability of the forward model to describe the true state.

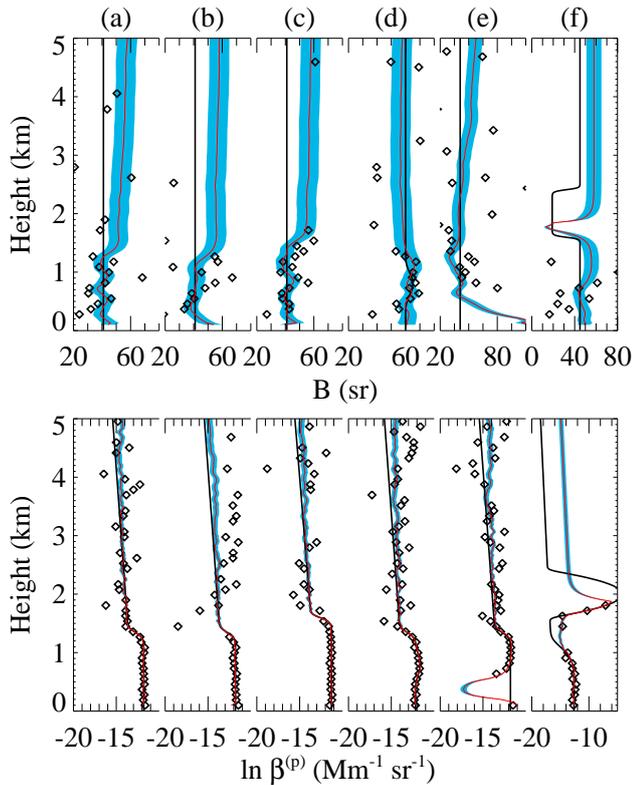
Concentrating on only the parameter error for the moment, this can be implemented by replacing all occurrences of  $\mathbf{S}_\epsilon$  with

$$\mathbf{S}_y = \mathbf{S}_\epsilon + \mathbf{K}_b \mathbf{S}_b \mathbf{K}_b^T. \quad (17)$$

This significantly increases the computing cost of the retrieval, as  $\mathbf{S}_y$  must now be inverted in each iteration. A reasonable approximation is to only re-evaluate  $\mathbf{S}_y$  after the last iteration. The full calculation is considered in this section but will be relaxed in Sect. 4 where the quantity of data increases.

The Ångström exponent can vary quite significantly but is commonly accepted to lie in the range 0.6–1.4, such that an error of 0.4 is reasonable (Klett, 1985). Radiosondes measure pressure and temperature at a given height with an accuracy of 0.5 hPa, 2 K, and 60 m, from which an error in number density and  $\mathcal{N}$  of 0.5% is expected (Kitchen, 1989). The height of the first observed bin,  $R_0$ , can be easily estimated to within 10 m. The standard deviation of the data used to estimate  $E_B$  can be easily derived.

The remaining parameters are estimated by some calibration procedure (e.g. Wandinger and Ansmann, 2002; Povey et al., 2012). For the purpose of demonstration, Fig. 11 shows the impact of each parameter on the total variance, assuming errors in  $C$  of 10% and  $\tau_d$  of 1 ns; these dominate the total.

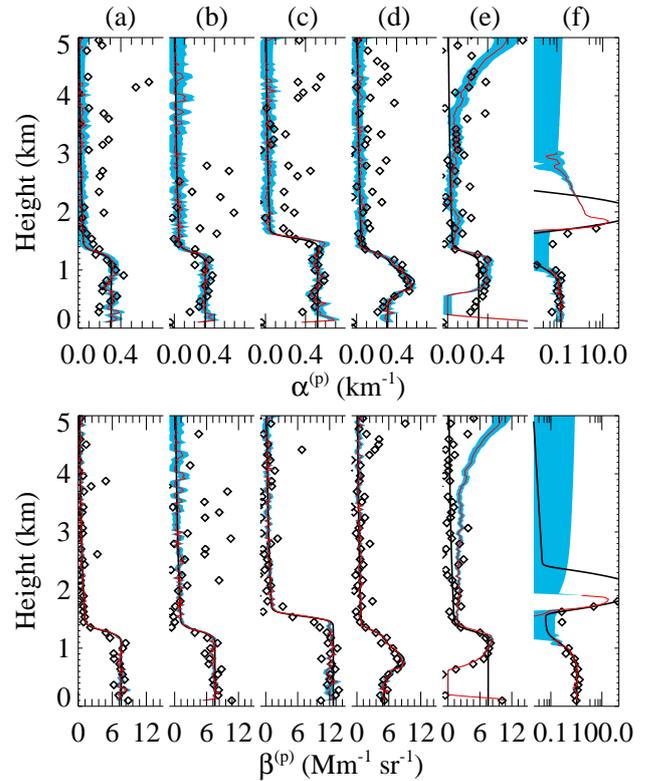


**Fig. 9.** Retrieval from the lidar ratio configuration (red) showing its error (blue) compared to that of the Ansmann method at 300 m resolution (diamonds) and the simulated profile (black).

For the impact of uncertainty in  $C$  to be of a similar order to the measurement error, it must be known to within 2% – an unrealistic expectation. However, these errors are unlikely to be correlated such that this term simply increases the total error.

The dead time is more troublesome for elastic measurements as it can introduce significant correlations within  $S_y$ . For the prototype system simulated with  $\tau_d = 50$  ns and a maximum count rate of 17 MHz, an error greater than 0.1 ns in its estimation significantly reduces the information content available and prevents the retrieval from converging. That is clearly an unrealistic expectation but is a fair representation of the impact that dead time has on the observations for this system. Most laboratory-standard systems will have much smaller dead times, which have a greater tolerance of around 1 ns.

The laser energy  $E_L$  behaves similarly to  $C$  but is considered separately as it can change significantly with time, whilst the calibration function should be fairly consistent. The laser energy may not always be accurately measured and so is retrieved as part of the state vector. As the a priori tends towards zero with height and the molecular component of the scattering is input into the retrieval as a parameter, the most favourable means for the algorithm to fit the signal at the top



**Fig. 10.** As Fig. 9 but for the extinction configuration.

of a profile is to vary  $E_L$ , effectively calibrating the signal against the Rayleigh scattering (as frequently done in Raman analyses). The process could be made more explicit by setting the a priori equal to zero with a small uncertainty at the top of the profile, but this has not been found necessary.

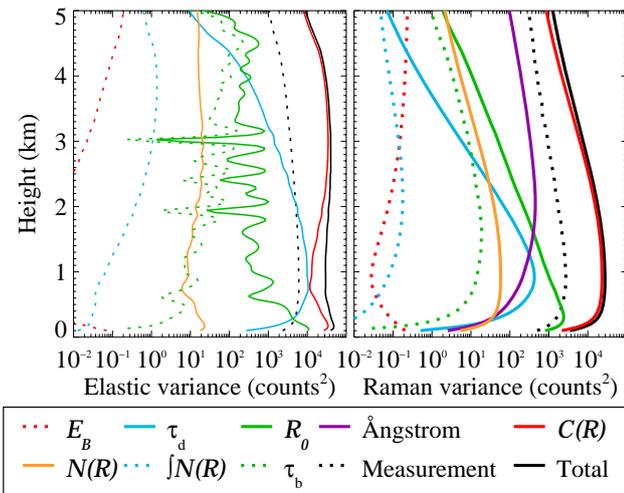
### 3.2.3 Further errors

The settings of the retrieval that have no bearing on the forward model should not affect  $\hat{x}$ . The initial value of  $\Gamma_i$  alters the number of iterations required to converge as it drives the size of each step in state space. A value of  $10^5$  appears to be optimal in most cases and  $\hat{x}$  appears to be independent of that choice provided it is not too large or small. Similarly, convergence thresholds of  $10^{-4}$  on change in cost or step and  $10^{-1}$  on error were selected as the highest order for which  $\hat{x}$  is not affected by the choice. The minimum retrieved height does have a small effect on the retrieval in its first few bins, so  $r_0 = 100$  m was chosen to concentrate these effects within a region where parameter errors will be large regardless.

Forward model error is defined in Sect. 3.2.3 of Rodgers (2000) as

$$\mathbf{G}_y[f(x, b, b') - F(x, b)], \quad (18)$$

where  $\mathbf{G}_y = \partial \hat{x} / \partial y$ , the sensitivity of the retrieved state to the measurement, and  $f$  is the exact, true profile including any processes the forward model  $F$  may not describe. This



**Fig. 11.** The contributions of model parameters to the total retrieved variance.

systematic error is generally difficult to estimate as, if  $f$  were known, it would most likely be used as the forward model instead.

There are some processes that are clearly not included within the current forward model. Multiple scattering has been neglected as it is mostly important for lidars with a wide footprint, such as space-based system, or for observations within clouds, where this algorithm is already known to perform poorly for other reasons. Though appropriate numerical models of multiple scattering exist (Eloranta, 1998), this is left as an area for future work if retrievals within clouds are desired.

There is a small difference between the bin-averaged backscatter that is sampled and the true backscatter defined by Mie theory, for which the error can be evaluated with Eq. (18). It is greatest in the entrainment layer (or at any other sharp gradient), being at most 1 % of the total error.

The models of the calibration function and detector nonlinearity are idealised versions of the truth. A rough estimate of these contributions can be produced by considering alternative models, such as that of Donovan et al. (1993). For case e, these are over 100 times larger than other errors. This is a circumstance that could benefit from the simultaneous processing of photon-counting and analogue data. The other cases are negligibly affected by the choice of nonlinear correction.

None of these errors describe the discrepancies shown in the free troposphere in Fig. 9 as that is dominated by the a priori uncertainty. In regions where the data are the dominant contribution to the retrieval (i.e. where the area of the averaging kernel is near unity), increasing the a priori variance does not affect the retrieved profiles. Where it is important, the error estimate should clearly be greater to better represent the uncertainty. Hence, the a priori uncertainty in  $B$  will be increased to 40 sr. This is effectively a uniform distribution in Fig. 2.

### 3.3 EARLINET intercomparison

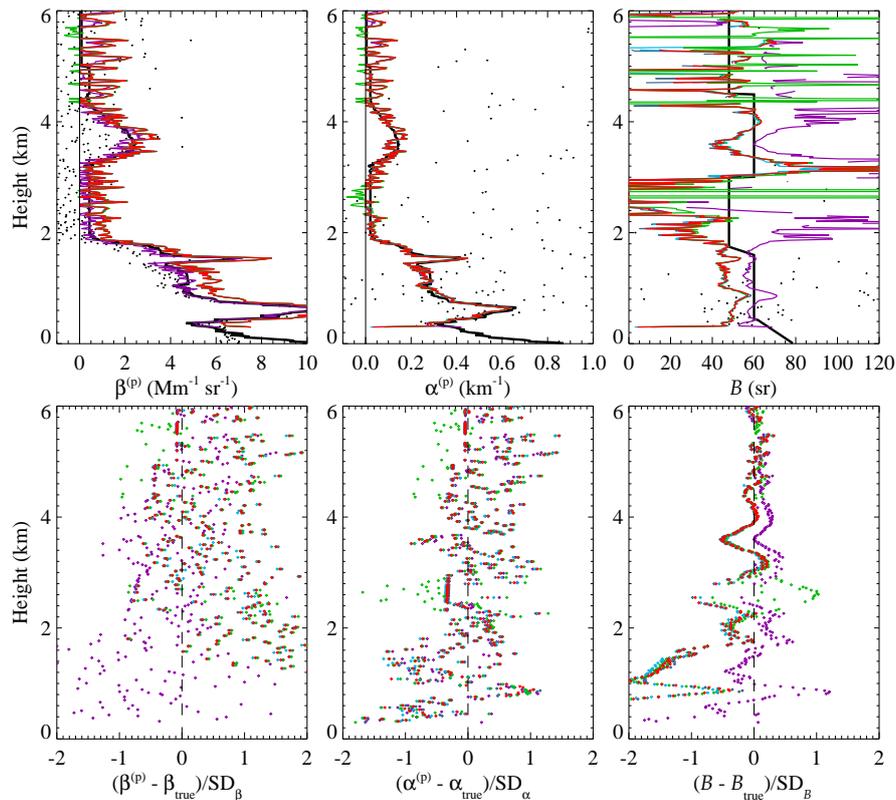
The European Aerosol Research Lidar Network (EARLINET) produced simulated data with which to compare the performance of the various algorithms used by its members (Pappalardo et al., 2004). These data have been processed with the proposed algorithm, shown in Fig. 12. The true profiles of backscatter, extinction, and the lidar ratio are shown over the top row in black, this retrieval in red, and the local formulation of the Ansmann method is shown with points. The bottom row plots the difference between the retrieved and synthetic profiles, normalised by the estimated error.

The extinction profile is consistent with that simulated throughout the retrieved range and performs equivalently to the EARLINET algorithms (Figs. 4 and 6 of Pappalardo et al., 2004). This retrieval fits the magnitude and shape of all three peaks more accurately than some of the EARLINET algorithms and does not indicate an erroneous peak above 5 km (though the magnitude of noise is similar).

The backscatter product is less satisfactory, with the retrieval overestimating the magnitude below 3 km. This is a matter of calibration. The EARLINET algorithms estimated the calibration constant for the elastic channel using a given value of the aerosol backscatter between 8 and 10 km (described as the Stage II comparison). As the extinction is derived from the gradient of the logarithm of the Raman channel, its calibration constant is not necessary. These are not ideal circumstances for the retrieval, as estimates of both constants are required. These were produced by additionally assuming  $\chi = 0.14$ , the most commonly observed value at the Chilbolton AERONET station (Woodhouse and Agnew, 2006–2011). The uncertainty of those estimates exceeds 60%. The retrieval of  $E_L$  found a suitable value of the Raman calibration constant (as the retrieval returns virtually identical results when both calibrations are multiplied by a random constant), but it has no mechanism to alter the ratio between the two calibrations resulting in the overestimation of the backscatter. The purple line of Fig. 12 shows a retrieval where the elastic calibration has been increased by 10%, giving a more accurate retrieval. This ratio should not change significantly over time and so would in practice be estimated from infrequent observation where  $\chi^{(p)}$  is known.

Two additional profiles are plotted in Fig. 12, where the a priori values (both  $x_a$  and  $S_a$ ) have been increased and decreased by 20% (grey and blue, respectively). The solutions cannot be visually distinguished from the original as the changes are less than  $10^{-7} \text{ Mm}^{-1} \text{ sr}^{-1}$  for backscatter and  $10^{-5} \text{ Mm}^{-1}$  for extinction, demonstrating that the proposed a priori is (as intended) a weak constraint on the solution. The differences can be resolved in the lower row of plots, showing the changes correspond to < 10% of the total error.

Figure 12 also includes an example of the retrieval without the requirement that all state vector elements be positive, shown in green. For this profile, the information content is sufficiently large that there is no difficulty in converging



**Fig. 12.** Performance of the retrieval (extinction mode) with data simulated by EARLINET. (Top) The simulated backscatter, extinction, and lidar ratio (black) compared to that from the Ansmann method (points) and retrieval (red). Also shown are the retrievals where the a priori is increased or decreased by 20 % (grey and blue), where the retrieval has been allowed to take negative values (green), and where the magnitude of the calibration function has been manually selected to improve the comparison. (Bottom) The difference between those retrievals and the simulated profile, normalised by the estimated standard deviation (Ansmann not shown).

(though this has been an issue elsewhere). Retrieved values increase slightly in magnitude throughout the profile, increasing its deviation from the simulated profile. This compensates for the negative values in the profile, such that the integral of the extinction profile differs by only 1 % from its value with the positive lower limit. The limit will therefore be retained.

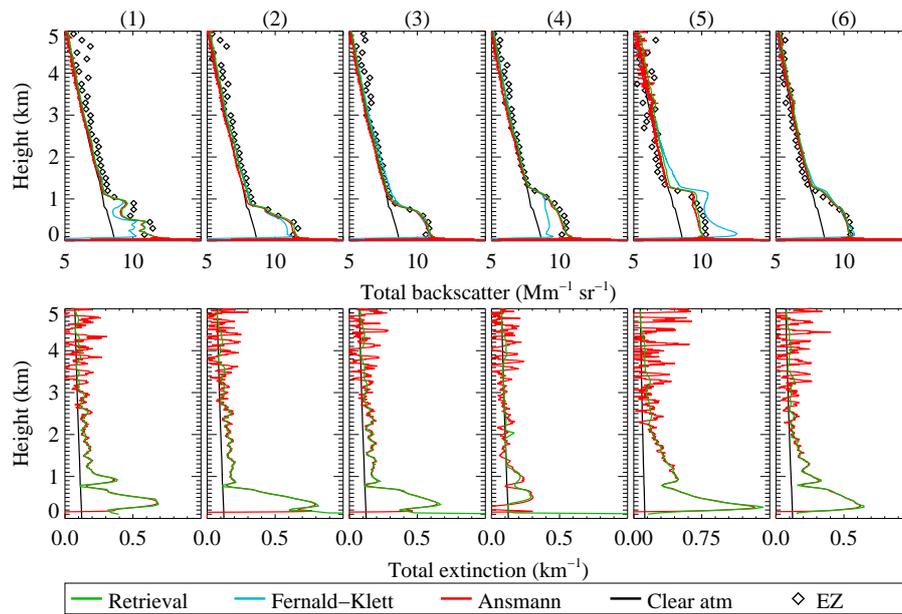
## 4 Application

### 4.1 Individual profiles

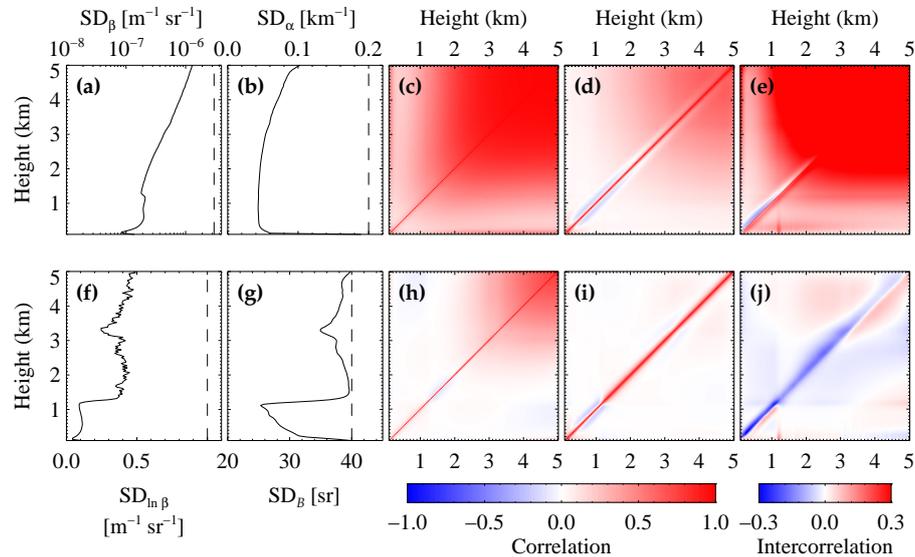
The retrieval is now applied to observations by the Chilbolton Ultraviolet Raman lidar (CUV; Agnew, 2003; Agnew and Wrench, 2006–2010), which is stationed at the Natural Environment Research Council (NERC) Chilbolton Facility for Atmospheric and Radio Research (CFARR; 51.1445° N, 1.4270° W; 84 m a.s.l.; STFC, 2006–2011). It uses a 355 nm Nd:YAG laser at 350 mJ and 50 Hz for water vapour profiling through the daytime boundary layer on a case-study basis, implementing both photon-counting and analogue data collection. Its observations can be directly compared to those

of a Leosphere EZ lidar operated continuously at the same site, which provides volume depolarisation ratio profiles instead of Raman observations. Radiosonde launches are available twice daily from Larkhill, 30 km northwest (UK Meteorological Office, 2006–2011).

Six profiles were selected from March 2010 for which the instrument's calibration has been thoroughly investigated using the techniques of Povey et al. (2012). Figure 13 compares the retrieved profiles to those given by the Fernald–Klett and Ansmann methods. A clear atmosphere is assumed between 4 and 5 km using a constant  $B$  to give an optical thickness consistent with sun photometer observations and the derivative is evaluated over 150 m. In the PBL, the retrieved backscatter is very similar to that given by the Ansmann ratio and an independent measurement by the EZ lidar. As the SNR decreases, the retrieval tends towards the Fernald–Klett solution. This is a proper response for the retrieval, giving answers similar to existing methods but tending from a two- to one-channel retrieval as the available information decreases. This is also expressed in the backscatter averaging kernels, which widen from 30 to 100 m.



**Fig. 13.** Various estimates of total backscatter (top) and two-way extinction (bottom) for six analogue profiles observed by the CUV during March 2010. The attenuated backscatter coefficient  $\beta \exp[-2(\chi + \sigma_R^{(L)} \mathcal{N})]$  reported by the EZ lidar is shown in diamonds for comparison to the total backscatter. The scattering that would be observed from a clear atmosphere is shown in black, highlighting negative  $\alpha$  returned by the Ansmann technique (which operated at 150 m resolution).



**Fig. 14.** As Fig. 8 but for real data from case (1) of Fig. 13. The large error correlations are a function of parameter error in  $E_B$ , which affects all levels equally and becomes dominant as the SNR decreases.

The retrieved extinction is consistent with the Ansmann solution but gives a much smoother solution. The averaging kernels confirm that there is little information available in the free troposphere but also show that the resolution in the PBL is better than that observed in simulations: 100 m. A tendency to find  $\alpha^{(p)} = 0$  at the top of the PBL in cases 1–4 is due to the number density profile. The radiosonde that morning

recorded a step decrease in temperature at the top of the PBL, but as that is a low-resolution measurement, linear interpolation overestimates  $N$  there. A standard atmosphere does no better. Due to factors such as these, the inclusion of parameter errors significantly reduces the information content in the free troposphere. This can occasionally produce unconstrained solutions (not shown) due to the relatively weak

a priori, but this does not significantly alter the result within the PBL. The errors, shown in Fig. 14, are similar to those of Fig. 8 but with the following differences:

- Larger background levels (being daytime observations) produce large correlations at the top of the profile. The two modes respond differently to this, with the logarithmic configuration reverting to the a priori covariance at the top of the PBL (as observed in simulations) and the linear configuration tending towards complete correlation (representing a more systematic error).
- The intercorrelation of extinction and backscatter is different. Bins above a point are still negatively correlated, but those below are more weakly correlated. The exact reason for this is not clear.

#### 4.2 Extended periods

Eleven hours of photon counting observations were processed from 2 March 2010. Figure 15 plots retrievals with an error less than 20 % or 30 sr ( $\beta^{(p)}$  and  $B$ , respectively). Row a shows the application of the Ansmann method, where the data were averaged over 30 m to give a similar resolution to the retrieval. Row b is the linear retrieval and c the logarithmic. The three backscatter fields are qualitatively similar before 14:00 GMT and after 18:00 GMT. Between these times, the measurement of laser energy has diverged increasingly from reality. As the retrieval has no knowledge of that, it retrieves smaller backscatter to compensate. The Ansmann method is not affected as it considers a ratio of channels. The difference between the Ansmann solution and the retrieval is effectively a constant factor of the failure in the calibration, with the results otherwise being consistent. For example, both methods observe larger backscatter in updrafts than downdrafts (where vertical wind was observed by a Doppler lidar).

The retrievals are consistent in their estimates of the lidar ratio and are no worse than the Ansmann method, which is greatly affected by overlap when estimating extinction. The aerosol layer near 1 km at 10:00 GMT gives a lidar ratio of around 30 sr. This is a residual layer lying above a developing mixed layer where lidar ratios are larger (around 50 sr). The low lidar ratio indicates large, likely spherical, particles which are reasonable for an aged residual layer. The results are better in the evening, observing a peak  $B$  of 70 sr over a background of 50 sr, indicating the appearance of smaller particles. By this time, convective mixing has collapsed into a persistent updraft, so the increase in depolarisation ratio could indicate that newer, non-spherical particles are being lofted from the surface or are advected over the site. Advected aerosol is more likely considering the brevity of the peak.

Figure 16 compares the retrieved  $\chi_{\infty}$  during that day to AERONET measurements (Woodhouse and Agnew, 2006–

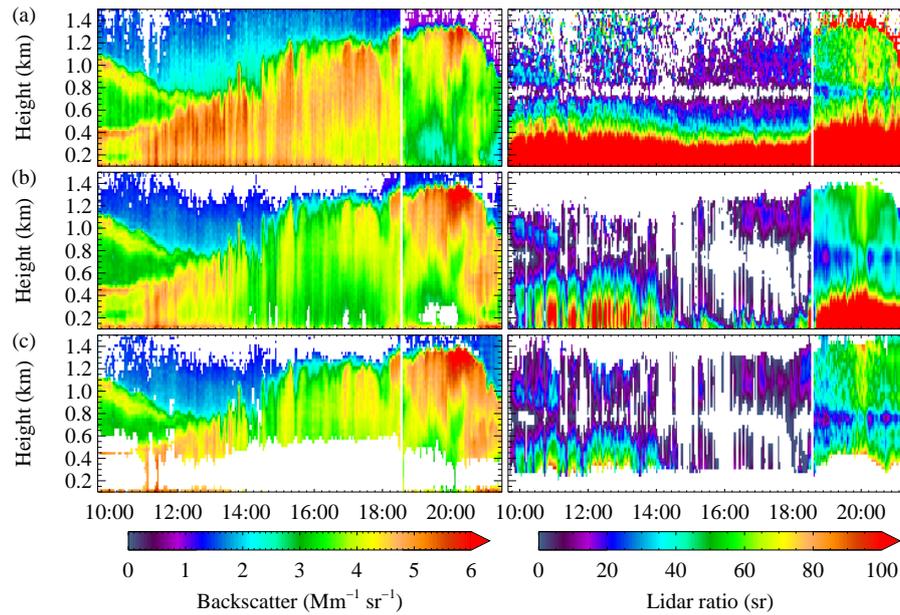
2011). Their agreement is reasonable if not impressive. The retrieval tends to return larger  $\chi$  than observed by AERONET, though it also contains substantial variability that the latter does not. This is likely due to the inaccurate measurement of laser energy, though this is under investigation. A calibration performed at 10:00 GMT was used throughout this day and that is the only AERONET measurement that was in any way input into the retrieval – the remainder are independent. Regardless, the retrieved values are equivalent to those given by the Ansmann algorithm, indicating that the retrieval is correct for the parameters it has been given. It is the calibration of the system, not the method of retrieval, producing the poor comparison.

#### 4.3 Eyjafjallajökull ash

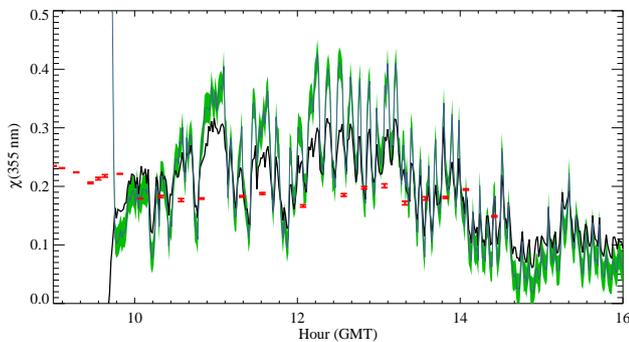
The eruptions of the Eyjafjallajökull volcano in southern Iceland during April and May of 2010 produced the single most significant volcanic ash event over northern Europe in the age of aviation. The closure of airspace cancelled around 100 000 flights, inconveniencing millions of travellers across the globe and resulting in massive losses for airlines and related industries. Owing to the density of personnel and instrumentation within the reach of this plume, it has become one of the most studied atmospheric events in history. The introduction of Johnson et al. (2012) provides a reasonable overview of the literature published to date and more will certainly be published over the years to come.

The CUV was operated, in addition to routine measurements, on 19 April to observe ash within the boundary layer. The data suffer similar difficulties to those already discussed, such as the impact of the calibration function being clear below 600 m. With these limitations in mind, the optimal estimation retrieval (in extinction mode) was applied to these observations, shown for analogue data in Figs. 17 and 18. Plot a presents the volume depolarisation ratio observed by the EZ lidar, while plots c and d show the retrieved backscatter and lidar ratio with errors outlined in Fig. 18. Ash particles have a large depolarisation due to their asphericity, and these measurements indicate the presence of a 400 m thick ash layer within the PBL, highlighted in Figs. 17c and d by plotting the 0.07 depolarisation isoline. It exhibits a low backscatter ( $< 10 \text{ Mm}^{-1} \text{ sr}^{-1}$ ) and lidar ratio (20–35 sr) compared to the remainder of the scene. These are well outside the range of 50–82 sr reported in the literature for similar ash in the free troposphere (Ansmann et al., 2010; Marengo and Hogan, 2011; Hervo et al., 2012), though slightly larger than found if the data is analysed with the Ansmann algorithm (10–25 sr). The retrieval indicates that the properties of the ash have changed significantly after 12–24 h within the PBL. The decreased  $B$  implies a growth or shape change of the particles, possibly due to sulfate coating.

A mixed layer forms beneath the aerosol (see the vertical velocity in plot b). There,  $B = 50\text{--}80$  sr with minimal depolarisation, which is broadly consistent with urban



**Fig. 15.** The backscatter (left) and lidar ratio (right) retrieved from photon-counting CUV observations on 2 March 2010. Results with an error greater than 20 % of  $\beta^{(p)}$  or 30 sr are plotted in white. (a) The Ansmann method after averaging the data over 30 m to give a similar resolution to the retrieval. (b) Retrieval of  $\beta^{(p)}$  and  $\alpha^{(p)}$ . (c) Retrieval of  $\ln\beta^{(p)}$  and  $B$ . The feature at 800 m in all lidar ratios is due to an inaccurate estimate of  $N$ .



**Fig. 16.** Retrieved aerosol optical thickness at 355 nm (grey/green) compared to that observed by AERONET at level 2.0 (red) for 2 March 2010. The analogue Ansmann solution integrated between 0.25 and 2 km is shown in black.

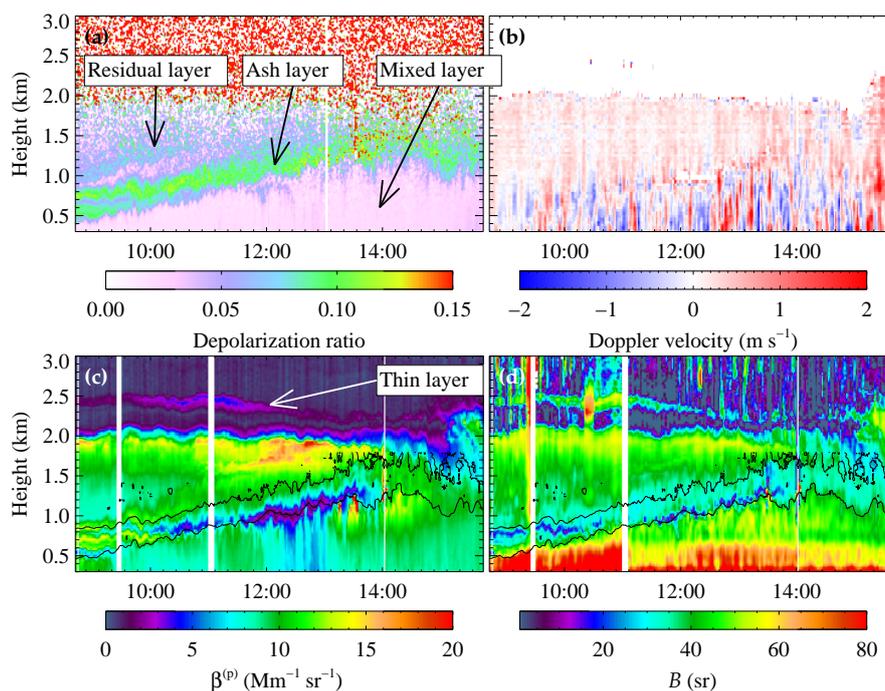
aerosols (Müller et al., 2007). Backscatter is fairly homogeneous throughout the layer except during a period of updrafts around 13:00 GMT when  $\beta^{(p)}$  decreases from 10 to  $6 \text{ Mm}^{-1} \text{ sr}^{-1}$ . The absence of a similar change elsewhere in the PBL gives some confidence that this is a real variation rather than a calibration artefact.

A more weakly depolarising aerosol resides in a poorly mixed residual layer above the ash layer. It persists until 14:00 GMT, when they mix. Backscatter and lidar ratios are large at the top of this layer and increase with height. This could be simple stratification within a poorly mixed layer

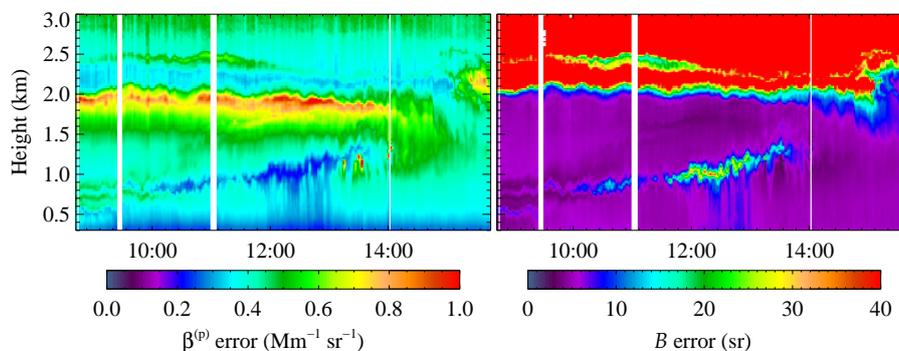
or smaller particles may have concentrated at the top of the layer whilst larger particles have begun to settle.

Finally, a thin layer of aerosol is present above the PBL (labelled in plot c). The EZ lidar did not resolve this, so no measure of the depolarisation is available. Expressing lidar ratios of 40–60 sr with low backscatter, the layer is consistent with aerosol typically observed at CFARR and there is no reason to label it as ash.

Figure 19 presents the distribution of retrieved extinction and backscatter for all points for which the error in the volume depolarisation ratio is  $< 100\%$ . Lines of constant  $B$  are added for reference. Points likely to contain ash are shown in the left plot by filtering for depolarisations greater than 0.07. The residual layer appears as a concentration of points around  $\beta^{(p)} = 10 \text{ Mm}^{-1} \text{ sr}^{-1}$  and  $B \simeq 40$  sr. The mixed layer appears in the right plot as a more continuous distribution between  $B = 40$  and 60 sr. The failure of the retrieval near the surface is evident in a vertical line of points at  $\beta^{(p)} = 8 \text{ Mm}^{-1} \text{ sr}^{-1}$ . In the free troposphere,  $\beta^{(p)} < 1 \text{ Mm}^{-1} \text{ sr}^{-1}$ , where poorly constrained retrievals produce a broad distribution in both plots. There are very few observations of the thin ash layer, but their presence is evident in observations near  $B \simeq 20$  sr in the left plot not expressed on the right.



**Fig. 17.** Observations of the Eyjafjallajökull ash plume at CFARR on 19 April 2010. **(a)** Volume depolarisation ratio observed by the EZ lidar. Values above 1.9 km are dominated by noise. **(b)** Vertical velocity observed by a Halo Doppler lidar. **(c)** Backscatter retrieved (in the linear mode) from analogue CUV measurements. The 0.07 isoline of volume depolarisation below 1.8 km is shown to highlight the ash layer. **(d)** Lidar ratio retrieved from same.

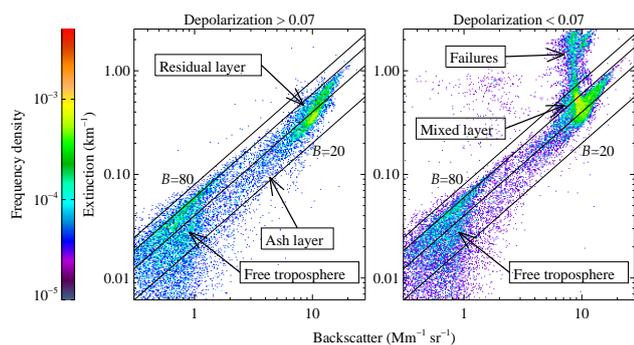


**Fig. 18.** Retrieved errors for plots c and d of Fig. 17.

## 5 Conclusions

An optimal estimation retrieval scheme for aerosol scattering properties from Raman lidar observations was proposed, using the lidar equations as a simple forward model. The a priori state and covariance matrix were based on the properties of aerosol outlined in the OPAC model to provide a weak constraint on the magnitude of scattering whilst assuming them to be vertically correlated over a scale height of 100 m. This is smaller than observed by balloon-borne measurements but ensures that the PBL and free troposphere are not coupled.

The state of the atmosphere can be described at each height by the aerosol backscatter and either the extinction or lidar ratio. These possibilities were assessed by considering their ability to process simulated data. The lidar ratio configuration was found to lose sensitivity in the free troposphere, relying excessively upon its a priori assumptions, as shown by the disappearance of the averaging kernel. If extinction is retrieved instead, it and backscatter should be retrieved linearly with a correlation assumed between them (95 % here, though more investigation of this value is necessary). This configuration maintains sensitivity throughout the profile. This choice is likely influenced by the desire to use a single a priori appropriate for all observations. In future, if aerosol type



**Fig. 19.** Distribution of extinction and backscatter for 19 April 2010. Lines delineate lidar ratios of 20, 40, 60, and 80 sr. (Left) Points observed to have a volume depolarisation ratio  $> 0.07$ , which likely contain some quantity of ash. (Right) The remaining points, corresponding to typical PBL aerosols and failed retrievals near the surface.

information was used to better estimate the lidar ratio before analysis, the lidar ratio configuration may perform more favourably.

In the analysis of simulated and real data, the proposed retrieval is consistent with existing analyses. Backscatter was always retrieved at the finest resolution allowed (mostly 33 m, but this remains true at the instrumental limit  $< 10$  m) and with an uncertainty between 2% in the most ideal circumstances and 20% in the least. Extinction and the lidar ratio are less well constrained, expressing resolutions of 300–500 m in simulations and 0.1–1 km with real data. Importantly, these are different from the scale of vertical correlations assumed a priori and increase as SNR decreases. The retrieval has selected the most suitable resolution independently, unlike the smoothing filters used in most studies. The uncertainty in extinction retrieved from real data is relatively large ( $> 15\%$ ) but that is in part due to the short timescales evaluated (1 min in Figs. 15 and 17). The integration time can be increased to reduce these errors to any desired level (at least until atmospheric variability begins to dominate). Regardless, errors are comparable to, if not smaller than, fair estimates of the error resulting from the standard Raman lidar technique of Ansmann et al. (1992) applied to the same data.

The magnitude of the uncertainty was shown to be dominated by the calibration of the instrument – primarily the offset of its vertical axis, the nonlinear response of its detectors, and the calibration function. Ideally, all of these would be determined with dedicated laboratory measurements, especially the offset and nonlinearity, which should change very little over time.

Optimal estimation is a poor technique for exploring unknown and unusual circumstances (such as dense smoke or desert dust plumes) as it can only retrieve states that are consistent with the a priori. (This is generally taken to mean that the retrieved state must be within one standard deviation of

the a priori, but it is more accurately considered a balance of the a priori and measurement uncertainties. A state far from the a priori can be returned if the uncertainty in its measurement is sufficiently small.) However, the retrieval will return a cost that indicates when a poor fit of the a priori is encountered. If observations are required there, the magnitude constraint can be altered or practically removed (by setting  $S_a$  to a very large value), analogous to satellite aerosol retrievals using several aerosol models from which only the result with the lowest cost is reported. It should be emphasised that the proposed a priori is a weak, order-of-magnitude constraint and there is no reason to suspect the retrieval would perform poorly except in the presence of unusually bright or dense aerosols.

The retrieval was applied to several hours of observation on 19 April 2010 of ash from the Eyjafjallajökull eruption. A depolarising ash layer was observed with a lidar ratio of 20–30 sr, much lower than observed in the free troposphere by previous studies and potentially indicating a growth of the particles after 12–24 h within the planetary boundary layer. More dispersed ash within a residual layer exhibited a backscatter of  $10 \text{ Mm}^{-1} \text{ sr}^{-1}$  and lidar ratio of 40 sr.

*Acknowledgements.* This work was supported by a NERC studentship in association with Hovemere Ltd. The synthetic EARLINET data were provided with the assistance of Valery Shcherbakov, Gelsomina Pappalardo, and Aldo Amodeo. Thanks to the two anonymous reviewers for their comments and recommendations that improved the discussions of this paper.

Edited by: G. Pappalardo

## References

- Agnew, J. L.: Lidar and radar tropospheric profiling at Chilbolton Observatory, in: Sixth International Symposium on Tropospheric Profiling: Needs and Technologies, 151–153, Leipzig, Germany, 2003.
- Agnew, J. and Wrench, C.: Chilbolton UV Raman lidar raw data, STFC Chilbolton Observatory, Rutherford Appleton Laboratory, 2006–2010.
- Althausen, D., Müller, D., Ansmann, A., Wandinger, U., Hube, H., Clauer, E., and Zörner, S.: Scanning 6-wavelength 11-channel aerosol lidar, *J. Atmos. Ocean. Tech.*, 17, 1469–1482, doi:10.1175/1520-0426(2000)017<1469:swcal>2.0.co;2, 2000.
- Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W.: Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar, *Appl. Optics*, 31, 7113–7131, doi:10.1364/AO.31.007113, 1992.
- Ansmann, A., Tesche, M., Gross, S., Freudenthaler, V., Seifert, P., Hiesch, A., Schmidt, J., Wandinger, U., Mattis, I., Müller, D., and Wiegner, M.: The 16 April 2010 major volcanic ash plume over central Europe: EARLINET lidar and AERONET photometer observations at Leipzig and Munich, Germany, *Geophys. Res. Lett.*, 37, L13810, doi:10.1029/2010gl043809, 2010.

- Colbeck, I.: Physical and Chemical Properties of Aerosols, Blackie Academic and Professional, London, UK, 1st Edn., 1998.
- Dacre, H. F., Grant, A. L. M., Hogan, R. J., Belcher, S. E., Thomson, D. J., Devenish, B. J., Marengo, F., Hort, M. C., Haywood, J. M., Ansmann, A., Mattis, I., and Clarisse, L.: Evaluating the structure and magnitude of the ash plume during the initial phase of the 2010 Eyjafjallajökull eruption using lidar observations and NAME simulations, *J. Geophys. Res.-Atmos.*, 116, D00U03, doi:10.1029/2011jd015608, 2011.
- Delanoe, J. and Hogan, R. J.: A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer, *J. Geophys. Res.-Atmos.*, 113, D07204, doi:10.1029/2007jd009000, 2008.
- Di Girolamo, P., Ambrico, P. F., Amodeo, A., Boselli, A., Pappalardo, G., and Spinelli, N.: Aerosol observations by lidar in the nocturnal boundary layer, *Appl. Optics*, 38, 4585–4595, doi:10.1364/AO.38.004585, 1999.
- Donovan, D. P., Whiteway, J. A., and Carswell, A. I.: Correction for nonlinear photon-counting effects in lidar systems, *Appl. Optics*, 32, 6742–6753, doi:10.1364/AO.32.006742, 1993.
- Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F., Volten, H., Munoz, O., Veihelmann, B., van der Zande, W. J., Leon, J. F., Sorokin, M., and Slutsker, I.: Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust, *J. Geophys. Res.-Atmos.*, 111, D11208, doi:10.1029/2005jd006619, 2006.
- Eloranta, E. W.: Practical model for the calculation of multiply scattered lidar returns, *Appl. Optics*, 37, 2464–2472, doi:10.1364/AO.37.002464, 1998.
- Ewell, D. M., Flocchini, R. G., Myrup, L. O., and Cahill, T. A.: Aerosol transport in the southern Sierra-Nevada, *J. Appl. Meteorol.*, 28, 112–125, doi:10.1175/1520-0450(1989)028<0112:atitss>2.0.co;2, 1989.
- Fernald, F. G.: Analysis of atmospheric lidar observations – Some comments, *Appl. Optics*, 23, 652–653, doi:10.1364/AO.23.000652, 1984.
- Ferrero, L., Perrone, M. G., Petraccone, S., Sangiorgi, G., Ferrini, B. S., Lo Porto, C., Lazzati, Z., Cocchi, D., Bruno, F., Greco, F., Riccio, A., and Bolzacchini, E.: Vertically-resolved particle size distribution within and above the mixing layer over the Milan metropolitan area, *Atmos. Chem. Phys.*, 10, 3915–3932, doi:10.5194/acp-10-3915-2010, 2010.
- Ferrero, L., Mocnik, G., Ferrini, B. S., Perrone, M. G., Sangiorgi, G., and Bolzacchini, E.: Vertical profiles of aerosol absorption coefficient from micro-Aethalometer data and {Mie calculation over Milan}, *Sci. Total Environ.*, 409, 2824–2837, doi:10.1016/j.scitotenv.2011.04.022, 2011.
- Fugii, T. and Fukuchi, T.: *Laser Remote Sensing*, Taylor and Francis, Boca Raton, FL, 2005.
- Grainger, R. G., Lucas, J., Thomas, G. E., and Ewen, G. B. L.: Calculation of Mie derivatives, *Appl. Optics*, 43, 5386–5393, doi:10.1364/ao.43.005386, 2004.
- Grant, J., Grainger, R. G., Lawrence, B. N., Fraser, G. J., von Biel, H. A., Heuff, D. N., and Plank, G. E.: Retrieval of mesospheric electron densities using an optimal estimation inverse method, *J. Atmos. Sol.-Terr. Phys.*, 66, 381–392, doi:10.1016/j.jastp.2003.12.006, 2004.
- Greenberg, J. R., Guenther, A. B., and Turnipseed, A.: Tethered balloon-based soundings of ozone, aerosols, and solar radiation near {Mexico City during MIRAGE-MEX}, *Atmos. Environ.*, 43, 2672–2677, doi:10.1016/j.atmosenv.2009.02.019, 2009.
- Guldner, J. and Spankuch, D.: Remote sensing of the thermodynamic state of the atmospheric boundary layer by ground-based microwave radiometry, *J. Atmos. Ocean. Tech.*, 18, 925–933, doi:10.1175/1520-0426(2001)018<0925:rsotts>2.0.co;2, 2001.
- Haywood, J. M. and Shine, K. P.: The effect of anthropogenic sulfate and soot aerosol on the clear-sky planetary radiation budget, *Geophys. Res. Lett.*, 22, 603–606, doi:10.1029/95gl00075, 1995.
- Hervo, M., Quennehen, B., Kristiansen, N. I., Boulon, J., Stohl, A., Fréville, P., Pichon, J.-M., Picard, D., Labazuy, P., Gouhier, M., Roger, J.-C., Colomb, A., Schwarzenboeck, A., and Sellegri, K.: Physical and optical properties of 2010 Eyjafjallajökull volcanic eruption aerosol: ground-based, Lidar and airborne measurements in France, *Atmos. Chem. Phys.*, 12, 1721–1736, doi:10.5194/acp-12-1721-2012, 2012.
- Hess, M., Koepke, P., and Schult, I.: Optical properties of aerosols and clouds: The software package OPAC, *B. Am. Meteorol. Soc.*, 79, 831–844, doi:10.1175/1520-0477(1998)079<0831:opoaac>2.0.co;2, 1998.
- Huang, Z. W., Huang, J. P., Bi, J. R., Wang, G. Y., Wang, W. C., Fu, Q. A., Li, Z. Q., Tsay, S. C., and Shi, J. S.: Dust aerosol vertical structure measurements using three MPL lidars during 2008 China-US joint dust field experiment, *J. Geophys. Res.-Atmos.*, 115, D00K15, doi:10.1029/2009jd013273, 2010.
- IPCC (Intergovernmental Panel on Climate Change): *Climate Change 2007: The Physical Basis of Climate Change, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, New York, 2007.
- Johnson, B., Turnbull, K., Brown, P., Burgess, R., Dorsey, J., Baran, A. J., Webster, H., Haywood, J., Cotton, R., Ulanowski, Z., Hesse, E., Woolley, A., and Rosenberg, P.: In situ observations of volcanic ash clouds from the FAAM aircraft during the eruption of Eyjafjallajökull in 2010, *J. Geophys. Res.-Atmos.*, 117, D00U24, doi:10.1029/2011jd016760, 2012.
- Kitchen, M.: Representativeness errors for radiosonde observations, *Q. J. Roy. Meteorol. Soc.*, 115, 673–700, doi:10.1256/smsqj.48712, 1989.
- Klett, J. D.: Stable analytical inversion solution for processing lidar returns, *Appl. Optics*, 20, 211–220, doi:10.1364/AO.20.000211, 1981.
- Klett, J. D.: Lidar inversion with variable backscatter extinction ratios, *Appl. Optics*, 24, 1638–1643, doi:10.1364/AO.24.001638, 1985.
- Li, W., Stamnes, K., Spurr, R., and Stamnes, J.: Simultaneous retrieval of aerosol and ocean properties by optimal estimation: {SeaWiFS case studies for the Santa Barbara Channel}, *International J. Remote Sens.*, 29, 5689–5698, doi:10.1080/01431160802007632, 2008.
- Lohmann, U. and Feichter, J.: Global indirect aerosol effects: a review, *Atmos. Chem. Phys.*, 5, 715–737, doi:10.5194/acp-5-715-2005, 2005.
- Lopatín, A., Dubovik, O., Chaikovskiy, A., Goloub, P., Lapyonok, T., Tanré, D., and Litvinov, P.: Enhancement of aerosol characterization using synergy of lidar and sun-photometer coincident

- observations: the GARRLiC algorithm, *Atmos. Meas. Tech.*, 6, 2065–2088, doi:10.5194/amt-6-2065-2013, 2013.
- Marchant, C. C., Moon, T. K., and Gunther, J. H.: An Iterative Least Square Approach to Elastic-Lidar Retrievals for Well-Characterized Aerosols, *IEEE T. Geosci. Remote*, 48, 2430–2444, doi:10.1109/tgrs.2009.2038903, 2010.
- Marchant, C. C., Wojcik, M. D., and Bradford, W. J.: Estimation of Aerosol Effective Radius by Multiwavelength Elastic Lidar, *IEEE T. Geosci. Remote Sens.*, 50, 645–660, doi:10.1109/tgrs.2011.2160725, 2012.
- Marengo, F. and Hogan, R. J.: Determining the contribution of volcanic ash and boundary layer aerosol in backscatter lidar returns: A three-component atmosphere approach, *J. Geophys. Res.-Atmos.*, 116, D00U06, doi:10.1029/2010jd015415, 2011.
- Marks, C. J. and Rodgers, C. D.: A retrieval method for atmospheric composition from limb emission measurements, *J. Geophys. Res.-Atmos.*, 98, 14939–14953, doi:10.1029/93jd01195, 1993.
- Measures, R. M.: *Lidar Remote Sensing: Fundamentals and Applications*, Krieger Publishing Company, Malabar, Florida, 2nd Edn., 1992.
- Müller, D., Wandinger, U., and Ansmann, A.: Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: theory, *Appl. Optics*, 38, 2358–2368, doi:10.1364/AO.38.002358, 1999.
- Müller, D., Ansmann, A., Mattis, I., Tesche, M., Wandinger, U., Althausen, D., and Pisani, G.: Aerosol-type-dependent lidar ratios observed with Raman lidar, *J. Geophys. Res.-Atmos.*, 112, D16202, doi:10.1029/2006jd008292, 2007.
- Müller, J. W.: Dead-time problems, *Nucl. Instrum. Methods*, 112, 47–57, doi:10.1016/0029-554x(73)90773-8, 1973.
- NDACC (Network for the Detection of Atmospheric Composition Change) [J. Rosen]: Backscattersonde data (Laramie, WY; Lauder, New Zealand; Thule, Greenland), Supported by NSF grants OPP-9423285, ATM-9500186, 1989–2000.
- Newsom, R. K., Turner, D. D., Mielke, B., Clayton, M., Ferrare, R., and Sivaraman, C.: Simultaneous analog and photon counting detection for Raman lidar, *Appl. Optics*, 48, 3903–3914, doi:10.1364/AO.48.003903, 2009.
- NOAA (National Oceanic and Atmospheric Administration): U.S. Standard Atmosphere, Tech. Rep. NOAA Doc. S/T 76-1562, U.S. Government Printing Office, Washington, D.C., 1976.
- Oke, T. R.: *Boundary Layer Climates*, Cambridge University Press, London, UK, 2nd Edn., 1987.
- Pappalardo, G., Amodeo, A., Pandolfi, M., Wandinger, U., Ansmann, A., Bösenberg, J., Matthias, V., Amirdis, V., De Tomasi, F., Frioud, M., Iarlori, M., Komguem, L., Papayannis, A., Rocadenbosch, F., and Wang, X.: Aerosol lidar intercomparison in the framework of the EARLINET project. 3. Raman lidar algorithm for aerosol extinction, backscatter, and lidar ratio, *Appl. Optics*, 43, 5370–5385, doi:10.1364/AO.43.005370, 2004.
- Pappalardo, G., Amodeo, A., Mona, L., and Pandolfi, M.: Systematic measurements of the aerosol extinction-to-backscatter ratio, in: *Lidar Remote Sensing for Industry and Environmental Monitoring V*, edited by: Singh, U. N. and Mizutani, K., Vol. 5653, 77–87, SPIE, doi:10.1117/12.578809, 2005.
- Pornsawad, P., D’Amico, G., Böckmann, C., Amodeo, A., and Pappalardo, G.: Retrieval of aerosol extinction coefficient profiles from Raman lidar data by inversion method, *Applied Optics*, 51, 2035–2044, doi:10.1364/AO.51.002035, 2012.
- Pounder, N. L., Hogan, R. J., Várnai, T., Battaglia, A., and Cahalan, R. F.: A Variational Method to Retrieve the Extinction Profile in Liquid Clouds Using Multiple-Field-of-View Lidar, *J. Appl. Meteorol. Climatol.*, 51, 350–365, doi:10.1175/jamc-d-10-05007.1, 2012.
- Povey, A.: The application of optimal estimation retrieval to lidar observations, Ph.D. thesis, University of Oxford, Oxford, UK, available at: <http://www.atm.ox.ac.uk/group/eodg/theses/Povey.pdf> (last access: 12 March 2014), 2013.
- Povey, A. C., Grainger, R. G., Peters, D. M., Agnew, J. L., and Rees, D.: Estimation of a lidar’s overlap function and its calibration by nonlinear regression, *Appl. Optics*, 51, 5130–5143, doi:10.1364/AO.51.005130, 2012.
- Press, W. H., Vetterling, W. T., Teukolsky, S. A., and Flannery, B. P.: *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New Delhi, India, second edn., 1992.
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding: Theory and Practice*, vol. 2, World Scientific, Singapore, second edn., 2000.
- Rosen, J. M. and Kjöme, N. T.: Backscattersonde — A new instrument for atmospheric aerosol research, *Appl. Optics*, 30, 1552–1561, doi:10.1364/AO.30.001552, 1991.
- Rosen, J., Young, S., Laby, J., Kjöme, N., and Gras, J.: Springtime aerosol layers in the free troposphere over Australia: Mildura Aerosol Tropospheric Experiment (MATE 98), *J. Geophys. Res.-Atmos.*, 105, 17833–17842, doi:10.1029/2000jd900208, 2000.
- Shcherbakov, V.: Regularized algorithm for Raman lidar data processing, *Appl. Optics*, 46, 4879–4889, doi:10.1364/AO.46.004879, 2007.
- Steyn, D. G., Baldi, M., and Hoff, R. M.: The detection of mixed layer depth and entrainment zone thickness from lidar backscatter profiles, *J. Atmos. Ocean.-Tech.*, 16, 953–959, doi:10.1175/1520-0426(1999)016<0953:TDOMLD>2.0.CO;2, 1999.
- STFC (Science and Technology Facilities Council) [C. L. Wrench]: Chilbolton Facility for Atmospheric and Radio Research (CFARR) data, NCAS British Atmospheric Data Centre, available at: [http://badc.nerc.ac.uk/view/badc.nerc.ac.uk\\_\\_ATOM\\_\\_dataent\\_chobs](http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_chobs) (last access: 9 November 2010), 2006–2011.
- Sugimoto, N., Matsui, I., Shimizu, A., Nishizawa, T., Hara, Y., Chenbo, X., Uno, I., Yumimoto, K., Zifa, W., and Soon-Chang, Y.: Lidar network observations of tropospheric aerosols, *Proceedings of the SPIE*, 7153, 71530A, doi:10.1117/12.806540, 2008.
- UK Meteorological Office [G. Parton]: Met Office Global Radiosonde Data, NCAS British Atmospheric Data Centre, available at: [http://badc.nerc.ac.uk/view/badc.nerc.ac.uk\\_\\_ATOM\\_\\_dataent\\_GLOBRADS](http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_GLOBRADS) (last access: 9 February 2011), 2006–2011.
- Veselovskii, I., Kolgotin, A., Griaznov, V., Müller, D., Wandinger, U., and Whiteman, D. N.: Inversion with regularization for the retrieval of tropospheric aerosol parameters from multiwavelength lidar sounding, *Appl. Optics*, 41, 3685–3699, doi:10.1364/ao.41.003685, 2002.
- Wandinger, U. and Ansmann, A.: Experimental determination of the lidar overlap profile with Raman lidar, *Appl. Optics*, 41, 511–514, doi:10.1364/AO.41.000511, 2002.

- Watts, P. D., Bennartz, R., and Fell, F.: Retrieval of two-layer cloud properties from multispectral observations using optimal estimation, *J. Geophys. Res.-Atmos.*, 116, D16203, doi:10.1029/2011jd015883, 2011.
- Welton, E. J., Campbell, J. R., Spinhirne, J. D., and Scott, V. S.: Global monitoring of clouds and aerosols using a network of micro-pulse lidar systems, in: *Conference on Lidar Remote Sensing for Industry and Environment Monitoring*, Vol. 4153, 151–158, Sendai, Japan, doi:10.1117/12.417040, 2000.
- Whiteman, D. N., Melfi, S. H., and Ferrare, R. A.: Raman lidar system for the measurement of water-vapor and aerosols in the Earth's atmosphere, *Appl. Optics*, 31, 3068–3082, doi:10.1364/AO.31.003068, 1992.
- Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, *J. Atmos. Ocean. Tech.*, 26, 2310–2323, doi:10.1175/2009JTECHA1281.1, 2009.
- Woodhouse, I. and Agnew, J.: Chilbolton AERONET Level 2.0 real time data, NASA GSFC, available at: <http://aeronet.gsfc.nasa.gov> (last access: 31 March 2013), 2006–2011.