# Estimation of the lidar overlap function by non-linear regression

Adam C. Povey,[1,*] Roy G. Grainger,[1] Daniel M. Peters,[1] Judith L. Agnew,[2] and David Rees[3]

[1]*Department of Atmospheric, Oceanic, and Planetary Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford, OX1 3PU, United Kingdom*

[2]*STFC Rutherford Appleton Laboratory, Harwell, Oxford, OX11 0QX, United Kingdom*

[3]*Hovemere Ltd., Tonbridge, TN9 1RF, United Kingdom*

[*]*Corresponding author: povey@atm.ox.ac.uk*

The overlap function of a Raman channel for a lidar system is retrieved by non-linear regression using an analytic description of the optical system and a simple model for the extinction profile, constrained by aerosol optical thickness. Considering simulated data, the scheme is successful even where the aerosol profile deviates significantly from the simple model assumed. Application to real data is found to reduce by a factor of 1.4 – 2.0 the root-mean-square difference between the attenuated backscatter coefficient as measured by the calibrated instrument and a commercial instrument. © 2012 Optical Society of America

*OCIS codes:* 010.0010, 010.3640, 120.0120, 120.0280, 280.0280, 280.3640.

## 1. Introduction

Lidar is an active remote sensor that observes backscattering from a pulsed laser beam by particulates in the atmosphere, measuring the number of photons scattered as a function of distance from the laser (determined from the delay between the laser pulse and observation) [1–3]. A lidar's overlap function describes the efficiency with which light is coupled into its detectors as a function of height. Incomplete coupling near the instrument limits the accuracy with which lidar can be used to investigate the planetary boundary layer (PBL), where aerosol is both most abundant and most variable. Further, many methods of lidar analysis are designed to only consider regions where the overlap function is constant and can incur significant errors if regions where it is not constant are considered [4, 5].

For a well-aligned system, the overlap function will tend smoothly from zero at the surface to unity at the height where the telescope's field-of-view (FOV) and laser beam overlap completely, above which it remains constant. If the system is not well-aligned, the overlap function can take a more complex form as a result of image defocusing near the telescope, the obstruction of the telescope's secondary mirror, and incomplete overlap of the FOV and the volume illuminated by the laser.

A wide range of techniques have been proposed to estimate the overlap function. When the arrangement of the system is well known, the overlap function can be determined directly using ray tracing techniques [6, 7] or approximated by an analytic formulation [8–10]. Otherwise, the overlap function can be estimated from measurements if the aerosol component of scattering is known. When the aerosol distribution is homogeneous in time or space, the signal from a reference height is sufficient correction [11, 12]. Independent measurement of the aerosol profile, such as by a ceilometer, can be used, though an iterative scheme is necessary to compensate for errors in that measurement [13]. For a Raman lidar, the elastic profile can be used in a similar fashion to correct the Raman channel used in deriving the extinction coefficient [14]. It has also been proposed to fit a polynomial to the range-corrected elastic profile in the region where complete overlap is expected and use this to estimate the scattering lower in the atmosphere [15]. If the capacity to scan the laser beam is available, it is possible to use multi-angle measurements to constrain the aerosol profile and derive a best-fit to the overlap function [16, 17].

The Robust And Compact Environmental Lidar (RACHEL) was developed by Hovemere Ltd. as a cost-effective and portable Raman lidar system for unattended monitoring of pollution by day and night. The prototype system was deployed at the NERC Chilbolton Facility for Atmospheric and Radio Research (CFARR) in southern England during Spring 2010. However, due to a series of minor faults, the laser operated at significantly less than its full power and had to be adjusted repeatedly in the field, such that the usually assumed analytic form could not be used. The low signal-to-noise ratio (SNR) of the data prevented iterative methods [13, 14] from converging to a sufficiently smooth solution to be useful. Further, the large, inhomogeneous aerosol loading ruled out integrating the signal over longer time periods to increase the SNR. A new method was sought to produce the best estimate of the overlap function possible under these challenging conditions.

The concept is to use non-linear regression as a framework to combine the noisy lidar profiles with a model of the measuring system and known information about the atmosphere, such as the aerosol optical thickness, to produce an approximation to the true overlap function that makes optimal use of all the information available. The methods of non-linear regression are applied across the scientific community [18, 19] such that it is easier for a scientist unfamiliar with lidar to comprehend this data analysis process than existing *ad hoc*

methods.

## 2. Method

Assuming a continuous, circular beam profile and by modelling a Cassegrain telescope as a thin lens with a circular obstruction viewed through a pinhole, it is shown in [8] that the overlap function can be written as,

$$\pi R_T^2 A(r) = \left[\frac{\gamma}{\nu(r)w(r)}\right]^2 \left\{ S_H\left[\frac{\nu(r)}{\gamma}R_T, r\right] - S_H\left[\frac{\nu(r)}{\gamma}R_o, r\right]\right\} \quad (1)$$

$$S_H(b,r) = \frac{1}{\pi}\int_{\mu=0}^{w(r)+d(r)} \mathcal{A}\left[\frac{R_p r}{f\gamma}, b; \mu\right] \mathrm{d}\mathcal{A}[\mu, w(r); d(r)] \quad (2)$$

$$\gamma = 1 + \frac{\Delta}{f} \quad (3)$$

$$\nu(r) = \left|\gamma - \frac{\Delta\, r}{f^2}\right| \quad (4)$$

$$w(r) = R_L + \phi_L\, r \quad (5)$$

$$d(r) = \sqrt{(\delta + \phi_\parallel\, r)^2 + (\phi_\perp\, r)^2}, \quad (6)$$

where $R_T$ is the radius of the telescope's primary mirror; $R_o$ is the radius of the central obstruction; $R_p$ is the radius of the pinhole; $f$ is the telescope's focal length; $\Delta$ is the distance of the pinhole from the telescope's focal plane away from the lens; $R_L$ is the initial radius of the laser beam; $\phi_L$ is the beam's divergence; $\delta$ is the distance between the optical axis of the telescope and laser with angles $\phi_{\parallel,\perp}$ between them parallel and perpendicular to the plane joining them; and $\mathcal{A}(R_1, R_2; \mu)$ gives the area of overlap of two circles, radii $R_1$ and $R_2$, with centres a distance $\mu$ apart (see the Appendix).

This formulation of the overlap function was selected over other possibilities as it most thoroughly describes a Cassegrain telescope whilst retaining an analytic form. It does, though, assume that the only relevant component in the optical system is the telescope. Different laser beam shapes can be considered, but are significantly more expensive to evaluate. Also, it would be expected that the overlap function would be affected by incomplete illumination of inhomogeneous components in the optical system after the telescope. This could increase or decrease the coupling at different heights and could be different for different channels of the instrument. The effect of these simplifications on the ability for this model to represent a true overlap function will be considered in the next section.

The energy observed by a Raman lidar system from a height bin ending at range $r$, after correcting for detector non-linearity and subtracting background, can be summarised by the

equations [1],

$$\begin{aligned} E_{\mathrm{el}}(r) &= E_0\, C_{\mathrm{el}}\, r^{-2} A_{\mathrm{el}}(r)\, [\beta_m(\lambda_L, r) + \beta_a(\lambda_L, r)] \\ &\quad \times \exp\left[-2 \int_0^r \alpha_m(\lambda_L, r') + \alpha_a(\lambda_L, r')\, \mathrm{d}r'\right] \qquad (7) \\ E_{\mathrm{ra}}(r) &= E_0\, C_{\mathrm{ra}}\, r^{-2} A_{\mathrm{ra}}(r)\, N_{\mathrm{X}}(r) \\ &\quad \times \exp\left[-\int_0^r \alpha_m(\lambda_L, r') + \alpha_a(\lambda_L, r') \right. \\ &\qquad \left. + \alpha_m(\lambda_{\mathrm{X}}, r') + \alpha_a(\lambda_{\mathrm{X}}, r')\, \mathrm{d}r'\right], \qquad (8) \end{aligned}$$

where the subscripts 'el' and 'ra' denote the elastic and Raman channels of the lidar; $\beta(\lambda, r)$ is the backscattering coefficient at a wavelength $\lambda$; $\alpha(\lambda, r)$ is the extinction coefficient; the subscripts $m$ and $a$ denote molecular and aerosol scattering; $\lambda_L$ and $\lambda_{\mathrm{X}}$ are the wavelengths of the laser beam and Raman scattered radiation; $E_0$ is the energy of the laser pulse; $A(r)$ is the overlap function; $N_{\mathrm{X}}(r)$ is the number density of the Raman scattering species; and $C$ is constant.

As the terms $N_{\mathrm{X}}(r)$ and $\alpha_m(\lambda, r)$ can be estimated from radiosonde measurements or a standard atmosphere, the elastic profile is a function of the unknowns $A$, $\beta$, and $\alpha$ whilst the Raman profile depends only on $A$ and $\alpha$. As a better posed problem, it is then proposed to determine the overlap function and the extinction from the Raman profile alone. In addition, the magnitude of (8) is significantly less dependent on the scattering properties than the overlap function as $\alpha$ only enters through the transmission term. This should ensure that errors due to $\alpha$ are minimal compared to errors due to $A$. Conversely, in the elastic channel, the overlap function and backscatter have similar impact on the observed profile and so errors in the aerosol profile will be more significant.

The retrieved $A$ and $\alpha$ can be used to approximately correct the elastic profile to estimate the backscatter, which can be validated against the retrieved extinction profile as a check of internal consistency. Also, though deducible from calibration of the Raman profile against molecular scattering where aerosol concentrations can be considered negligible, the constant $C_{\mathrm{ra}}$ will also be considered unknown to account for variations in the value of $A$ as $r \to \infty$.

The overlap function and extinction coefficient are both a function of height, technically giving a number of unknowns equal to twice the number of measurements in the lidar profile, which is a highly under-determined problem. This can be significantly reduced by expressing the two unknown profiles as functions of some set of parameters. Rather than use simple polynomials as in [15], it is believed more reasonable to select a functional form appropriate to the known physics of the system.

The overlap function is so described by equations (1 – 6) in terms of ten variables. Some of these can be measured and are not expected to change over time ($R_T, R_o, R_p, f, R_L, \phi_L$). The remainder, describing the alignment of the telescope and laser, are not easily measured and

could potentially change after any adjustment to the system $(\Delta, \delta, \phi_\parallel, \phi_\perp)$. Hence, known values for the former parameters, given in table 1, will be assumed and the four unknown parameters will be retrieved.

The evaluation of eqn. (2) in the form originally presented in [8] is computationally expensive and unstable for particular combinations of the input parameters. By considering the limits under which $\mathcal{A}$ describes the complete overlap of two circles, as outlined in the Appendix, eqn. (2) can be reduced to a form for which the computational cost is sufficiently low that the calculation of the function and its derivatives takes of order one second with minimal error.

An idealised aerosol scattering profile is then proposed. It is hypothesised from observations [20–22] that stable PBLs, approximately, exhibit a common extinction profile as a function of vertical height above the lidar ($z = r\cos\theta$, where $\theta$ is the zenith angle of the laser beam). This is represented as a constant extinction up to some height $z_0$ (which is not necessarily the top of the PBL) and a rapid exponential decay above that over a scale height $H$. This is similar to the form proposed in [23], but with a discontinuity of gradient at the top of the mixed layer. As the aerosol profile is only expected to introduce a minor correction, this simplification is not expected to be important, but any model profile could in theory be used.

As plotted in fig. 1(a), the model extinction profile can be expressed as,

$$\alpha_a(\lambda, z) = \begin{cases} \alpha_0, & z \leq z_0; \\ \alpha_0 \exp\left[\frac{z_0 - z}{H}\right], & z > z_0. \end{cases} \tag{9}$$

This profile can then be constrained by observations of the atmosphere. One possibility is a measurement of the aerosol optical thickness, $\chi$, with a sun photometer. Integrating,

$$\int_0^z \alpha_a(\lambda_L, r')\,\mathrm{d}r' = \begin{cases} \frac{\chi}{H+z_0} z, & z < z_0; \\ \frac{\chi}{H+z_0}\left[z_0 + H\left(1 - \exp\frac{z_0 - z}{H}\right)\right], & z \geq z_0. \end{cases} \tag{10}$$

Eqns. (1), (8), and (10) can then be combined to produce a forward model of a Raman lidar system for use within an optimal estimation framework. Optimal estimation is a non-linear retrieval scheme with rigorous incorporation of any prior information about the state of the system [18]. It solves for $\mathbf{x}$ the inverse problem,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\varepsilon}, \tag{11}$$

where $\mathbf{y}$ is a column vector describing a set of measurements; 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; and $\boldsymbol{\varepsilon}$ gives the noise on that measurement.

In this case, $\mathbf{y}$ will contain the Raman profile $E_{\mathrm{ra}}(r)$; $\mathbf{x}$ will be the parameters $\Delta, \delta, \phi_\parallel, \phi_\perp, z_0, \ln H, C_{\mathrm{ra}}$; and auxiliary information, such as the atmospheric pressure and

temperature profiles, $\chi$, and the remaining parameters of the overlap function are contained in $\mathbf{b}$. The forward model, $\mathbf{F}(\mathbf{x}, \mathbf{b})$, will be eqns. (1) and (10) substituted into (8).

The values of $\mathbf{x}$ can be expected to fall within some physically reasonable range, such as the distance $\delta$ being of similar order to the measured distance between the telescope and the laser. These expectations are summarised with an *a priori* state vector $\mathbf{x}_a$. The variations about these expected values are described by the covariance matrix $\mathbf{S}_a$, here assumed to be a diagonal matrix with each value giving the variance of the corresponding parameter. This prior information constrains the algorithm to consider solutions that are deemed physically reasonable, such as $z_0$ and $H$ being of order 0.1 – 1 km. Solutions using $z_0$ of tens of kilometres would clearly be unphysical and should be suppressed.

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,

$$\begin{aligned} -2\ln \mathrm{P}(\mathbf{x}|\mathbf{y}) &= [\mathbf{y}-\mathbf{F}(\mathbf{x},\mathbf{b})]^T\mathbf{S}_\varepsilon^{-1}[\mathbf{y}-\mathbf{F}(\mathbf{x},\mathbf{b})] \\ &\quad +[\mathbf{x}-\mathbf{x}_a]^T\mathbf{S}_a^{-1}[\mathbf{x}-\mathbf{x}_a]+c, \end{aligned} \tag{12}$$

where $c$ is an arbitrary constant. This quantity is hereafter referred to as the cost. For the purpose of comparison, it will always be quoted normalised by the number of measurements in the vector $\mathbf{y}$. The covariance matrix $\mathbf{S}_\varepsilon$ describes both the random experimental error and any known systematic errors in the forward model, such as uncertainties in the parameters $\mathbf{b}$. Estimating the PDFs of the parameters of $\mathbf{x}$ using the Metropolis-Hastings algorithm [24] revealed that only $H$ was not well described by a Gaussian. However, $\ln H$ was found to be approximately Gaussian and so will be retrieved instead.

The state that minimises this cost function (that with maximal probability) solves,

$$-\mathbf{K}^T(\hat{\mathbf{x}})\mathbf{S}_\varepsilon^{-1}[\mathbf{y}-\mathbf{F}(\hat{\mathbf{x}},\mathbf{b})]+\mathbf{S}_a^{-1}[\hat{\mathbf{x}}-\mathbf{x}_a]=0, \tag{13}$$

where $\mathbf{K}(\hat{\mathbf{x}})=\nabla_{\mathbf{x}}\mathbf{F}(\hat{\mathbf{x}},\mathbf{b})$.

It is shown in [18] that the iteration,

$$\mathbf{x}_{i+1}=\mathbf{x}_i+[(1+\Gamma_i)\mathbf{S}_a^{-1}+\mathbf{K}_i^T\mathbf{S}_\epsilon^{-1}\mathbf{K}_i]^{-1}\{\mathbf{K}_i^T\mathbf{S}_\epsilon^{-1}[\mathbf{y}-\mathbf{F}(\mathbf{x}_i,\mathbf{b})]-\mathbf{S}_a^{-1}(\mathbf{x}_i-\mathbf{x}_a)\}, \tag{14}$$

converges to $\hat{\mathbf{x}}$ with corresponding covariance matrix,

$$\mathbf{S}_{\mathbf{x}}=(\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}+\mathbf{S}_a^{-1})^{-1}. \tag{15}$$

A schematic of the iteration is presented in figure 2.

General practice is that after an iteration, if the cost function, eqn. (12), has increased, the scaling factor $\Gamma_i$ is increased by a factor of ten. Otherwise, it is reduced by a factor of

two. Iteration ceases when either the cost function or all elements of the state vector change by less than some threshold after a step. In this work, an initial $\Gamma$ of $10^5$ is used over a maximum of 30 iterations, which are considered converged if the relative change in cost is less than $10^{-6}$ or the change in all parameters is less than 10 % of the estimated error (the square root of the diagonal elements of $\mathbf{S_x}$) in that parameter. Most retrievals were found to converge due to the latter condition.

The first guesses used for the optical parameters (rows labelled $\mathbf{x}_0$ in tables 2 and 3) were selected by experimentation to minimise the number of iterations to achieve convergence in simulated data. Other guesses are possible, but were generally found to take longer to converge. An *a priori* state of a well-aligned, monostatic system was used in all retrievals presented, with large variances selected to allow a wide range of possible states (row $\sigma_a$ in table 2). A bistatic system would instead use the known separation of the laser and telescope axis as the first guess and *a priori* of $\delta$. Though such systems are not presented in this paper, the method has performed equivalently in simulation. Equal first guesses and *a priori* values were used for aerosol parameters and these were estimated from the expected conditions, which vary from case to case.

The number density profile, $N(r)$, can be estimated from interpolation of nearby radiosonde profiles or a standard atmosphere calibrated to surface observations. Using the Rayleigh scattering cross-section derived in [25], the molecular components of scattering were estimated from these data. The retrieval is found to be insensitive to the expected errors of $\sim 2$ K for sondes launched within a day and 200 km of the lidar observation [26], though these are included within the estimate of $\mathbf{S}_\epsilon$.

Superficially, this method resembles that recently presented in [17]. That used a variation of the mathematical model in [10], which is itself a modification of the work of [8], but they neglect the obstruction of the telescope's secondary mirror. This is most important very near the telescope and the desire to simplify the problem can be appreciated. However, the lack of its inclusion for a monostatic system does make the definition of the range of minimum overlap ($r_{\min}$ in [10]) poorly defined and will be shown (within this retrieval scheme) to produce inconsistent overlap functions.

The methods presented in this section, though, differ significantly in the particulars of the evaluation of the retrieval. Profiles at two different zenith angles are used in [17], effectively using the more slanted profile as an estimate of the true aerosol profile to correct the other within an *ad hoc* iterative correction. Though that method was clearly successful, the approach of optimal estimation proposed in this paper has a simple, statistical definition for the retrieved state (that with the maximal probability given the measurements) and has been proved in several decades of satellite retrievals [27–30]. The use of a more widely understood and applied method of iteration is considered worth the slight increase of the complexity of

the calculations.

## 3. Simulation

The behaviour of the retrieval scheme was investigated through the use of data simulated for the RACHEL platform. Such simulations are easily produced with the forward model. Measurement noise is assumed uncorrelated and to be entirely Poisson-distributed counting noise. It is simulated for the addition of 60,000 laser shots at the average power observed in field tests to give a sufficient signal-to-noise ratio up to 3 km for all models considered.

Four alignments of the system are considered, highlighting the expected range of states. The exact parameters are summarised in the rows labelled $\mathbf{x}_{\text{true}}$ in table 2. The Raman profile simulated for each case is plotted in fig. 3, marked with $\Diamond$, $\triangle$, $\square$, and $\times$ to indicate models 1 – 4, respectively. Plotted in black over these is the forward modelled profile output after a retrieval. The calibration function, defined by the product,

$$\mathcal{C} = C_{\text{ra}} A(r), \tag{16}$$

is plotted using the same symbols for both simulated profiles and retrievals in fig. 4.

Models 1 ($\Diamond$) and 4 ($\times$) correspond to poorly-aligned systems, with large misalignments between the laser and the telescope optical axis. As seen in fig. 4, these exhibit low level maxima and tend to a constant value with height, such that features can be observed throughout the expected range, but with inferior coupling to the detectors.

Models 2 ($\triangle$) and 3 ($\square$) present well-aligned systems, though model 2 uses a telescope focused to 860 m rather than several kilometres. These overlap functions develop in the expected fashion of increasing monotonically to a maximum, but greater coupling below 800 m is achieved in the lower focused system at the expense of reduced coupling above that.

All retrievals used the same first guess and *a priori* state vectors, given by the rows labelled $\mathbf{x}_0$ and $\mathbf{x}_a$ of table 2. The *a priori* covariance matrix is diagonal, with elements given by the square of row $\sigma_a$. The parameters retrieved from each model's simulated profile are given in rows labelled $\hat{\mathbf{x}}$, with the errors in brackets.

The four retrievals demonstrate a high quality fit in all cases, with normalised costs in the range 0.94 – 0.98. Good fits are expected to have a normalised cost of approximately unity. The correlation with the true profile in all cases is very high and profiles are consistent with the truth at all levels, where the error for each point of $\mathcal{C}$ is estimated by the diagonal elements of the covariance matrix,

$$\mathbf{S}_{\mathcal{C}} = \mathbf{K}' \mathbf{S}_{\mathbf{x}'} (\mathbf{K}')^T, \tag{17}$$

where $\mathbf{x}' = (\Delta, \delta, \phi_{\parallel}, \phi_{\perp}, C_{\text{ra}})^T$; $\mathbf{K}' = \nabla_{\mathbf{x}'} \mathcal{C}$; and $\mathbf{S}_{\mathbf{x}'}$ is identical to $\mathbf{S}_{\mathbf{x}}$ but with the rows and columns corresponding to $z_0$ and $H$ removed.

The retrieval does have a slight tendency towards underestimation with height, most pronounced in the retrieval of model 4. This is to be expected as this model has the calibration function of lowest magnitude. This being directly proportional to the number of photons observed, the SNR will be lower for this profile, corresponding to a lower information content in the retrieval. When the integration time of this profile is increased, the fit is found to be equivalent to the others.

To investigate the suitability of the idealised extinction profile, a variety of perturbations were added, shown in fig. 1(b–c), and simulated using the well-aligned parameters of model 3. All are constructed to have an equal $\chi$ of 0.4 at 355 nm. The first perturbation considered is the introduction of a Gaussian peak at 500 m of width 25 m, analogous to an optically thin cloud in the boundary layer. Then, two periods of a sinusoid are introduced between 230 and 590 m, roughly analogous to aerosol layers. The profile is multiplied by noise normally-distributed with mean 1 and standard deviation 0.05 in the third case. Finally, the non-physical but illustrative case of linearly decreasing extinction over a height 1.3 km is considered.

The retrieved calibration functions are plotted in fig. 5 with the retrieved parameters summarised by the rows labelled $\bar{\mathbf{x}}$ in table 2. In the first three cases, the retrieval successfully produces a calibration function consistent with that used for simulation. Even in the case of a linear decrease in extinction with height, the retrieval is accurate to within 5 % despite the significant difference in extinction profile, though this difference is greater than the predicted error.

Hypothetically, the algorithm should work equally well for all values of $\chi$, the important point being if the profile is well represented by eqn. (9). This was justified in simulations explored up to $\chi = 2$ giving equivalent answers. However, in reality, large aerosol loadings are more likely to be significantly variable and would be expected to be less well described by the simple model, which should be evident from a poor quality of fit from the retrieval.

It should be noted that in all retrievals thus far, there has been insufficient cause for the retrieval to update the estimate of $\phi_\perp$ from the *a priori*, as indicated by the errors remaining at their *a priori* value. The consistency of the $\mathcal{C}$ retrieved with its true form implies that the impact of this parameter within the forward model can be duplicated by the other parameters. It would therefore be advisable in future to either not retrieve this parameter at all or to select another appropriate parameter if some observed states are found to be poorly modelled.

These profiles of fig. 5 also demonstrate the significant degeneracy present in the model, whereby a particular functional form can be produced by a variety of parameters (for example, compare $\hat{\mathbf{x}}$ to $\mathbf{x}_{\text{true}}$ for model 4 in table 2). Due to this, a local minima of the cost function can occur (which the algorithm might converge to) where a very particular set of

parameters produce a $\mathcal{C}$ more consistent than others in that area of state space. Such local minima can be identified by the high cost of their solution. Avoiding these is a fundamental difficulty of optimal estimation. Applying several different first guesses to the retrieval and selecting the result with the lowest cost can be useful or, in extreme cases, simulated annealing can be used [18]. As such, it is important in this retrieval to evaluate its success against the calibration function and not against the values of the parameters, which may not be physically meaningful but produce a calibration function consistent with the true form (and therefore useful).

Though the retrieval was seen to be robust against the perturbations of 5, there will always be a difference between the idealised model extinction profile and the true profile. This will introduce an error to the retrieval. This is referred to as forward model error and is very difficult to evaluate, both because it requires a knowledge of the true state vector and because it requires the very expensive evaluation of the change in the retrieved state as a function of perturbations to the measurement vector, $\mathbf{G}_y$. To estimate the magnitude of its effect, a profile was simulated using the extinction profile of [23] (using the error function), a retrieval performed to obtain $\hat{\mathbf{x}}$, and $\mathbf{G}_y$ was evaluated by rerunning the retrieval with each measurement perturbed by one count in turn. The error was then estimated by,

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

where $\mathbf{f}'$ is the true, noiseless simulated profile. A selection of other alternative model profiles, such as a linear decrease in extinction, were also evaluated, though the same $\mathbf{G}_y$ was used.

The forward model error was found to be up to the same magnitude as the originally predicted error, but mostly about an order of magnitude smaller. When cast as the error on $\mathcal{C}$, it was found to insignificantly increase the predicted error by only a few percent above 500 m. Below this height, the forward model error at most doubled the total error. As such, for real data, this error component could be conservatively estimated by doubling the given error below 500 m if a full calculation is impractical. This height appears to be driven by the form of the overlap function and not by the aerosol profile.

It is also possible to investigate the impact of a different beam profile on the retrieval. A commonly accepted beam profile is a Gaussian, for which [8] demonstrated that the integral across the beam can be expressed as a polynomial expansion (Section VI.B). Using this, the overlap function was recalculated for each of the four model states. The retrieval was then applied assuming a continuous profile. The results are plotted in fig. 6. For model 3, the retrieved calibration function is very similar to the true profile with a low normalised cost of 1.3. The remaining models converge, though less well with costs of 6.0, 2.3, and 2.8, respectively. The discrepancies are primarily in the region 200 – 1000 m, where the overlap of the beam and FOV is most rapidly changing and the beam profile is most important.

The low error estimates indicate that this particular forward model error is not properly accounted for in the retrieval.

To approximate the effect of the optical system, the overlap function from 0 – 3 km was perturbed by up to 5 % using a polynomial, also shown in fig. 6. The polynomial was fit to the difference between the overlap functions of the elastic and Raman channels for a lidar at CFARR (determined by their own method resembling [14]). Retrieval for this case, with cost 1.0, performs better than for the Gaussian beam shape, with difference below 1 km only just above the levels of quoted error. This indicates that, though not explicitly included in the forward model, it is still capable of fitting more realistic forms.

In both these cases, the retrieval is attempting to fit behaviour not explicitly added to the forward model. There are likely several local minima, explaining the fits with high costs. In all of these cases, by using the true parameter values as a first guess, fits with lower costs that are more consistent with the truth can be obtained, demonstrating that the proposed model can approximate more complex systems. In practice, though, randomising the first guess will be an inefficient means to find a suitable minima. Hence, in the analysis of real data, if the algorithm does not converge to a state with cost near unity for a set of reasonable first guesses, it would be advisable to return to the forward model and develop it further to explain these effects.

## 4. Application

The retrieval was then applied to real data. Ideally, observations at night would be used as these have a higher SNR and, sufficiently long after sunset, the PBL will generally be stable. However, as it was not possible to measure $\chi$ at night, a balance was sought by considering the early morning and dawn of two days that showed the minimal variation in $\chi$ within an hour of sunrise — April 8th and 10th, 2010. AERONET measurements at the site gave $\chi$ of 0.13 and 0.48 at 355 nm, respectively. Profiles from 0530 – 0730 were totalled for each morning and corrected for detector non-linearity [31], which was conservatively assumed to introduce a 5 % error. Measurements greater than 4.8 MHz were neglected as poorly corrected for non-linearity. The background was estimated as the mean of measurements above 30 km, where the lidar signal is negligible.

Table 3 outlines for each retrieval the first guess given to the retrieval (rows $\mathbf{x}_0$), the retrieved parameters (rows $\hat{\mathbf{x}}$), the error on those values (in brackets), and the normalised cost of that solution against the data. Larger values of $z_0$ and $H$ were used as a first guess than with the simulated data, based on the experience of technicians at the CFARR. The same *a priori* state and covariance were assumed as in simulations.

For illustration of the effects of the first guess, three variations are presented for April 10th — $\hat{\mathbf{x}}^1$, $\hat{\mathbf{x}}^2$, and $\hat{\mathbf{x}}^3$. The first is the same as used for April 8th, the second assumes the

extinction profile is almost entirely exponential with a greater scale height, and the third is the same as the first, but exchanges the values of $\phi_{\parallel}$ and $\phi_{\perp}$.

The retrieved calibration function from April 8th is plotted in fig. 7(a). For comparison, a simple arithmetic inversion of eqn. (8) is also shown, calculated by correcting the measurements for range, background, molecular scattering, and the *a priori* aerosol profile. Firstly, $\mathcal{C}$ tends smoothly to a constant with height. This would seem to indicate RACHEL was well aligned at this time. The retrieved $\mathcal{C}$ is slightly smaller from 0.5 – 1 km than would be expected from the data without retrieval. This is because the retrieved scale height is greater than initially guessed, such that the transmission in this range is lower (as the aerosol optical thickness is inversely proportional to $H + z_0$ in eqn. (10)). It is further evident that the retrieval is returning larger errors than observed in the simulated data. The reason for this is clear from the broad scatter of data points above 1 km. The ability to fit a physically consistent function regardless is one of optimal estimation's strengths.

The validity of the retrieved extinction profile can be explored through use of the elastic channel. In particular, if a constant lidar ratio is assumed, the attenuated backscatter coefficient,

$$\beta^*(r) = \frac{E_{\rm el}(r)\, r^2}{C_{\rm el} E_0 A(r)} \exp\left[2\int_0^r \alpha_m(\lambda_L, r')\,{\rm d}r'\right] \tag{19}$$

$$\equiv [\beta_m(\lambda_L, r) + \beta_a(\lambda_L, r)] \exp\left[-2\int_0^r \alpha_a(\lambda_L, r')\,{\rm d}r'\right], \tag{20}$$

can be calculated for the elastic channel and compared to that from direct substitution of the retrieved $\alpha_a$ into eqn. (20). This is presented in fig. 7(b), with the retrieved profile in black and the elastic data in blue ($C_{\rm el}$ was calculated by calibration against molecular scattering above 8 km). In addition, plotted in red is the $\beta^*$ published by CFARR from a Leosphere EZ lidar operated continuously at the site, evaluated using Leosphere's commercial algorithm (a lidar ratio of 15 was chosen to give coincidence of all profiles above 6 km).

We can see that in this case, the retrieved $\alpha_a$ is reasonable up to 1.5 km, but then underestimates the scale height. However, the fairly good correspondence between the published $\beta^*$ and that determined from RACHEL's elastic channel (the RMS deviation between them reduces from $2.4 \times 10^{-5}\,{\rm m}^{-1}\,{\rm sr}^{-1}$ without overlap correction to $1.4 \times 10^{-5}$ with it) gives confidence that despite the extinction profile, $\mathcal{C}$ is useful. The difference in $\alpha_a$ may be due to a change in the lidar ratio between the PBL and free troposphere.

Similar plots are then presented for April 10th in fig. 8 with $\mathbf{x}_0^1$, $\mathbf{x}_0^2$, and $\mathbf{x}_0^3$ plotted in black, orange, and green, respectively. RACHEL is clearly not well aligned at this time, with all solutions presenting a maximum around 400 m and then decreasing to a value about half that observed on April 8th. This reduced efficiency has also reduced the SNR, such that suitable data is only available up to about 2 km, half the range of the previous day.

Regardless, substituting $\mathbf{x}_0^2$ into eqn. (20) produces a $\beta^*$ profile qualitatively consistent with independent observations so this solution was then used to correct the elastic profile (with a lidar ratio of 50). Again, this is more consistent with observations than without the correction (RMS deviation reduces from $2.4\times10^{-5}$ to $1.7\times10^{-5}$) such that the retrieval appears to have successfully calibrated the profile even in these poor conditions. However, the large errors on $z_0$ and $C_{\mathrm{ra}}$ practically mean that these data should be used with care as these errors will dominate any attempt to determine the extinction and backscatter coefficients.

Of the other two solutions presented, $\mathbf{x}_0^3$ corresponds to a local minimum, with a cost of 2.3. The calibration function is not consistent with the data and the aerosol profile deviates greatly from the truth. The first case, $\mathbf{x}_0^1$, though, presents a spurious minima of the cost function. An underestimation of the extinction profile above 700 m has been compensated for by increasing the magnitude of the peak of the calibration function to produce a solution that happens to fit the data well. Such solutions are always a possibility with optimal estimation retrieval. Hence, it is important to properly evaluate any solutions to ensure they are physically sensible. Here, the check against the EZ lidar product is very useful, though if this is not available the elastic channel does provide some validation. Further, the low errors predicted for this solution are inconsistent with the quality of the data used. This overconfidence can be an indication of a spurious result.

A brief consideration was also given to the impact of $R_o$ on the retrieval. This, the last term of (1), is neglected in many previous implementations of the formulation of [8]. The retrieval of April 8th was repeated with four alterations of the original model, plotted in figure 9: $R_o = 0$; $R_o = 0$ with $R_T$ also reduced to maintain the original detector area; and $R_o \pm 5$ mm. Perturbation to the size of the secondary mirror produces variation within the quoted error, but its removal entirely gives a significantly different result, which when used to correct the elastic channel does not give results consistent with the independent measurements. Further, the fits without the secondary mirror have a higher cost of 1.9.

However, for data simulated on bistatic systems with $R_o > 0.1$ m, it was found that the difference between the situations described above was within expected error. For such systems, as considered in [10, 17], the last term of (1) can be neglected with minimal error to improve the computational time by a factor of up to two.

It is again clear that $\phi_\perp$ is not well retrieved. The changed initial condition in $\mathbf{x}_0^3$ confirms that this is related to a degeneracy within the forward model of the impact of either angle and indicates that, in a future version of the retrieval, it would be advisable to remove this parameter. The impact of instead retrieving a different parameter, such as the laser's divergence (the most poorly constrained 'known' parameter), is beyond the scope of this paper.

## 5. Conclusions

An optimal estimation scheme has been proposed for the retrieval of the overlap function of a Raman channel of a lidar system. In optimal estimation, a mathematical formulation of the atmosphere and optical system is used to produce a model observation. The product of the retrieval is the state input into this formulation that produces a model observation most consistent with actual measurements and any other known information in a least squares sense. The method was intended to be applied to observations with a low SNR and in conditions with a rapidly varying aerosol distribution where averaging in time to improve the SNR is unreasonable. Such conditions are not well managed by existing methods of determining the overlap function from a single measurement.

By modelling a telescope as a thin lens with a central obstruction and assuming a continuous beam profile, [8] presented an analytic formulation of the overlap function, here eqn. (1). A rearrangement of this equation outlined in the Appendix reduced the processing time and numerical errors associated with the evaluation of this function and its derivatives sufficiently to permit its use within the optimal estimation framework. A simple model of the extinction profile in conditions of a stable PBL was proposed in eqn. (9). This assumes that extinction is constant up to some height, then decreases exponentially over some scale height.

The retrieval scheme was successful in simulated stable conditions for which the model extinction profile could be expected to be suitable. Various perturbations to the assumed aerosol profile were not found to significantly affect the retrieval. Errors introduced due to effects not included within the forward, such as beam profile and the optical system after the telescope, were found to be only important below 500 m and could be conservatively estimated to double the error in that region. These effects primarily introduce local minima which can make finding a suitable solution difficult, though this has not been an issue for studies with the RACHEL system.

The retrieval was then applied to two days of measurements with the RACHEL system. The retrieved overlap function and extinction profile from the Raman channel were used to approximately correct the observations from the elastic channel of the lidar to produce an attenuated backscatter coefficient. This was found to be more consistent with independent observations than without the overlap correction by a factor of 1.4 – 2.0, despite the expected differences between the overlap function for different channels.

The retrieved extinction profiles, though functional, were clearly only approximations to the truth. This first implementation of non-linear regression to the problem of calibrating a lidar can clearly be improved by using more detailed models of the instrument, covering the impact of optics after the telescope, or more involved models of the extinction profile. However, it has hopefully been demonstrated that, despite the relatively simplistic models used, it is still possible to retrieve useful calibrations through the methods of non-linear

regression.

Practically, the retrieved overlap function will be used within some algorithm to derive the extinction and backscatter. For a single-channel, Klett-Fernald scheme [4], as the overlap function is effectively a multiplicative correction, it will simply add to the fractional error in the backscatter coefficient. As is clear from a comparison of figures 4, 7(a), and 8, the exact error depends on the quality of data used in the retrieval. For the conditions considered in Section 4, this varies from about 2 % when the system is well aligned, which is slightly better than that expected from existing techniques of determining the overlap function, to over 10 % when it is not well aligned, which would be the dominant source of error in the retrieval, though this was in part due to the lower SNR of the data used in the retrieval.

The impact of this error on a Raman lidar scheme [5] is more subtle as there the correction is the derivative of the overlap function. In its current form, the dominant error in this derivative is from numerical integration within $S_H$ near regime changes in its formulation and can be over 100 %. However, if an analytic form of the derivative of $S_H$ with respect to $r$ can be found, the errors should be equivalent to those introduced into the single-channel scheme. This is an aim for future work.

This retrieval scheme is most useful for calibration of low-powered lidars, where the integration time required to evaluate other methods of estimating the overlap function can be restrictive. It is also fast enough to be evaluated on a modern laptop in a few minutes and so can be used in the field to confirm optimal alignment of the instrument without the need of external information, such as a ceilometer profile. Further, the method is primarily limited by the physics described in the forward model and can likely be improved by extending the model to describe the impact of other processes, such as a more complex aerosol model, the non-linearity of the detectors, or inhomogeneities in the optical components.

## A. Evaluation of $S_H$

For brevity in this section, the $r$ dependence of all variables will be dropped and the following definition shall be used,

$$\rho = \frac{R_p r}{f\gamma}. \tag{21}$$

The formulation of the overlap function considered requires evaluating the integral,

$$S_H(b, z) = \frac{1}{\pi} \int_{\mu=0}^{w+d} \mathcal{A}(\rho, b; \mu)\, \mathrm{d}\mathcal{A}[\mu, w; d], \tag{22}$$

where,

$$\mathcal{A}(R_1,R_2;\mu) = \begin{cases} 0, & \mu \geq R_1+R_2; \\ \pi\min[R_1^2,R_2^2], & \mu \leq |R_1-R_2|; \\ \mathbb{A}(R_1,R_2;\mu) & \text{otherwise,} \end{cases} \tag{23}$$

$$\begin{aligned} \mathbb{A}(R_1,R_2;\mu) &= R_1^2\cos^{-1}\left(\frac{\mu^2+R_1^2-R_2^2}{2\mu R_1}\right) \\ &\quad +R_2^2\cos^{-1}\left(\frac{\mu^2+R_2^2-R_1^2}{2\mu R_2}\right) \\ &\quad -\tfrac{1}{2}\sqrt{\Upsilon(R_1,R_2;\mu)} \end{aligned} \tag{24}$$

$$\Upsilon(R_1,R_2;\mu) = [(R_1+R_2)^2-\mu^2][\mu^2-(R_1-R_2)^2]. \tag{25}$$

It can be shown that,

$$\mathrm{d}\mathcal{A}(R_1,R_2;\mu) = \begin{cases} -\sqrt{\Upsilon(R_1,R_2;\mu)}\frac{\mathrm{d}\mu}{\mu}, & |R_1-R_2| \leq \mu \leq R_1+R_2; \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

When $w+d \leq |\rho-b|$ the integrand of eqn. (22) is constant, corresponding to where the laser beam is completely within the telescope's FOV. Hence,

$$S_H(b,z,w+d \leq |\rho-b|) = \pi w^2 \mathrm{min}[\rho^2,b^2]. \tag{27}$$

Otherwise, integration by parts using eqn. (26) gives,

$$\begin{aligned} S_H(b,z) &= w^2\mathcal{A}[\rho,b;w+d] \\ &\quad +\frac{1}{\pi}\int_{|\rho-b|}^{\Lambda}\mathcal{A}[\mu,w;d]\sqrt{\Upsilon(\rho,b;\mu)}\,\frac{\mathrm{d}\mu}{\mu}, \end{aligned} \tag{28}$$

where the upper limit $\Lambda = \min[w+d,\rho+b]$ now accounts for where $\mathrm{d}\mathcal{A}=0$. The first term is zero for $w+d<\rho+b$. Hence,

$$S_H(b,z,|\rho-b|>w+d) = 0. \tag{29}$$

Further, the integrand of eqn. (28) is zero where $\mu \leq d-w$. Hence, the lower limit is instead $\Gamma = \max[|\rho-b|,d-w]$. By considering the relative values of $\mu$ and $w-d$, eqn. (23) can be used to write this integral as,

$$\begin{aligned} \kappa_H(b,z) &= \int_{\Gamma}^{w-d}\mu\sqrt{\Upsilon(\rho,b;\mu)}\,\mathrm{d}\mu \\ &\quad +\int_{w-d}^{\Lambda}\mathbb{A}[\mu,w;d]\sqrt{\Upsilon(\rho,b;\mu)}\,\frac{\mathrm{d}\mu}{\pi\mu}. \end{aligned} \tag{30}$$

The remaining cases of $S_H$ then derive from considering various limits, where the remaining integral terms are evaluated by Romberg integration of the ninth order to a fractional accuracy of $10^{-8}$. These forms are practically found to be the most stable.

- $\rho + b < w - d$ — The second term of eqn. (30) does not exist and $S_H$ can be evaluated exactly,

$$\begin{aligned} S_H(b,z) &= w^2 \mathcal{A}[\rho, b; w+d] \\ &\quad + \int_{|\rho-b|}^{\rho+b} \mu \sqrt{\Upsilon(\rho, b; \mu)}\, \mathrm{d}\mu \qquad (31) \\ &= \pi \rho^2 b^2. \qquad (32) \end{aligned}$$

- $|w - d| < |\rho - b| < \Lambda$ — The first term of eqn. (30) does not exist. For $\rho + b < w + d$, $\kappa_H$ is then integrated by parts to avoid rounding errors for $b \ll 1$,

$$\kappa_H(b,z) = \int_{|\rho-b|}^{\Lambda} \mathbb{A}[\mu, w; d] \sqrt{\Upsilon(\rho, b; \mu)}\, \frac{\mathrm{d}\mu}{\pi\mu} \qquad (33)$$

$$= \min[\rho^2, b^2] \mathbb{A}(|\rho - b|, w; d) \qquad (34)$$

$$+ \frac{2}{\pi} \int_{|\rho-b|}^{\rho+b} \mu \cos^{-1}\left(\frac{d^2 + \mu^2 - w^2}{2d\mu}\right) \mathbb{A}(\rho, b; \mu)\, \mathrm{d}\mu. \qquad (35)$$

- $|\rho - b| < d - w < \Lambda$ — The first term of eqn. (30) does not exist, giving ,

$$\kappa_H(b,z) = \int_{d-w}^{\Lambda} \mathbb{A}[\mu, w; d] \sqrt{\Upsilon(\rho, b; \mu)}\, \frac{\mathrm{d}\mu}{\pi\mu}. \qquad (36)$$

- $d - w < |\rho - b| < w - d < \Lambda$ — Both terms of eqn. (30) exist and the first can be evaluated exactly to give,

$$\begin{aligned} \kappa_H(b,z) &= \frac{(w-d)^2 - \rho^2 - b^2}{4} \sqrt{\Upsilon(\rho, b; w-d)} \\ &\quad + 2\rho^2 b^2 \tan^{-1} \sqrt{\frac{(w-d)^2 - (\rho-b)^2}{(\rho+b)^2 - (w-d)^2}} \\ &\quad + \int_{w-d}^{\Lambda} \mathbb{A}[\mu, w; d] \sqrt{\Upsilon(\rho, b; \mu)}\, \frac{\mathrm{d}\mu}{\pi\mu} \qquad (37) \\ &= \mathbb{B} + \int_{w-d}^{\Lambda} \mathbb{A}[\mu, w; d] \sqrt{\Upsilon(\rho, b; \mu)}\, \frac{\mathrm{d}\mu}{\pi\mu}. \qquad (38) \end{aligned}$$

In summary,

$$S_H(b,r) = \begin{cases} 0, & \rho + b \le d - w; \\ \hline \pi w^2 \min[\rho^2, b^2], & w + d \le |\rho - b|; \\ \hline \pi \rho^2 b^2, & \rho + b \le w - d; \\ \hline \int_{|\rho-b|}^{\rho+b} \sqrt{\Upsilon(\rho,b;\mu)}\mathbb{A}(\mu,w;d)\,\frac{\mathrm{d}\mu}{\pi\mu}, & \rho + b \le w + d \\ & \text{and } |w-d| \le |\rho-b|; \\ \hline \int_{d-w}^{\rho+b} \sqrt{\Upsilon(\rho,b;\mu)}\mathbb{A}(\mu,w;d)\,\frac{\mathrm{d}\mu}{\pi\mu}, & \rho + b \le w + d \\ & \text{and } |\rho-b| < d-w; \\ \hline \mathbb{B} + \int_{w-d}^{\rho+b} \sqrt{\Upsilon(\rho,b;\mu)}\mathbb{A}(\mu,w;d)\,\frac{\mathrm{d}\mu}{\pi\mu}, & \rho + b \le w + d \\ & \text{and } |\rho-b| < w-d; \\ \hline w^2\mathbb{A}(\rho,b;w+d) & w + d < \rho + b \\ + \int_{|\rho-b|}^{w+d} \sqrt{\Upsilon(\rho,b;\mu)}\mathbb{A}(\mu,w;d)\,\frac{\mathrm{d}\mu}{\pi\mu}, & \text{and } |w-d| \le |\rho-b|; \\ \hline w^2\mathbb{A}(\rho,b;w+d) & w + d < \rho + b \\ + \int_{d-w}^{w+d} \sqrt{\Upsilon(\rho,b;\mu)}\mathbb{A}(\mu,w;d)\,\frac{\mathrm{d}\mu}{\pi\mu}, & \text{and } |\rho-b| < d-w; \\ \hline w^2\mathbb{A}(\rho,b;w+d) + \mathbb{B} & w + d < \rho + b \\ + \int_{w-d}^{w+d} \sqrt{\Upsilon(\rho,b;\mu)}\mathbb{A}(\mu,w;d)\,\frac{\mathrm{d}\mu}{\pi\mu}, & \text{and } |\rho-b| < w-d. \end{cases} \tag{39}$$

Thanks to the staff of the NERC-funded Chilbolton Facility for Atmospheric and Radio Research for their assistance during the deployment of the RACHEL lidar. Radiosonde data provided by the UK Meteorological Office from `http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_GLOBRADS`. AERONET level 2.0 data available at `http://aeronet.gsfc.nasa.gov`. Immense gratitude must be extended to the two anonymous reviewers, whose comments significantly improved this paper. This study was supported by a NERC CASE studentship with Hovemere Ltd.

## References

1. R. M. Measures, *Lidar remote sensing: Fundamentals and applications* (Krieger Publishing Company, Malabar, Florida, 1992), 2nd ed.
2. T. Fujii and T. Fukuchi, *Laser remote sensing* (Taylor and Francis, Boca Raton, FL, 2005).
3. C. Weitkamp, *Lidar: range-resolved optical remote sensing of the atmosphere*, Springer Series in Optical Sciences (Springer, New York, 2005).
4. J. D. Klett, "Lidar inversion with variable backscatter extinction ratios," Applied Optics **24**, 1638–1643 (1985).
5. A. Ansmann, M. Riebesell, and C. Weitkamp, "Measurement of atmospheric aerosol extinction profiles with a Raman lidar," Optics Letters **15**, 746–748 (1990).

6. R. Velotta, B. Bartoli, R. Capobianco, L. Fiorani, and N. Spinelli, "Analysis of the receiver response in lidar measurements," Applied Optics **37**, 6999–7007 (1998).
7. I. Berezhnyy, "A combined diffraction and geometrical optics approach for lidar overlap function computation," Optics and Lasers in Engineering **47**, 855–859 (2009).
8. T. Halldórsson and J. Langerholc, "Geometrical form factors for the lidar function," Applied Optics **17**, 240–244 (1978).
9. G. M. Ancellet, M. J. Kavaya, R. T. Menzies, and A. M. Brothers, "Lidar telescope overlap function and effects of misalignment for unstable resonator transmitter and coherent receiver," Applied Optics **25**, 2886–2890 (1986).
10. K. Stelmaszczyk, M. Dell'Aglio, S. Chudzynski, T. Stacewicz, and L. Woste, "Analytical function for lidar geometrical compression form-factor calculations," Applied Optics **44**, 1323–1331 (2005).
11. Y. Sasano, H. Shimizu, N. Takeuchi, and M. Okuda, "Geometrical form-factor in the laser-radar equation — Experimental determination," Applied Optics **18**, 3908–3910 (1979).
12. K. Tomine, C. Hirayama, K. Michimoto, and N. Takeuchi, "Experimental determination of the crossover function in the laser-radar equation for days with a light mist," Applied Optics **28**, 2194–2195 (1989).
13. J. L. Guerrero-Rascado, M. J. Costa, D. Bortoli, A. M. Silva, H. Lyamani, and L. Alados-Arboledas, "Infrared lidar overlap function: An experimental determination," Optics Express **18**, 20350–20359 (2010).
14. U. Wandinger and A. Ansmann, "Experimental determination of the lidar overlap profile with Raman lidar," Applied Optics **41**, 511–514 (2002).
15. S. W. Dho, Y. J. Park, and H. J. Kong, "Experimental determination of a geometric form factor in a lidar equation for an inhomogeneous atmosphere," Applied Optics **36**, 6009–6010 (1997).
16. M. Adam, V. A. Kovalev, C. Wold, J. Newton, M. Pahlow, W. M. Hao, and M. B. Parlange, "Application of the Kano-Hamilton multiangle inversion method in clear atmospheres," Journal of Atmospheric and Oceanic Technology **24**, 2014–2028 (2007).
17. G. Biavati, G. Di Donfrancesco, F. Cairo, and D. G. Feist, "Correction scheme for close-range lidar returns," Applied Optics **50**, 5872–5882 (2011).
18. C. D. Rodgers, *Inverse methods for atmospheric sounding: Theory and practice*, vol. 2 of *Series on Atmospheric, Oceanic, and Planetary Physics* (World Scientific, Singapore, 2000), 2nd ed.
19. A. Gelb, *Applied optimal estimation,* (The Analytic Sciences Corporation, 1974).
20. Q. S. He, C. C. Li, J. T. Mao, A. K. H. Lau, and D. A. Chu, "Analysis of aerosol vertical distribution and variability in Hong Kong," Journal of Geophysical Research —

Atmospheres **113**, 13 (2008).

21. M. Wiegner, S. Emeis, V. Freudenthaler, B. Heese, W. Junkermann, C. Munkel, K. Schafer, M. Seefeldner, and S. Vogt, "Mixing layer height over Munich, Germany: Variability and comparisons of different methodologies," Journal of Geophysical Research — Atmospheres **111**, 17 (2006).
22. R. Maurya, P. K. Dubey, B. C. Arya, and S. L. Jain, "Aerosol vertical profile measurements using micro pulse lidar at New Delhi, India," in "International Symposium for the Advancement of Remote Sensing of the Boundary Layer," (2010).
23. D. G. Steyn, M. Baldi, and R. M. Hoff, "The detection of mixed layer depth and entrainment zone thickness from lidar backscatter profiles," Journal of Atmospheric and Oceanic Technology **16**, 953–959 (1999).
24. S. Chib and E. Greenberg, "Understanding the Metropolis–Hastings algorithm," American Statistician **49**, 327–335 (1995).
25. A. Bucholtz, "Rayleigh-scattering calculations for the terrestrial atmosphere," Applied Optics **34**, 2765–2773 (1995).
26. M. Kitchen, "Representativeness errors for radiosonde observations," Quarterly Journal of the Royal Meteorological Society **115**, 673–700 (1989).
27. C. Marks and C. Rodgers, "A retrieval method for atmospheric composition from limb emission measurements," Journal of Geophysical Research — Atmospheres **98**, 14939–14953 (1993).
28. P. Palmer, J. Barnett, J. Eyre, and S. Healy, "A nonlinear optimal, estimation inverse method for radio occultation measurements of temperature, humidity, and surface pressure," Journal of Geophysical Research — Atmospheres **105**, 17513–17526 (2000).
29. W. Li, K. Stamnes, R. Spurr, and J. Stamnes, "Simultaneous retrieval of aerosol and ocean properties by optimal estimation: SeaWiFS case studies for the Santa Barbara Channel," International Journal of Remote Sensing **29**, 5689–5698 (2008).
30. P. D. Watts, R. Bennartz, and F. Fell, "Retrieval of two-layer cloud properties from multispectral observations using optimal estimation," Journal of Geophysical Research — Atmospheres **116** (2011).
31. D. P. Donovan, J. A. Whiteway, and A. I. Carswell, "Correction for nonlinear photon-counting effects in lidar systems," Applied Optics **32**, 6742–6753 (1993).

Table 1. RACHEL system specification.

| Parameter | Value |
|---|---|
| Laser type | Nd-YAG |
| Emitted wavelength | 354.7 nm |
| Pulse rate | 20 Hz |
| Average pulse energy | 45 mJ |
| Beam radius, $R_L$ | 17.5 mm |
| Beam divergence, $\phi_L$ | 0.3 mrad |
| Telescope model | Meade LXD-75 |
| Primary mirror radius, $R_T$ | 101.5 mm |
| Secondary mirror radius, $R_o$ | 37.5 mm |
| Fibre radius, $R_p$ | 0.2 mm |
| Focal length, $f$ | 2.0 m |
| Resolution | 10.5 m |

Table 2. State vectors used in Section 3. A hat indicates the result of a retrieval, a bar indicates the result of a retrieval on a profile with the aerosol perturbation denoted by the superscript, and the subscript 'true' indicates values used to simulate data.

| Model | Vector | $\Delta$ (mm) | $\delta$ (mm) | $\phi_{\parallel}$ ($10^{-3}$ deg) | $\phi_{\perp}$ ($10^{-3}$ deg) | $z_0$ (m) | $\ln H$ ($\ln$ m) | $C_{\mathrm{ra}}$ ($10^{-17}$ m$^5$ J$^{-1}$) | Cost |
|---|---|---|---|---|---|---|---|---|---|
| All | $\mathbf{x}_0$ | 0.001 | 0.001 | 0.1 | 0.001 | 500 | 3.00 | 1.40 | |
| All | $\mathbf{x}_a$ | 0 | 0 | 0 | 0 | 500 | 5.30 | 4.40 | |
| All | $\sigma_a$ | 10.0 | 10.0 | 10.0 | 10.0 | 1000 | 4 | 10.0 | |
| 1 | $\mathbf{x}_{\mathrm{true}}$ | 12.8 | -2.79 | -11.7 | -3.56 | 906 | 5.15 | 1.50 | |
| 1 | $\hat{\mathbf{x}}$ | 12.84 (0.05) | 2.71 (0.09) | 12.22 (0.02) | 0 (10.0) | 1029 (132) | 4.77 (1.08) | 1.491 (0.009) | 0.96 |
| 2 | $\mathbf{x}_{\mathrm{true}}$ | 4.66 | 4.78 | -1.28 | 0.862 | 176 | 5.47 | 1.36 | |
| 2 | $\hat{\mathbf{x}}$ | 4.61 (0.02) | 1.32 (0.88) | 1.33 (1.58) | 0.37 (5.83) | 662 (97) | 3.30 (3.41) | 1.347 (0.008) | 0.95 |
| 3 | $\mathbf{x}_{\mathrm{true}}$ | 0.617 | 1.26 | -0.700 | 5.97 | 642 | 3.63 | 1.96 | |
| 3 | $\hat{\mathbf{x}}$ | 0.78 (0.04) | 0.056 (0.08) | 5.90 (0.04) | 0 (10.0) | 635 (40) | 3.17 (1.66) | 1.92 (0.01) | 0.95 |
| 4 | $\mathbf{x}_{\mathrm{true}}$ | 0.0113 | -0.356 | 7.89 | -9.88 | 195 | 5.60 | 0.824 | |
| 4 | $\hat{\mathbf{x}}$ | -0.38 (0.13) | -0.39 (0.10) | 1.29 (0.08) | 0 (10.0) | 509 (83) | 3.07 (3.69) | 8.48 (1.07) | 0.98 |
| 3 | $\bar{\mathbf{x}}^1$ | 0.78 (0.05) | 0.035 (0.097) | 5.90 (0.06) | 0 (10.0) | 649 (18) | 3.16 (0.76) | 1.92 (0.02) | 1.00 |
| 3 | $\bar{\mathbf{x}}^2$ | 0.64 (0.11) | 0.02 (0.12) | 6.00 (0.09) | 0 (10.0) | 620 (13) | 4.01 (0.23) | 1.95 (0.03) | 1.03 |
| 3 | $\bar{\mathbf{x}}^3$ | 0.75 (0.07) | -0.10 (0.10) | 5.97 (0.07) | 0 (10.0) | 655 (19) | 3.11 (0.82) | 1.94 (0.02) | 0.96 |
| 3 | $\bar{\mathbf{x}}^4$ | -0.10 (0.14) | 0.06 (0.11) | 6.27 (0.10) | 0 (10.0) | 562 (15) | 5.82 (0.04) | 2.05 (0.04) | 1.00 |

Table 3. State vectors used in Section 4. A hat indicates the result of a retrieval and a superscript denotes the various initial conditions considered.

| Day | Vector | $\Delta$ (mm) | $\delta$ (mm) | $\phi_\parallel$ ($10^{-3}$ deg) | $\phi_\perp$ ($10^{-3}$ deg) | $z_0$ (m) | $\ln H$ ($\ln$ m) | $C_{\mathrm{ra}}$ ($10^{-16}$ m$^5$ J$^{-1}$) | Cost |
|---|---|---|---|---|---|---|---|---|---|
| 8th | $\mathbf{x}_0$ | 0.001 | 0.001 | 10.0 | 0.001 | 1000 | 6.00 | 2.21 | 59.1 |
| 8th | $\hat{\mathbf{x}}$ | 4.50 (0.02) | 13.0 (0.7) | 8.58 (0.11) | 0 (10.0) | 1480 (292) | 5.73 (1.00) | 3.18 (0.04) | 1.33 |
| 10th | $\mathbf{x}_0^1$ | 0.001 | 0.001 | 10.0 | 0.001 | 1000 | 6.00 | 2.21 | 42.9 |
| 10th | $\hat{\mathbf{x}}^1$ | 5.02 (0.27) | -24.5 (1.6) | 14.2 (0.2) | 0.02 (9.99) | 490 (35) | 4.97 (0.27) | 6.92 (0.49) | 0.72 |
| 10th | $\mathbf{x}_0^2$ | 0.001 | 0.001 | 10.0 | 0.001 | 50 | 7.00 | 2.21 | 44.7 |
| 10th | $\hat{\mathbf{x}}^2$ | 6.03 (0.01) | -23.5 (1.8) | 14.3 (0.2) | 0 (10.0) | 54 (971) | 6.59 (0.75) | 5.70 (1.03) | 0.78 |
| 10th | $\mathbf{x}_0^3$ | 0.001 | 0.001 | 0.001 | 10.0 | 1000 | 6.00 | 2.21 | 42.9 |
| 10th | $\hat{\mathbf{x}}^3$ | 2.46 (0.08) | -0.002 (10.0) | 0 (10.0) | 13.0 (0.1) | 734 (68) | 5.17 (0.54) | 7.60 (0.18) | 2.30 |

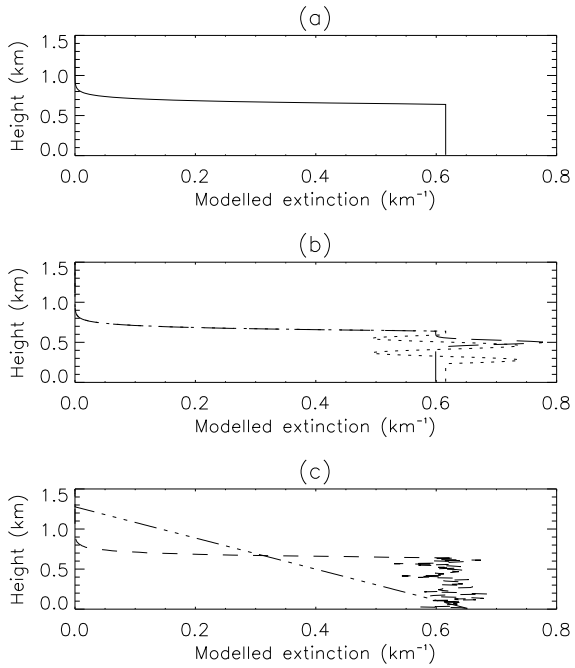


Fig. 1. Aerosol extinction profiles used for simulating data. All have $\chi = 0.4$ at 355 nm. (a) Unperturbed model profile, where extinction is constant to 640 m and decreases exponentially above that with a scale height of 38 m; (b) Addition of a Gaussian peak at 500 m (dashed) or sinusoidal variations from 230 – 590 m (dotted); (c) Addition of normally-distributed multiplicative noise (short dashes) or a linear decrease in extinction (dot-dash) over 1.3 km. model-ext.eps

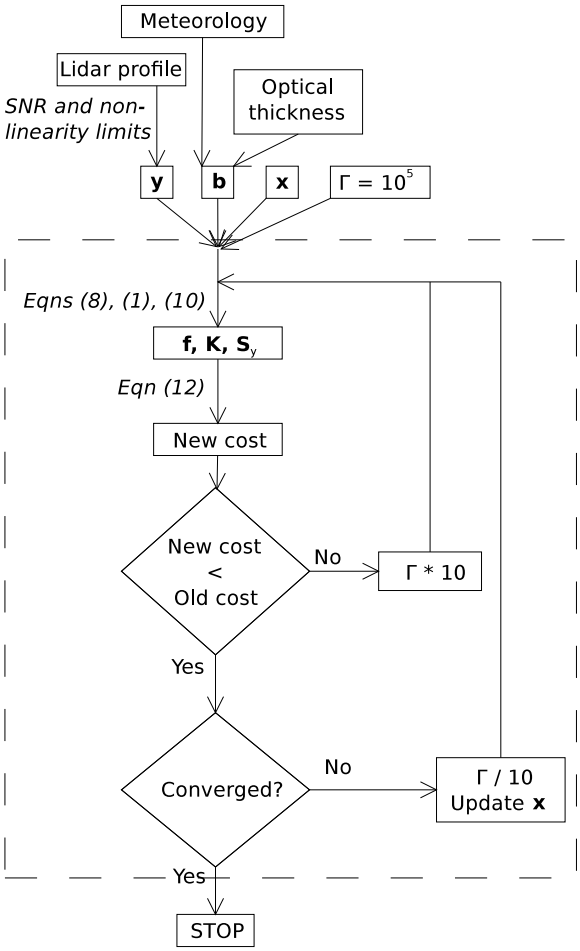


Fig. 2. Schematic of the optimal estimation algorithm. iter.eps

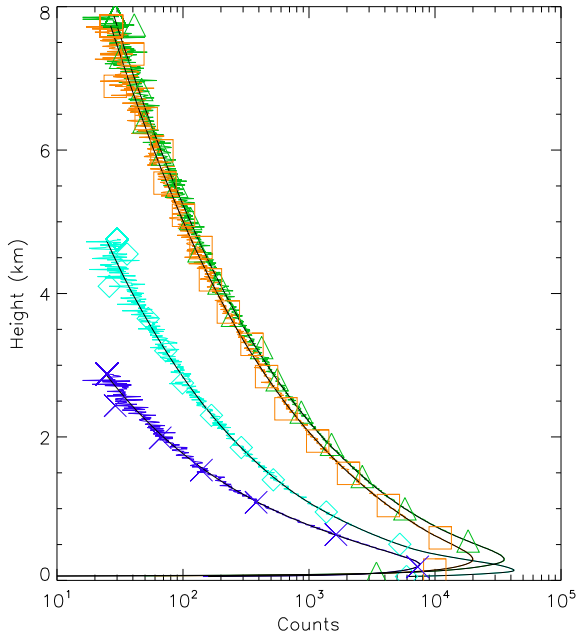


Fig. 3. (Color online) Raman lidar profiles, $\mathbf{y}$, simulated for four representative alignments of the RACHEL system (table 2). Models 1 – 4 are marked with blue $\Diamond$, green $\triangle$, orange $\square$, or purple $\times$, respectively where only points with an SNR greater than two are shown. Plotted in black over each is the forward modelled profile, $\mathbf{F}(\mathbf{x}, \mathbf{b})$, resulting from a retrieval on these data. fits-forward_model.eps

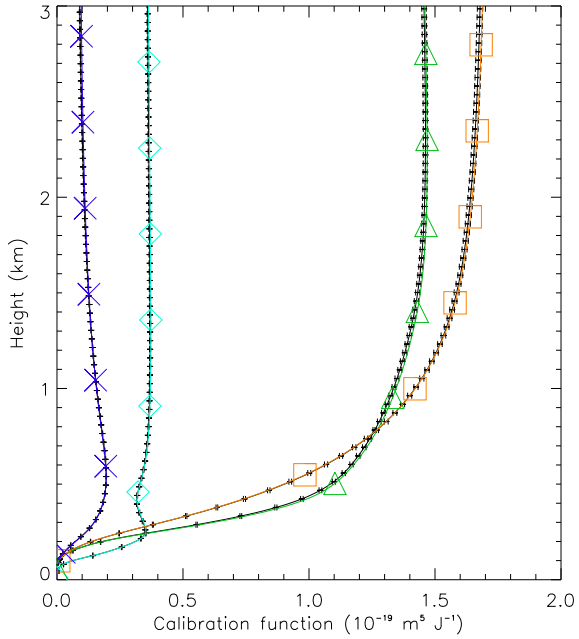


Fig. 4. (Color online) Retrieved calibration function, $C_{\mathrm{ra}}A(r)$, from simulated data of fig. 3, with errors derived from eqn. (15). The profile used to simulate these data is plotted with the symbols from the previous figure. fits-perfect.eps

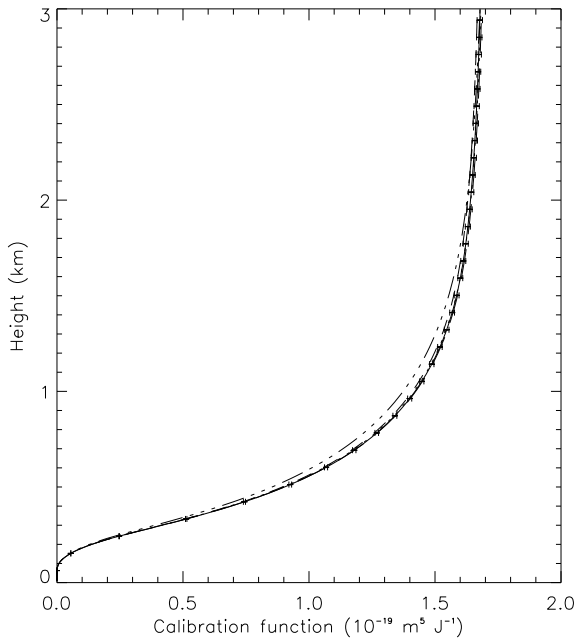


Fig. 5. As fig. 4, but for profiles simulated with a perturbation to their extinction profile, as introduced in fig. 1. The degeneracy of the model states is evident from the very similar profiles, despite differing parameter values (table 2). fits-aerosol.eps

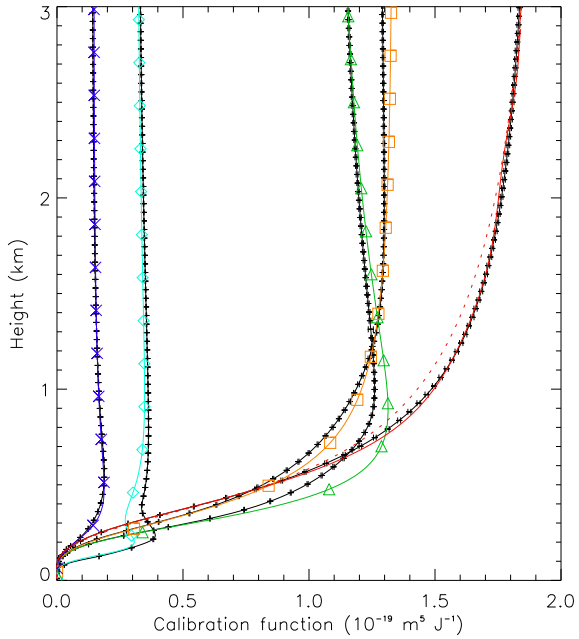


Fig. 6. (Color online) As fig. 4, but for data simulated using features not included in the forward model. Curves marked with characters used a Gaussian beam profile and the red curve perturbed the overlap function (unperturbed function as dotted line). Discrepancies are concentrated in the region below 2 km where the overlap of the beam and FOV is incomplete, such that beam shape is potentially important. fits-gaus.eps

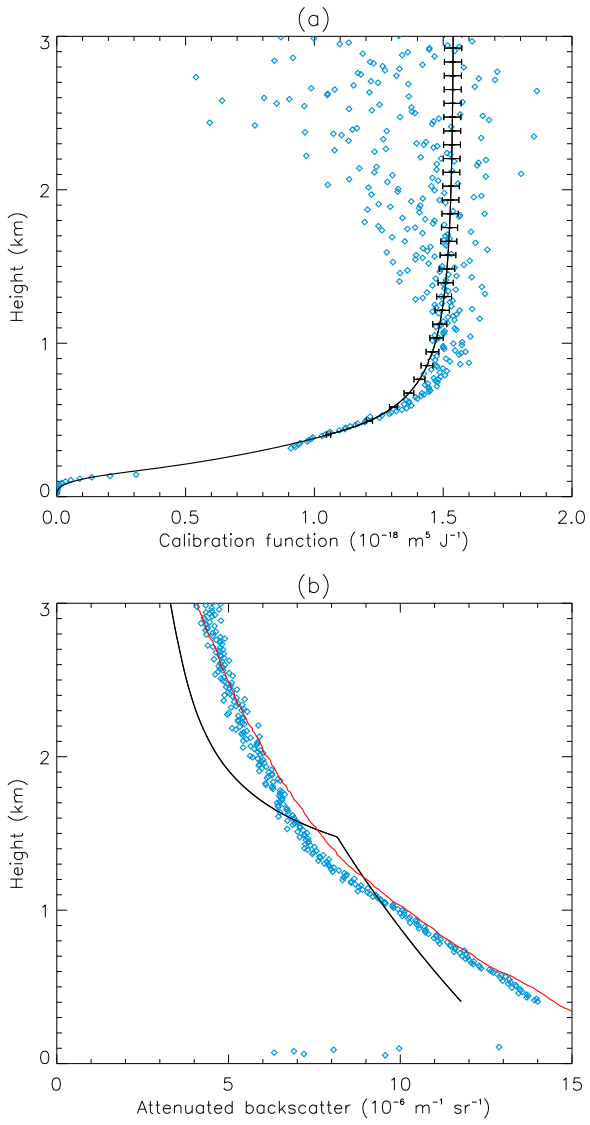


Fig. 7. (Color online) Analysis of April 8th, 2010 when RACHEL was reasonably well aligned. (a) The retrieved calibration function with errors. Plotted as points is an arithmetic inversion of the measurement, eqn. (8), with the *a priori* extinction profile. Measurements beyond the reasonably linear range of the detectors are not plotted and were not used in the retrieval. (b) The attenuated backscatter coefficient at 355 nm for the retrieved aerosol profile (black), the elastic profile corrected with the retrieved overlap function (blue), and as reported independently by an EZ lidar at the same site (red). A lidar ratio of 15 was chosen to give consistency between the three signals above 6 km. retrieved_profile.eps

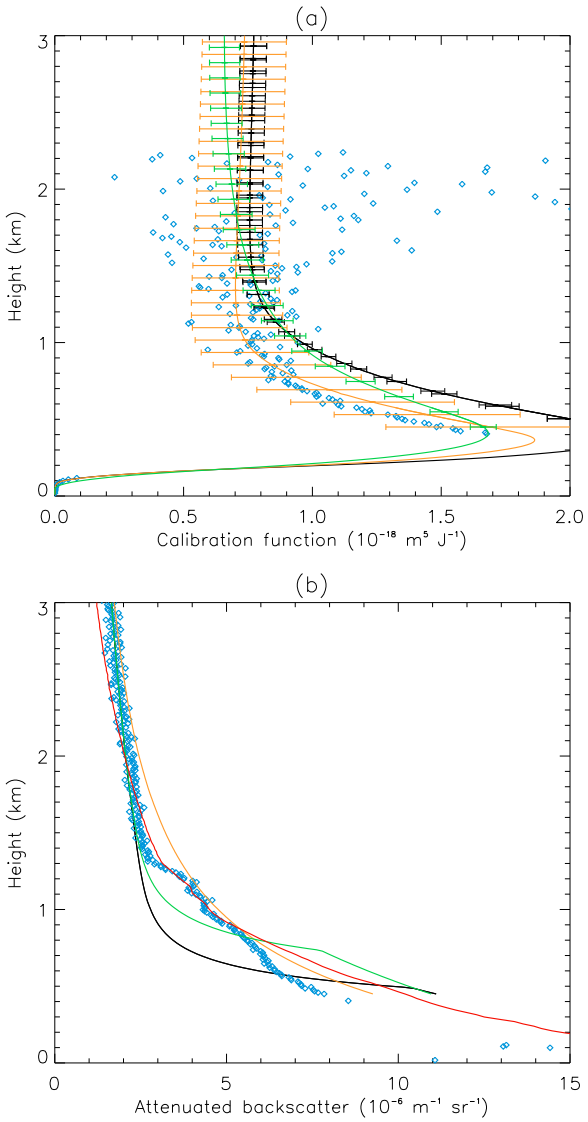


Fig. 8. (Color online) As fig. 7, but for April 10th when RACHEL was significantly misaligned. In addition, the results for three different initial conditions are presented (table 3), conditions 1 – 3 plotted in black, orange, and green, respectively. Measurements with an SNR of less than two are not plotted and were not used in the retrieval. A lidar ratio of 50 was used. Note the greater error introduced by the reduced SNR of the data at 2 km compared to fig. 7. initial_conditions.eps

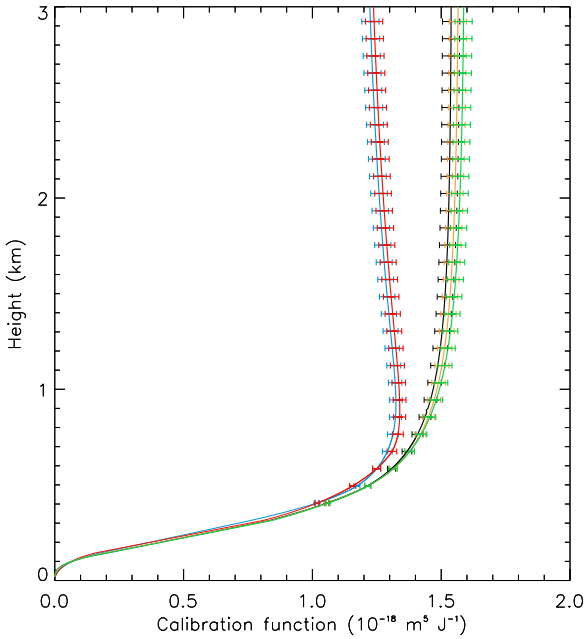


Fig. 9. (Color online) The effect of $R_o$ on retrieval of April 8th, 2010. The calibration function presented in fig. 7 is shown in black, with the functions retrieved with a $\pm 5$ mm perturbation of $R_o$ in orange and green, respectively. Retrieval with $R_o = 0$ is shown in blue and red shows the effect of further reducing $R_T$ to maintain a constant effective telescope area. odia_perturb.eps