

# Development of an optimal estimation retrieval scheme for a Raman lidar system

A.C. Povey

Supervisors: R. G. Grainger, D. M. Peters  
Word Count 7,280 (7,458 with appendices)

## Summary of acronyms

ABC	Attenuated backscatter coefficient
CUV	Chilbolton ultraviolet lidar
DIAL	Differential absorption lidar
LIDAR	Light detection and ranging
MRH	Maximum resolvable height
PMT	Photomultiplier tube
RACHEL	Robust and compact hybrid environmental lidar
RMS	Root-mean square deviation

## Summary of mathematical symbols used

Symbol	Description	Units
$\lambda_{0,x}$	Emission/observation wavelength	m
$r$	Range from instrument	m
$z$	Vertical height above the instrument	m
$c$	Speed of light in vacuum	$\text{m s}^{-1}$
$H[x]$	Heaviside step function, which equals unity for $x > 0$ and zero otherwise	—
$P(\lambda, r)$	Number of photons observed at wavelength $\lambda$ from range $r$	counts
$\eta(\lambda)$	Instrument detection efficiency at $\lambda$	—
$\sigma(\lambda)$	Scattering cross-section at $\lambda$	$\text{m}^2$
$\Omega$	Solid angle	sr
$\mathcal{T}(\lambda, r_0, r_1)$	Atmospheric transmission from $r_0$ to $r_1$ . If only one coordinate given, instead, from 0 to $r$	—
$\alpha(\lambda, r)$	Volume extinction coefficient	$\text{m}^{-1}$
$\beta(\lambda, r)$	Volume backscattering coefficient	$\text{m}^{-1} \text{sr}^{-1}$
$A(r)$	Effective area of system	$\text{m}^2$
$\mathcal{A}(r_1, r_2; \mu)$	Area of overlap between two circles, radii $r_1$ and $r_2$ , with distance $\mu$ between their centres	$\text{m}^2$
$\chi(\alpha, \rho, \mu)$	$[(\alpha + \rho)^2 - \mu^2][\mu^2 - (\alpha - \rho)^2]$	$\text{m}^4$
$S(z)$	$\ln[z^2 P_s(\lambda, z)]$	
$S_H(\rho)$	The Halldórsson $S$ function	$\text{m}^4$
$B$	Lidar ratio	sr
$k$	Ångstrom coefficient	—
$w(r)$	Half-width of the laser beam at range $r$	m
$\tau$	Dead time of the PMT	s
$\Delta t$	Laser pulse duration	s
$\mu$	Radial distance from the optical axis of the telescope	m

Symbol	Description	Units
$\mu^*$	Radial distance from the optical axis of laser	m
$R$	Radius of the primary mirror of the telescope	m
$b$	Radius of the secondary mirror of the telescope	m
$a$	Radius of the exit aperture of the telescope	m
$f$	Focal length of the telescope	m
$z^*$	Displacement of the telescope aperture from its focal plane	m
$\delta$	Distance between telescope axis and laser pupil	m
$\phi_{\parallel,\perp}$	Angle between the axes parallel/perpendicular to the plane defined by the telescope axis and laser pupil	rad
$h$	Planck's constant	J s
$I$	Radiance	$\text{W m}^{-2} \text{sr}^{-1}$
$n$	Refractive index of standard air	—
$p$	Pressure	Pa
$T$	Temperature	K
$N$	Number density	$\text{m}^{-3}$
$M$	Molar mass	$\text{kg mol}^{-1}$
$N_A$	Avagadro's constant	$\text{mol}^{-1}$
$g$	Acceleration due to gravity	$\text{m s}^{-1}$
$\Gamma$	Atmospheric lapse rate	$\text{K km}^{-1}$
$\mathbf{y}$	Measurement vector	counts
$\mathbf{x}$	State vector	
$\mathbf{x}_a$	<i>A priori</i> state vector	
$\mathbf{F}(\mathbf{x})$	Forward model	counts
$\mathbf{K}$	$\nabla_{\mathbf{x}}\mathbf{F}(\mathbf{x})$	
$\mathbf{S}_\epsilon$	Error covariance of measurement	
$\mathbf{S}_a$	Covariance matrix of <i>a priori</i> state vector	
$P(\mathbf{x})$	Probability of state $\mathbf{x}$	—

**Acknowledgements** *This work would not be possible without the data collected by J. Agnew at the Rutherford Appleton Laboratory nor without the support and advice of D. Rees and everyone at Hovemere. Thanks to J. Lass and S. Harrington for administration, J. Storey for his invaluable instruction, C. Wrench for guidance and support, D. McKay for assistance with data collection, C. Leonard for assistance with the measurements at the end of April, and to everyone else at the Chilbolton Observatory for their patience, conversations, and cable ties. This work is supported by a NERC CASE studentship with Hovemere Ltd.*

# 1 Introduction

## 1.1 Basic Concepts

Active observations of the atmosphere through its scattering of a light source have been performed since the 1930's by photographing the beam of a searchlight [Hulburt, 1937]. Later work used pulsed lights to determine the height of the cloud base from the time-of-flight, after which the technique came to be known as lidar (LIght Detection And Ranging) due to its similarity to radar. The development of the photomultiplier tube (PMT) enabled more robust and continuous measurement [Friedland et al., 1956], but the technique was most widely exploited after the advent of the Q-switched laser. The vastly lower beam divergence and higher energies allowed measurements over much greater ranges with shorter observations, enabling researchers to probe the structure and evolution of the atmosphere in unprecedented detail [Collis, 1966]. Since then, improvements in technology and greater usage of the spectral properties of the atmosphere have ensured the technique remains virtually unrivalled in its ability to resolve atmospheric motions [Amodeo et al., 2007].

A basic lidar system comprises of a pulsed laser beam directed into the atmosphere, observed by a telescope. The telescope is focused onto a system of filters, dichroics, and/or interferometers to separate the profile into channels by polarisation and wavelength, as required, which are measured by PMTs. Time-gating of data collection enables the system to collect a profile of the backscattered power as a function of range from the instrument,  $r = \frac{1}{2}ct$ , where  $c$  is the speed of light and  $t$  is the delay since the emission of the pulse. The received signal attenuates with the square of the range from the instrument and can be observed to ranges of tens of kilometers, such that lidar profiles span several orders of magnitude. Multiple PMTs at different gains and/or arrays of telescopes can be used to observe different regions of the atmosphere reliably. In order to overcome noise, profiles can be averaged over seconds, minutes, or hours, depending on the application and conditions.

The most widespread usage of meteorological lidar currently are as ceilometers — single-wavelength systems from which cloud base and boundary layer height are determined from a profile at the transmitted wavelength (known as the elastic return). These systems, though technically simpler than the alternatives, present distinct challenges when attempting to qualitatively characterise the atmosphere as they make only one measurement at set of heights to characterise a complex system. In order to better constrain the problem, more advanced lidar systems make measurements at multiple wavelengths:

- Depolarising lidar monitor the depolarising properties of the atmosphere on the beam. This gives an indication of the degree of sphericity in aerosol particles and can very useful in distinguishing clouds from aerosols;
- Differential Absorption Lidar (DIAL) transmit two similar wavelengths of light, one on a significant absorption feature of a gas under investigation and another away from this feature. Such systems have been used to monitor trace gas distributions since the 1970's [Grant, 1991];
- Raman lidar monitors specific species by observing the strongest line of their Raman spectrum. Raman scattering is an inelastic molecular scattering process whereby incoming light of an arbitrary wavelength excites

vibrational and/or rotational motions in the molecule. Photons are then emitted across a discrete spectrum relative to the original wavelength. As the cross-section for Raman scattering is several orders of magnitude smaller than that for elastic scattering, such systems require a high degree of optical blocking to prevent cross-talk between the elastic and inelastic channels;

- Doppler lidar measures the frequency shift of the backscattered light to monitor wind speeds along the beam line.

Several different methods of analysis for lidar data have been proposed, but none give a rigorous derivation of the resulting errors and frequently rely on extensive calibration measurements that researchers cannot practically perform with a common lidar system. This research aims to develop a new method of analysis using an optimal estimation retrieval scheme. Such a scheme can be adapted to include as little calibration data is available and returns rigorously derived estimates of the resulting errors. Data retrieved from this scheme could then be more readily compared to other data sources, vastly increasing the value of lidar data in the atmospheric sciences.

## 1.2 RACHEL

Transmitter	Receiver
Wavelength, 354.7 nm	Primary mirror diameter, 203 mm
Average pulse energy, 65 mJ	Secondary mirror diameter, 75 mm
Pulse length, 5 – 7 ns	Focal length, 2 m
Repetition rate, 20 Hz	Field of view, 0.2 mrad
Beam diameter, 35 mm	Efficiency, 0.03% (elastic); 0.9% (Raman)
Beam divergence, < 0.3 mrad	Dead time, 10 ns (elastic); 70 ns (Raman)

Table 1: Summary of the theoretical parameters of the RACHEL system, deduced from measurements and the manufacturer’s specifications.

The Robust and Compact Hybrid Environmental Lidar (RACHEL), developed by Hovemere Ltd, is a four-channel, coaxial Raman lidar system designed for automated operation and easy transportation with scanning across the entire sky. It was originally designed after the introduction of the Kyoto Protocol as a portable detector to measure CO<sub>2</sub> emissions from businesses to enable monitoring and enforcement of the carbon trading scheme. The optical system (outlined in figure 1) was designed to be highly adaptable, based on a series of interchangeable dichroics and interference filters such that, in theory, any trace gas may be monitored through an appropriate selection. This has enabled the design, over time, to move towards a more generic aim of pollution monitoring, intending to use nitrogen Raman measurements to calibrate the elastic channel for measurements of aerosols and trace gas measurements from other Raman channels, as described in section 2.4.

Currently, a frequency-tripled Nd:YAG laser is used, with the return signal collected by a Schmidt-Cassegrain telescope coaxial with the beam. The relevant parameters are summarised in table 1. The system is currently configured to observe the elastic backscatter and Raman backscattering from nitrogen, water

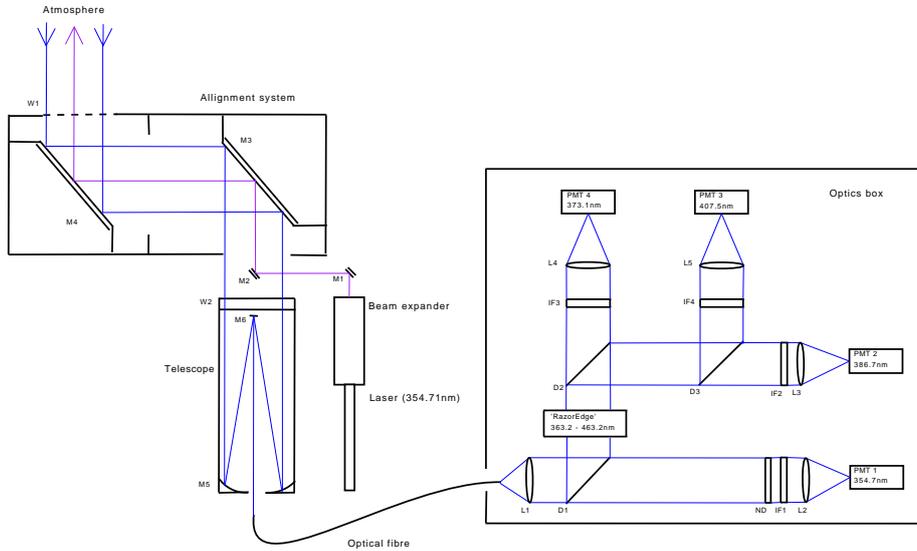


Figure 1: Schematic of the optical components of the RACHEL system, showing the scanning system, telescope, and optics box.

vapour, and carbon dioxide (channels 1 – 4 at wavelengths 354.7, 386.7, 407.5 nm, and 373.1 nm, respectively) with photon-counting PMTs. A simple scanning system has been built consisting of two  $45^\circ$  mirrors mounted on motorised axes. This enables observation of the entire  $2\pi$  of the sky with the capacity for automated scanning programs. This combination of automated, scanning Raman lidar in an easily transported system is currently unique.

In order to keep costs, size, and power consumption to a minimum, a lower powered laser than would usually be utilised in Raman lidar is implemented in combination with a relatively small telescope. Further, a lower-than-optimal dynamic range is derived from the PMTs. This requires careful budgeting of the power in the elastic channel so as to not damage the detectors. The combination of these results in the signal-to-noise ratio (SNR) of the RACHEL system being smaller than typical research-grade systems. As will be shown, though, this range is more than sufficient for its aims, pollution monitoring, which is primarily important in the PBL.

**This report contains five sections and one appendix.** Section 2 summarises the lidar equation, some methods for its solution, and presents a basic model for an instrument. Section 3 introduces a field study conducted at the STFC Chilbolton Observatory in early 2010. An outline of the optimal estimation retrieval method and the current development of a forward model for lidar systems are given in section 4, with conclusions and plans for future work presented in section 5. The appendix contains details from derivations in section 2.

## 2 Response of a lidar system

### 2.1 Attenuation

The most basic unit of a lidar measurement is a single laser pulse. It is transmitted through the atmosphere, backscattered by particles, and transmitted again to be observed by the lidar's telescope. During the propagation, the pulse is attenuated by the atmosphere. For radiation of wavelength  $\lambda$  travelling from position  $\mathbf{r}_0$  to  $\mathbf{r}_1$ , this is summarised by the transmission function,

$$\mathcal{T}(\lambda, r_0, r_1) = \exp \left[ - \int_{\mathbf{r}_0}^{\mathbf{r}_1} \alpha(\lambda, r') \cdot d\mathbf{r}' \right], \quad (1)$$

where  $\alpha(\lambda, r')$  is known as the extinction coefficient.

The extinction coefficient can be written as the sum of the extinctions due to each component of the atmosphere. In lidar studies, the primary components of interest are molecular and aerosol scattering. Molecules are substantially smaller than the laser wavelength and so are well described by Rayleigh scattering. Any residual extinction is attributed to aerosols as the laser wavelength is generally selected to avoid any significant gas absorption features of the atmosphere.

Rayleigh scattering is a well-known and characterised process, allowing calculation of the molecular component of the extinction from knowledge of the temperature and pressure of the atmosphere. The total Rayleigh scattering cross-section per molecule is given by King [1923] as,

$$\sigma_m(\lambda) = \frac{24\pi^3}{\lambda^4 N_s^2} \left( \frac{n^2 - 1}{n^2 + 2} \right)^2 F_k(\rho_n), \quad (2)$$

where  $n$  is the refractive index of standard air at  $\lambda$ ,  $N_s$  is the molecular number density of standard air. Standard air is defined by Penndorf [1957] as “dry air containing 0.03% CO<sub>2</sub> by volume at normal pressure 1013.25 mb and having an air temperature of 15°C.” The King correction factor,  $F_k(\rho_n)$ , is a function of the depolarisation factor  $\rho_n(\lambda_0)$ , and is given by King [1923] as,

$$F_k(\rho_n) = \frac{6 + 3\rho_n(\lambda)}{6 - 7\rho_n(\lambda)}. \quad (3)$$

For computational simplicity, this report calculates the Rayleigh scattering cross-section from the empirical fit derived in Bucholtz [1995],

$$\sigma_m(\lambda) = A\lambda^{-(B+C\lambda+D/\lambda)}, \quad (4)$$

where the constants are given in table 2. This demonstrates a fit of better than 0.2% to full calculations in the wavelength region this report will consider.

The volume-extinction coefficient can then be expressed as,

$$\alpha_m(\lambda_0, z) = N(z)\sigma_m(\lambda_0), \quad (5)$$

where  $N(z)$  is the number density of the air.

$\lambda_0(\mu\text{m})$	$A$	$B$	$C$	$D$
0.2 – 0.5	$3.01577 \times 10^{-28}$	3.552142	1.35579	0.11563
> 0.5	$4.01061 \times 10^{-28}$	3.99668	$1.10298 \times 10^{-3}$	$2.71393 \times 10^{-2}$

Table 2: The parameters of equation 4 for the empirical fit of the Rayleigh scattering cross-section from Bucholtz [1995].

## 2.2 The lidar equation

Figure 2 presents a schematic of the scattering of a lidar pulse, duration  $\Delta t$  and wavelength  $\lambda_0$ , observed at a wavelength  $\lambda_x$ . A parcel of air, volume  $V = \frac{1}{2}c\Delta t A_L(r)$  where  $A_L(r)$  is the cross-sectional area of the laser beam, backscatters a fraction  $\frac{V}{A_L}\beta(\lambda_0, \lambda_x, r)\frac{A(r)}{r^2}$  of the incident power into the telescope (where  $\frac{A(r)}{r^2}$  is the solid angle subtended by the telescope). The volume backscattering coefficient,  $\beta(\lambda_0, \lambda_x, r)$ , gives the differential cross-section for backscattering per unit volume per steradian.

The term  $A(r)$  is known as the effective area and describes the coupling of the laser pulse into the detector as a function of range  $r$ . This is most greatly affected by the focusing of the telescope, with the lowest ranges being poorly focused onto the detectors and so having lower effective areas. The term is often decomposed as  $AO(r)$ , where  $A$  is the physical area of the telescope and  $O(r)$  is a dimensionless function of height that varies from zero to unity, known as the overlap function<sup>1</sup>.

When observing a point a distance  $r$  from the instrument, as outlined in section 2.1, the signal is attenuated by a factor  $\mathcal{T}(\lambda_0, 0, r)$  after emission and by a factor  $\mathcal{T}(\lambda_x, r, 0)$  after scattering. In addition, the optical efficiency of the observing system,  $\eta(\lambda_0, \lambda_x)$ , summarises the losses within the optical system and the efficiency of the detectors. Thus, for a monochromatic laser pulse containing  $P_0$  photons<sup>2</sup>, one can express the number of photons observed from a range  $r$  after subtracting background signals as most commonly presented [Argall and Sica, 2003; Measures, 1992],<sup>3</sup>

$$P_s(\lambda, r) = P_0(\lambda_0)\eta(\lambda_0, \lambda_x)\frac{c\Delta t}{2}\frac{A(r)}{r^2}\beta(\lambda_0, \lambda_x, r) \times \exp\left[-\int_0^r \alpha(\lambda_0, r') dr'\right] \exp\left[-\int_0^r \alpha(\lambda_x, r') dr'\right]. \quad (6)$$

## 2.3 The Klett solution for elastic lidar

The most basic mode of operation for any lidar is its elastic mode. This collects a single profile of the atmosphere, such that they must derive values for both the backscatter and attenuation coefficients from a single measurement. As this mode should provide the signal of the greatest magnitude, it is important to be able to analyse it without additional information. This is obviously impossible

<sup>1</sup>Though not strictly correct, the term ‘overlap function’ is often used to describe the effective area as well.

<sup>2</sup>Virtually all lidar literature uses the symbol  $P$  to represent the observations of the instrument, despite the fact that generally this is a measure of energy or photon counts.

<sup>3</sup>This formulation assumes that the angle subtended by the telescope is small enough that variations in the phase function and multiple scattering can be neglected.

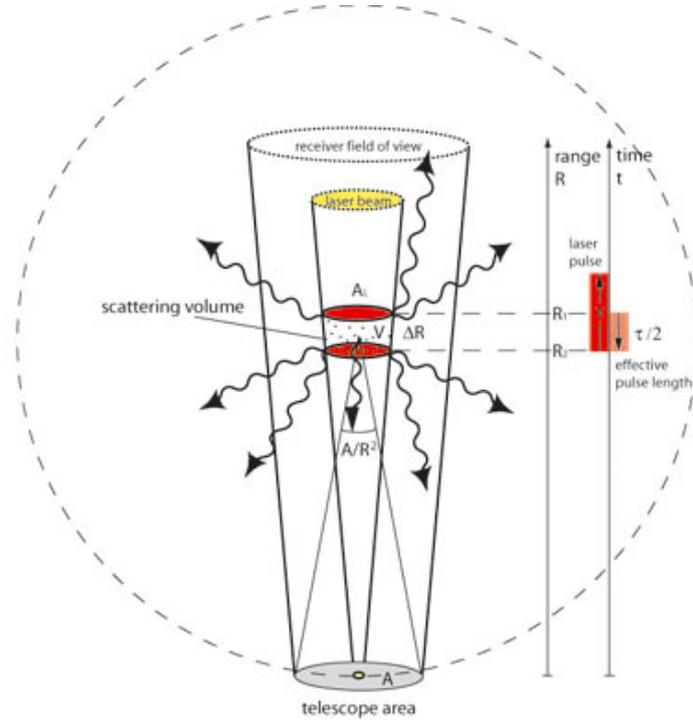


Figure 2: A schematic of lidar backscattering, showing a volume  $V$  backscattering a fraction  $\frac{V}{A_L} \beta(\lambda_0, \lambda_x, r) \frac{A(r)}{r^2}$  of the incident power into the telescope. (Image credit to the Leibniz Institute for Tropospheric Research [<http://lidar.tropos.de/en/research/lidar.html>].)

without assuming a relation between the backscatter and attenuation coefficients and the most common of these was first outlined in Klett [1981] and expanded in Klett [1985].

For a vertically-pointing ( $r \equiv z$ ), elastic backscatter lidar ( $\lambda_0 = \lambda_x \equiv \lambda$ ), equation 6 may be written,

$$z^2 P_s(\lambda, z) = P_0(\lambda) \frac{c \Delta t}{2} \eta(\lambda) A(z) \beta(\lambda, z) \exp \left[ -2 \int_0^z \alpha(\lambda, z') dz' \right], \quad (7)$$

$$\ln[z^2 P_s(z)] = S(z) = S_0 + \ln \frac{A(z)}{A(0)} + \ln \frac{\beta(\lambda, z)}{\beta_0} - 2 \int_0^z \alpha(\lambda, z') dz', \quad (8)$$

where  $S_0$  and  $\beta_0$  are constants. Differentiating with respect to  $z$  and assuming the variation of  $A(z)$  is negligible above some limiting height,

$$\frac{dS}{dz} = \frac{1}{\beta} \frac{d\beta}{dz} - 2\alpha. \quad (9)$$

The Klett method assumes that when particulate backscattering dominates over molecular scattering,

$$\beta = B^{-1} \alpha^k, \quad (10)$$

where the lidar ratio,  $B$ , and the Ångström coefficient,  $k$ , are determined em-

pirically from field studies.<sup>4</sup> As the extinction and backscatter coefficients are integrals over the size distributions of gas molecules and aerosols, this assumption is equivalent to stating that the spectral shape and composition of the scatterers are spatially invariant and any variations in backscatter are due solely to variations in scatterer density.

Using this relation, equation 9 can be written as,

$$\frac{dS}{dz} = \frac{k}{\alpha} \frac{d\alpha}{dz} - 2\alpha. \quad (11)$$

This is a Bernoulli equation which can be solved by introducing the change of variable  $w = \alpha^{-1}$  to give,

$$\alpha = \exp\left(\int \frac{1}{k} \frac{dS}{dz'} dz'\right) \left[ C - \int \frac{2}{k} \exp\left(\int \frac{1}{k} \frac{dS}{dz''} dz''\right) dz \right]^{-1}, \quad (12)$$

where  $C$  is an integration constant.

Using the obvious boundary condition at  $z = 0$  regrettably produces a solution that is not stable to numerical integration and demonstrates an exponential increase in errors with increasing range [Hughes et al., 1985; Klett, 1981]. By instead integrating from a reference range  $z_m$  to  $z$ , a stable solution is obtained,

$$\alpha(z) = \frac{\exp\left(\frac{S(z)-S(z_m)}{k}\right)}{\alpha(z_m)^{-1} + \frac{2}{k} \int_z^{z_m} \exp\left(\frac{S(z')-S(z_m)}{k}\right) dz'}. \quad (13)$$

When using equation 10, it is common to assume  $k = 1$  [Gaumet et al., 1998; Reagan et al., 1989], though the work of Klett [1985] demonstrates that for transitional regions where spectral characteristics are changing it is better to use  $k = 1.3$ .

The accurate determination of  $B$  is a great challenge for single channel lidar. Generally, standard values are used, using knowledge of the atmospheric composition from the source of the air mass. However, as shown by figure 3 of Ansmann et al. [1992], the lidar ratio can vary by up to a factor of five across an atmosphere with a variety of different scatterers. Klett [1985] attempts to avoid this difficulty by producing an empirical model for the variation of  $\beta$ , but obviously the preferred solution is to collect data from more than one channel such that  $\alpha$  and  $\beta$  can be determined separately.

Because of these difficulties, it is common in modern lidar studies of atmospheric dynamics to forgo the calculation of the fundamental backscatter or extinction and instead calculate an attenuated backscatter coefficient,

$$ABC(z) = \frac{P_s(\lambda, z)z^2}{P_0(\lambda)A(z)} \exp\left[2 \int_0^z \alpha_m(\lambda, z') dz'\right]. \quad (14)$$

By substituting from equation 6, it can be shown that,

$$ABC(z) \equiv C\beta(\lambda, z) \exp\left[-2 \int_0^z \alpha_a(\lambda, z') dz'\right],$$

where  $C$  is a constant and  $\alpha_{m/a}$  is the extinction due to molecular Rayleigh scattering or aerosols, respectively.

<sup>4</sup>This was an extension of the work of Fernald et al. [1972], adding the constant  $k$ .

## 2.4 The Ansmann solution for Raman lidar

It is then necessary to develop a method to analyse the other profiles observed by the instrument. For RACHEL, these are the Raman channels. The lidar equation for Raman scattering is written following equation 6, but separates the backscatter coefficient into a cross-section and a number density,

$$P_s(\lambda_X, r) = P_0(\lambda_0) \frac{A(r)}{r^2} \eta(\lambda_X) \frac{c\Delta t}{2} \left[ N_X(r) \left| \frac{d\sigma_R}{d\Omega} \right|_{\lambda_0, \pi} \right] \times \exp \left[ - \int_0^r \alpha(\lambda_0, r') dr' \right] \exp \left[ - \int_0^r \alpha(\lambda_X, r') dr' \right], \quad (15)$$

$$= \frac{C_X}{r^2} \sigma'_{R,X} N_X(r) \exp \left[ - \int_0^r \alpha(\lambda_0, r') + \alpha(\lambda_X, r') dr' \right], \quad (16)$$

where  $\lambda_0$  is the wavelength of the laser pulse,  $\lambda_X$  is the wavelength of the Raman scattered radiation from a species X,  $\left| \frac{d\sigma_R}{d\Omega} \right|_{\lambda, \pi} \equiv \sigma'_{R,X}$  is the differential cross-section for Raman backscatter by species X, and the range dependence of effective area has been neglected.

By taking the ratio of the Raman profiles for two species, it is possible to eliminate the absorption of the outgoing pulse,  $\alpha(\lambda_0, r)$ . Nitrogen is most frequently used as the reference signal. As a well-mixed gas, its density profile,  $N_{N_2}$ , is proportional to the density profile of the atmosphere.

Thus,

$$\frac{N_X}{N_{N_2}} = C \frac{P_s(\lambda_X, r)}{P_s(\lambda_{N_2}, r)} \exp \left[ \int_0^r \alpha(\lambda_X, z') - \alpha(\lambda_{N_2}, z') dz' \right], \quad (17)$$

where  $C$  is a constant. This formulation has the advantage that the cross-sections are included in the calibration factor, which is determined by comparison with radiosonde or other measurements, such that standard values are not required, which are only known to  $\pm 10\%$ . The exponential can be modeled with standard atmospheres or neglected, depending on the conditions [Guerrero-Rascado et al., 2008].

In addition, the nitrogen (or oxygen) Raman profile alone can be used to derive the extinction coefficient of the atmosphere. As outlined in Ansmann et al. [1990], equation 16 is rearranged as,

$$\alpha(\lambda_0, r) + \alpha(\lambda_X, r) = \frac{d}{dr} \left[ \ln \frac{N_X(r)}{P_s(\lambda_X, r)r^2} \right], \quad (18)$$

$$\alpha_a(\lambda_0, r) + \alpha_a(\lambda_X, r) = \frac{d}{dr} \left[ \ln \frac{N_X(r)}{P_s(\lambda_X, r)r^2} \right] - \alpha_m(\lambda_0, r) - \alpha_m(\lambda_X, r), \quad (19)$$

where in (19) the extinction coefficients have been separated into components due to scattering by aerosols and that by molecular Rayleigh scattering.

A wavelength dependence for the aerosol coefficients is assumed,

$$\frac{\alpha_a(\lambda_0, r)}{\alpha_a(\lambda_X, r)} = \left( \frac{\lambda_X}{\lambda_0} \right)^k, \quad (20)$$

where  $k = 1$  for particles comparable in size to  $\lambda$  and  $k = 0$  for larger particles [Ansmann et al., 1992].

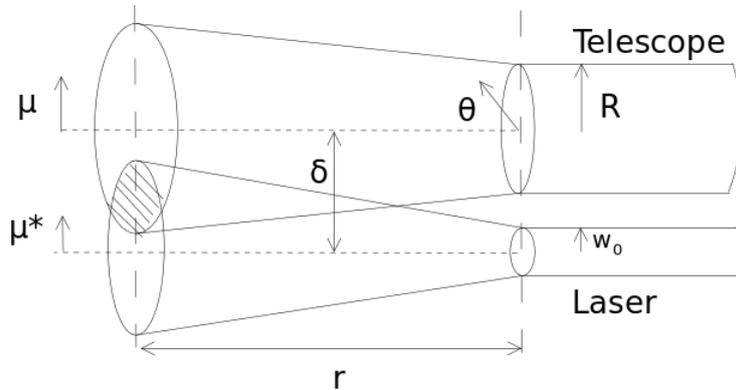


Figure 3: A schematic of the optical axes of a lidar system.

Hence,

$$\alpha_a(\lambda, r) = \frac{\frac{d}{dr} \left[ \ln \frac{N_X(r)}{P_s(\lambda_X, r)r^2} \right] - \alpha_m(\lambda, r) - \alpha_m(\lambda_X, r)}{1 + (\lambda/\lambda_X)^k}. \quad (21)$$

Using this result, the backscatter coefficient can then be determined from the elastic signal by evaluating equation 6 directly. Many current Raman lidar studies use these to determine the lidar ratio,  $B = \alpha/\beta$ , and (with additional channels) the Ångström coefficient,  $k^5$ . These are dependent on the size distribution of the aerosol and can be used to identify its origin [Mattis et al., 2008; Noh et al., 2007].

## 2.5 Model of effective area

The above methods of solution of the lidar equation neglect the effective area of the instrument as constant above some height. For work in the free troposphere, this assumption is generally found to be valid [Velotta et al., 1998]. However, when considering the planetary boundary layer (PBL) and ranges less than 2 km, it is possible that the effective area will be varying in a significant way. A knowledge of the instrument's effective area is required to avoid underestimating the true backscatter in these regions. If the optical system is well parametrised, this can be determined from ray tracing software. If not, geometrical optics can be used to determine an approximate functional form or the function can be estimated from measurements with the instrument in well-known conditions.

This section describes the method introduced in Halldórsson and Langerholc [1978] (though there are a great number of alternatives, this is one of the most generic). An area backscattering coefficient is introduced,

$$\eta = \frac{c\Delta t}{2} \beta(z) \equiv \frac{1}{I} \frac{dP_s}{d\Omega dA}, \quad (22)$$

such that the differential power scattered out of a beam of radiance  $I$  in the solid angle  $d\Omega$  by an area  $dA$  is,

$$dP_s = \eta I d\Omega dA. \quad (23)$$

<sup>5</sup>This is distinct from the constant in equation 10

A simple model of the intensity distribution of a laser in a high excitation mode assumes constant intensity across any cross-section of the circular beam,

$$I(\mu^*, z) = \frac{P_0}{\pi w^2(z)} H[w(z) - \mu^*], \quad (24)$$

where  $w(z)$  is the radius of the beam at height  $z$ ,  $\mu^*$  is the radial distance from the centre of the beam, and  $H[x]$  is the Heaviside step function.

A telescope of radius  $R$  observes this distribution. This is represented by a cylindrical coordinate system  $(\mu, \theta, z)$ , as shown in figure 3. Substituting (24) into (23),

$$dP_s = \eta \frac{P_0}{\pi w^2(z)} H[w(z) - \mu^*] \left( \frac{\pi R^2}{z^2} \right) \mu d\mu d\theta, \quad (25)$$

where  $d\Omega = \frac{\pi R^2}{z^2}$  was used.

The distance from the laser axis,  $\mu^*$ , can be expressed in terms of the coordinates from the optical axis as,

$$\mu^* = \sqrt{\mu^2 + \delta^2 - 2\delta\mu \cos \theta}, \quad (26)$$

where  $\delta$  is the perpendicular distance between the optical axes of the laser and telescope.

Combining equations 6, 22, and 23 whilst neglecting absorption ( $\alpha = 0$ ),

$$A(z) = \frac{z^2}{\eta P_0} P_s(z), \quad (27)$$

$$= \frac{R^2}{w^2(z)} \int_0^\infty \int_0^{2\pi} \psi(\mu, z) H[w(z) - \mu^*] \mu d\mu d\theta, \quad (28)$$

where  $\psi(\mu, z)$  is the fraction of scattered flux  $dP_s$  that is actually detected.

The flux that couples with the detector,  $\psi$ , is the fraction of the image spot that passes through the exit aperture. The function  $\mathcal{A}(r_1, r_2; \mu)$  is defined as the overlap area of two circles, radii  $r_1$  and  $r_2$ , with distance  $\mu$  between their centres. Then, if the aperture has radius  $a$ ,

$$\psi(r, z) = \frac{\mathcal{A}(a, R'; \mu') - \mathcal{A}(a, b'; \mu')}{\pi R'^2}, \quad (29)$$

where the second term represents the area obscured by the secondary mirror,  $R'$  and  $b'$  are the radii of image spots in the focal plane produced by the entire lens and the obstruction, respectively, and  $\mu'$  is the distance between their centres in that plane.

For a Cassegrain telescope, focal length  $f$ , it is shown in Appendix A.1 that a circular object of radius  $p$  at distance  $z$  from the optic plane of the telescope produces an image of radius,

$$p' = \frac{fp}{z}. \quad (30)$$

Hence,

$$\psi(\mu, z) = \frac{\mathcal{A}\left(a, \frac{fR}{z}, \frac{f\mu}{z}\right) - \mathcal{A}\left(a, \frac{fb}{z}, \frac{f\mu}{z}\right)}{\pi (fR/z)^2}. \quad (31)$$

It is shown in Appendix A.2 that,

$$\mathcal{A}(r_1, r_2; \mu) = \begin{cases} 0, & \mu \geq r_1 + r_2; \\ \pi \min[r_1^2, r_2^2], & \mu \leq |r_1 - r_2|; \\ r_1^2 \cos^{-1}\left(\frac{\mu^2 + r_1^2 - r_2^2}{2\mu r_1}\right) + r_2^2 \cos^{-1}\left(\frac{\mu^2 + r_2^2 - r_1^2}{2\mu r_2}\right) & \\ -\frac{1}{2}[(r_1 + r_2)^2 - \mu^2][\mu^2 - (r_1 - r_2)^2]^{1/2}, & \text{otherwise.} \end{cases} \quad (32)$$

Using the fourth-order symmetry of this expression, equation 31 is rewritten as,

$$\psi(\mu, z) = \frac{\mathcal{A}\left(\frac{az}{f}, R; \mu\right) - \mathcal{A}\left(\frac{az}{f}, b; \mu\right)}{\pi R^2}. \quad (33)$$

**Coaxial approximation:** Substituting into (28) for the perfectly aligned, coaxial case ( $\delta = 0 \Rightarrow \mu^* = \mu$ ) where  $H[w(z) - \mu^*]$  becomes solely a limit on the radial integral,

$$A(z) = \frac{1}{\pi w^2(z)} \int_0^{w(z)} \int_0^{2\pi} \mathcal{A}(\alpha, R; \mu) - \mathcal{A}(\alpha, b; \mu) \mu \, d\mu \, d\theta, \quad (34)$$

$$\equiv \frac{S_H(R) - S_H(b)}{w^2(z)}, \quad (35)$$

where  $\alpha = \frac{az}{f}$  and  $S_H(\rho) = \int_0^{w^2(z)} \mathcal{A}(\alpha, \rho; \mu) \, d\mu^2$ .

From (32), it can be shown that for  $w(z) \leq |\alpha - \rho|$ ,

$$S_H(\rho) = \int_0^{w^2(z)} \pi \min[\alpha^2, \rho^2] \, d\mu^2, \quad (36)$$

$$= \pi w^2(z) \min[\alpha^2, \rho^2]. \quad (37)$$

Noting that  $\mathcal{A}(\alpha, \rho; \mu) = 0$  for  $\mu \geq \alpha + \rho$ , integration by parts gives,

$$S_H(\rho) = \int_0^{\min[w^2(z), (\alpha + \rho)^2]} \mathcal{A}(\alpha, \rho; \mu) \, d\mu^2, \quad (38)$$

$$= [\mu^2 \mathcal{A}(\alpha, \rho; \mu)]_0^{\min[w(z), \alpha + \rho]} - \int_{\mu=0}^{\mu=\min[w(z), \alpha + \rho]} \mu^2 \, d\mathcal{A}(\alpha, \rho; \mu). \quad (39)$$

It is shown in Appendix A.3 that,

$$d\mathcal{A}(\alpha, \rho; \mu) = \begin{cases} -\sqrt{\chi(\alpha, \rho, \mu)} \frac{d\mu}{\mu}, & |\alpha - \rho| \leq \mu \leq \alpha + \rho \\ 0, & \text{otherwise,} \end{cases} \quad (40)$$

where  $\chi(\alpha, \rho, \mu) = [(\alpha + \rho)^2 - \mu^2][\mu^2 - (\alpha - \rho)^2]$ .

Substituting this into (39) for  $|\alpha - \rho| \leq w(z) \leq \alpha + \rho$ ,

$$S_H(\rho) = w^2(z) \mathcal{A}[\alpha, \rho; w(z)] + \frac{1}{2} \int_{|\alpha - \rho|^2}^{w^2(z)} \sqrt{\chi(\alpha, \rho, \mu)} \, d\mu^2. \quad (41)$$

This integral is evaluated in Appendix A.4 to give,

$$S_H(\rho) = w^2(z)\mathcal{A}[\alpha, \rho; w(z)] + (\alpha\rho)^2 \cos^{-1} \left( \frac{\alpha^2 + \rho^2 - w^2(z)}{2\alpha\rho} \right) - \frac{\alpha^2 + \rho^2 - w^2(z)}{2\alpha\rho} \sqrt{\chi[\alpha, \rho, w(z)]}. \quad (42)$$

Hence, the effective area of a lidar system with telescope of radius  $R$  and a central obstruction of radius  $b$  observing a coaxial, perfectly aligned laser beam of half-width  $w(z)$  is,

$$A(z) = \frac{S_H(R) - S_H(b)}{w^2(z)}, \quad (43)$$

where,

$$S_H(\rho) = \begin{cases} \pi w^2(z) \min[\alpha^2, \rho^2], & w(z) \leq |\alpha - \rho|; \\ \pi(\alpha\rho)^2, & w(z) \geq \alpha + \rho; \\ w^2(z)\mathcal{A}[\alpha, \rho; w(z)] + (\alpha\rho)^2 \cos^{-1} \left( \frac{\alpha^2 + \rho^2 - w^2(z)}{2\alpha\rho} \right) - \frac{\alpha^2 + \rho^2 - w^2(z)}{2\alpha\rho} \sqrt{\chi[\alpha, \rho, w(z)]}, & \text{otherwise,} \end{cases}$$

for  $\alpha = ar/f$ , where  $a$  is the radius of the exit aperture of the telescope (i.e. the diameter of the optical fibre) and  $f$  is the telescope's focal length.

**Generally:** When there is a misalignment in the axes of the laser and telescope, there is no longer an analytic expression for  $A(z)$  and one must write,

$$\delta(z) = \sqrt{(\delta + \phi_{\parallel}z)^2 + (\phi_{\perp}z)^2}, \quad (44)$$

where  $\phi_{\parallel}$  and  $\phi_{\perp}$  are the angles between the laser and telescope axes, parallel and perpendicular to the plane defined by the telescope axis and the laser pupil. It may also be necessary to include the factors  $\gamma$  and  $\nu$  outlined in Appendix A.1 to account for an aperture not positioned in the optical plane. This can be implemented in the calculations that follow through the transformations,

$$\alpha \rightarrow \frac{az}{\gamma f} \quad ; \quad R \rightarrow \frac{\nu}{\gamma} R \quad ; \quad b \rightarrow \frac{\nu}{\gamma} b. \quad (45)$$

Using (44), equation 35 can be rewritten,

$$A(z) = \frac{1}{\pi w^2(z)} \int_0^{\infty} \int_0^{2\pi} H[w(z) - \mu^*] [\mathcal{A}(\alpha, R; \mu) - \mathcal{A}(\alpha, b; \mu)] \mu d\mu d\theta. \quad (46)$$

As  $H[w(z) - \mu^*]$  represents a circle of radius  $w(z)$  centred at a radial distance  $\delta(z)$  from the optical axis and  $\mu d\mu d\theta$  represents a circle of radius  $\mu$  about that axis, the function  $\mathcal{A}$  can be used to rewrite (46) as,

$$A(z) = \frac{1}{\pi w^2(z)} \int_{\mu=0}^{\mu=w(z)+\delta(z)} [\mathcal{A}(\alpha, R; \mu) - \mathcal{A}(\alpha, b; \mu)] d\mathcal{A}[\mu, w(z); \delta(z)]. \quad (47)$$

The form of equation 35 can then be used if one defines,

$$S_H(\rho) = \frac{1}{\pi} \int_{\mu=0}^{\mu=w(z)+\delta(z)} \mathcal{A}(\alpha, \rho; \mu) d\mathcal{A}[\mu, w(z); \delta(z)]. \quad (48)$$

Integrating by parts, following the methods within (38),

$$\begin{aligned}
S_H(\rho) &= \frac{1}{\pi} \left\{ \mathcal{A}[\mu, w(z); \delta(z)] \mathcal{A}[\alpha, \rho; \mu]_0^{\min[w(z)+\delta(z), \alpha+\rho]} \right. \\
&\quad \left. - \int_{\mu=0}^{\mu=\min[w(z)+\delta(z), \alpha+\rho]} \mathcal{A}[\mu, w(z); \delta(z)] d\mathcal{A}(\alpha, \rho; \mu) \right\}, \quad (49) \\
&= w^2(z) \mathcal{A}[\alpha, \rho; w(z) + \delta(z)] \\
&\quad + \frac{1}{\pi} \int_{\mu=|\alpha-\rho|}^{\mu=\min[w(z)+\delta(z), \alpha+\rho]} \mathcal{A}[\mu, w(z); \delta(z)] \sqrt{\chi(\alpha, \rho, \mu)} \frac{d\mu}{\mu}. \quad (50)
\end{aligned}$$

This method can be adapted to any misalignment of the system through numerical integration of equation 50. Figure 4 shows examples of the overlap functions determined using the expected parameters of the RACHEL system (see section 1.2) for continuous and Gaussian beam profiles and both aligned and misaligned beams. Further, a generic beam profile may be applied to the method by replacing the  $H[x]$  function of equation 46 and evaluating both the radial and angular integrals.

In this work, the true beam profile of the laser is not known and cannot be readily measured. As such, a theoretical function will be fit to measurements of the overlap function. Over merely using the measured function, this method has the advantage of producing a function that can be applied to any range resolution exactly (such that the measured function could be derived from a reduced vertical resolution to improve the SNR). Further, the parameters of the fit should indicate the degree of misalignment in the system, giving feedback to the operator to improve the alignment.

## 2.6 Photomultiplier tubes

Virtually all existing lidar systems use photomultiplier tubes to detect the return profile. To achieve the greatest dynamic range from the elastic channel, many systems operate two detectors on that channel — one in analogue mode, where the photon flux is proportional to the voltage output, and one in photon counting mode, where the detector emits a pulse whenever a photon is detected. The latter mode can only measure very low photon fluxes (typically less than two photons per range bin), but does so with substantially less noise than the analogue mode (which has an output proportional to the tube voltage and temperature).

However, the reduced noise of the photon counting mode comes at the expense of reduced linearity of the detector as, after detecting a photon, there is a finite time,  $\tau$ , during which the detector cannot detect another photon whilst the detector ramps back to full voltage. As reviewed in Whiteman [2003, Appendix B], there are two models for this response:

- Paralyzable counters, where if an additional photon strikes the detector during the dead time, that dead time is extended a further time  $\tau$ . Assuming the number of photons striking the detector in a given time interval is described by a Poisson distribution, the measured number of photons can be written as,

$$N_{\text{meas}} = N_{\text{real}} \exp(-\tau N_{\text{real}}). \quad (51)$$

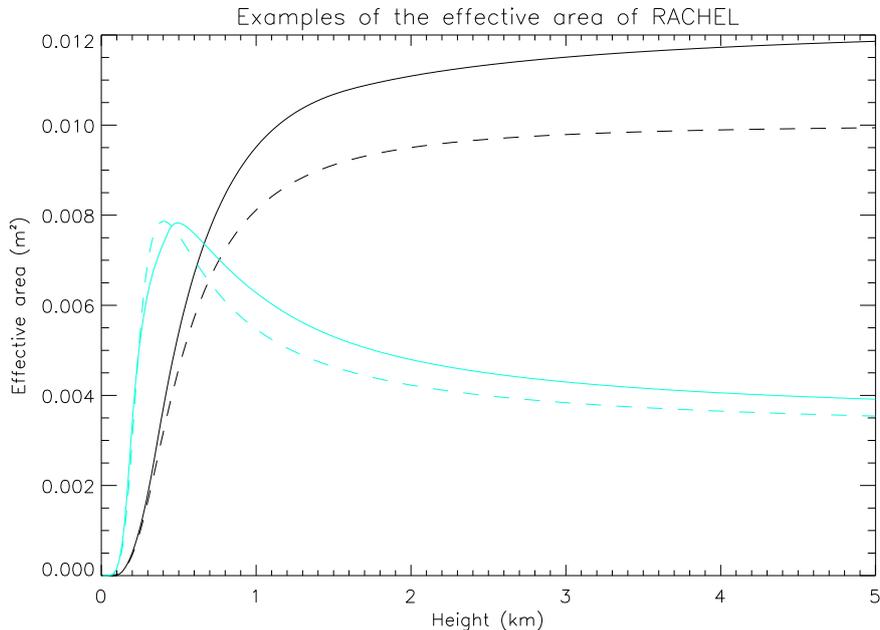


Figure 4: Examples of the Halldórsson and Langerholc [1978] model of effective area. The solid lines use a continuous beam profile and the dashed a Gaussian beam profile. The black lines show a perfectly aligned RACHEL system (see table 1) whilst the blue lines introduce misalignments (see sec. 3.2 and table 3).

- Non-paralyzable counters, where the dead time is a constant. Hence, if the detector measures a flux  $N_{\text{meas}}$ , it will be ‘dead’ for a time fraction  $\tau N_{\text{meas}}$  and it can be shown that [Whiteman, 2003],

$$\begin{aligned}
 N_{\text{meas}} &= (1 - \tau N_{\text{meas}}) N_{\text{real}}, \\
 N_{\text{real}} &= \frac{N_{\text{meas}}}{1 - \tau N_{\text{meas}}}.
 \end{aligned}
 \tag{52}$$

The actual response of the detector will lie somewhere between these two extremes, though (52) is most frequently used as it is more easily evaluated.

## 3 Field work

### 3.1 Chilbolton Observatory

The RACHEL system was given its first extensive field test at the Chilbolton Observatory, in Hampshire (51.1445° N, 1.4370° W, 84 m) [Council, 2010] from February to April of 2010. This rural site hosts a variety of continuously operating instruments, including a 905 nm Vaisala lidar ceilometer, the 35 GHz Copernicus cloud radar, a 355 nm Leosphere Raman lidar, a 1.5 micron Doppler lidar, radiometers, and surface meteorology measurements. A higher powered 355 nm Raman lidar system is also operated on a case-study basis (hereafter,

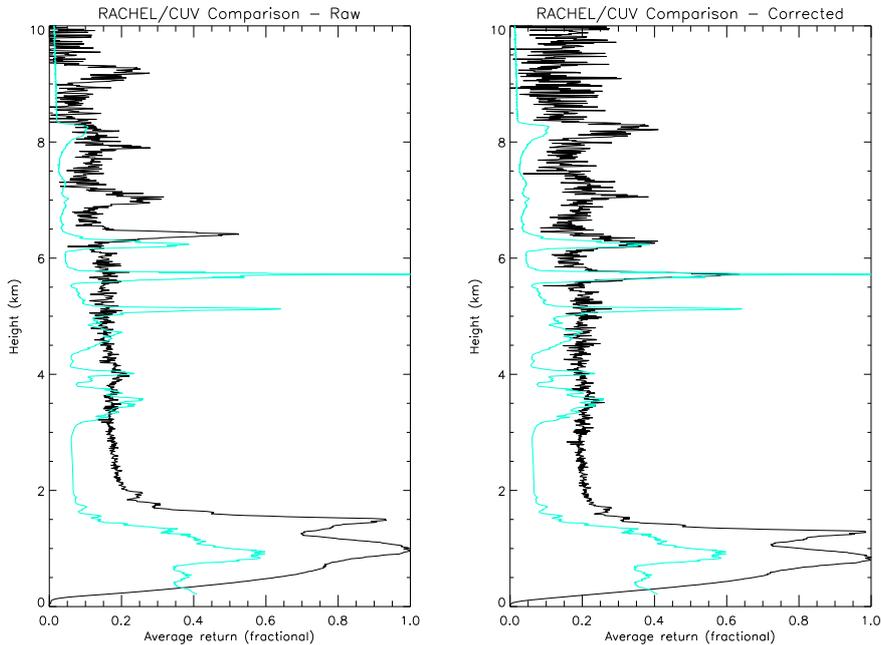


Figure 5: Example of the correction of the RACHEL system parameters, showing profiles of the attenuated backscatter coefficient averaged over three hours of the morning of Feb. 26, 2010 measured by both RACHEL (black) and the CUV (blue). The left plot shows the poor correlation between the instruments using the raw parameters and the right shows the substantially improved correlation produced by fitting of the parameters.

known as the CUV) and is currently the primary system against which the performance of RACHEL is evaluated.

Initial results from the field were promising, with RACHEL observing similar structures to the other instruments. However, more detailed analysis revealed an error in the vertical measurements of features in the RACHEL data, relative to the other instruments at the site (which were in agreement with each other). This error increased with height, being sufficiently small at low levels as to have not been obvious from the raw profiles available during the early stages of data analysis. This increase implied that the error was in the bin size of the RACHEL data, rather than a more general difference of observation.

The vertical height of bin  $i$  is given by,

$$z_i = \frac{1}{2}cd \cos \theta \times (i - i_0), \quad (53)$$

where  $d$  is the duration of the bin in seconds,  $\theta$  is the zenith angle of observation, and  $i_0$  is the subscript of the first significant bin. This offset is necessary as the data collection system is not currently triggered by the emission of the laser pulse, resulting in a delay between the initialization of data collection and the true zero of range.

During the Chilbolton test, observations were made almost exclusively vertically ( $\theta = 0$ ). This was confirmed by measurements using an electronic spirit

level to  $\pm 0.2^\circ$ . Hence, there are two unknowns in the system — the bin duration,  $d$ , and the offset,  $i_0$ . The former is intended to be specified by the control software as 67 ns, but it is possible that coding bugs and hardware timing features could introduce an error. The data acquisition system is currently undergoing testing to independently specify this error, but as a first guess, a fitting scheme was developed.

For each period of data available, an averaged attenuated backscatter coefficient profile was obtained from RACHEL, CUV, and the Vaisala ceilometer. The multidimensional simplex minimisation algorithm of Press et al. [1992, Section 10.4] was applied to minimise the root-mean square (RMS) deviation between the profile of RACHEL and another instrument<sup>6</sup>, taking the bin duration and bins offset as free parameters. The fits were then visually inspected to ensure a sensible minimum was found. From these, durations were found in the range 58 – 62 ns with offsets of 0 – 50 bins.

In an attempt to reduce the range of results, the period of averaging was reduced to coincide with specific features (such as simple cloud fields) observed in both data sets. Eventually, though, it was decided to accept the results of a fit to CUV on the morning of Feb. 26, 2010, as shown in figure 5. The period was selected as it was the only data set to display clouds at all levels of the atmosphere. The results of 60.375 ns and 7 bins were found to produce visually satisfying fits in all other days of data and to produce correlations similar to the maxima found from direct fitting. The effect of the change in parameters is shown on the right of fig. 5. These values will be compared to the results of hardware and software evaluation, currently underway.

### 3.2 Effective area

Having specified the vertical axis of the instrument, it was possible to produce an estimate of the instrument’s overlap function. Equation 15 can be rearranged to show that,

$$\begin{aligned} P_s(\lambda_x, r) &= \frac{C'_x}{r^2} A(r) \sigma'_{R,X} N_x(r) \exp \left[ - \int_0^r \alpha(\lambda_0, r') + \alpha(\lambda_x, r') dr' \right], \\ A(r) &= \frac{P_s(\lambda_x, r) r^2}{C'_x N_x(r)} \exp \left[ \int_0^r \alpha(\lambda_0, r') + \alpha(\lambda_x, r') dr' \right], \end{aligned} \quad (54)$$

where  $C'_x = \frac{1}{2} P_0(\lambda_0) \eta(\lambda_x) c \Delta t \sigma'_{R,X}$ .

Hence, if measurements are taken of a sufficiently clear sky that the extinction can be taken to be purely due to Rayleigh scattering, the effective area can be estimated from the instrument parameters and a standard atmosphere. However, it is likely that the overlap function will change each time the instrument is adjusted or aligned and the overlap function should be estimated separately for each period.

Consideration of the Chilbolton data set showed that 1 – 2 AM on Feb. 20, 2010 to be potentially the clearest skies in the first data block, based on estimates of the aerosol extinction coefficient from a simple implementation of the Klett and Ansmann algorithms. The overlap function derived from this time is

---

<sup>6</sup>As the CUV and ceilometer data products were presented after noise filtering, it was necessary to introduce a maximum height of 10 km to these fits to ensure the noise-dominated region of the RACHEL data did not bias the result.

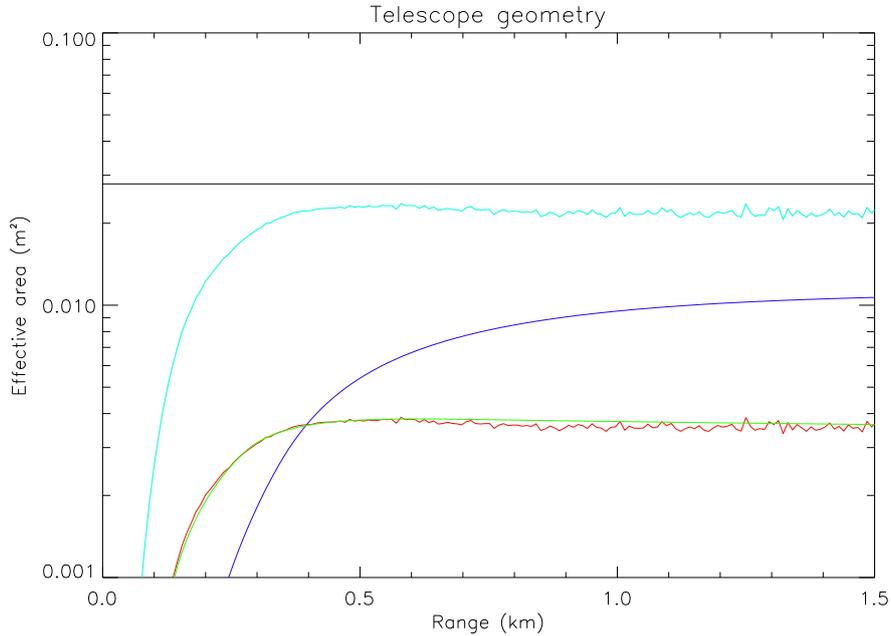


Figure 6: The variation of effective area for the RACHEL system with range. The blue line shows the effective area derived from measurements from 1 – 2 AM on Feb. 20, 2010, assuming average laser power and a worst-case optical efficiency. A naive estimate of the effective area, derived from telescope area and the beam divergences, is shown in black as an upper limit. The purple curve shows the theoretical effective area calculated from estimates of the system parameters assuming a constant beam profile. The green and red lines show the observed and theoretical effective areas resulting from an RMS fitting for the misalignment of the beam.

shown in figure 6, against the theoretical relation of equation 43. The theoretical model is calculated using estimates of the instrument parameters and a lower limit for the system’s optical efficiency. It further assumes perfect alignment. A second set of curves (yellow and red) are derived from fitting the model to the observations, allowing for small misalignments (see table 3).

While in the field, there was a short exploitation of RACHEL’s scanning capacity. On a clear, still night, measurements were collected at a range of zenith angles across the sky over the course of an hour. Using the assumption that the atmosphere is horizontally homogeneous and slowly varying over that time, each measurement observes the extinction and backscatter coefficients along a different slant path but superimposed on the same overlap function (being a function of the absolute range from the instrument). The best method to determine the overlap function from this data is currently under consideration.

### 3.3 Data processing

Having determined the overlap function and specified the instrument’s height axis, it is possible to begin data processing. For the purpose of comparison with

Property	Value
$z^*$	0.0101 m
$\delta$	-0.0177 m
$\phi_{\parallel}$	$-5.73 \times 10^{-3}$ rad
$\phi_{\perp}$	$6.75 \times 10^{-3}$ rad
$\eta(\lambda_x)$	0.0528

Table 3: Parameters of the fitting of the model equation 43 to the measured overlap function, as shown by the green and red curves of figure 6.

the data collected by the Chilbolton Observatory, the attenuated backscatter coefficient is determined as per equation 14.

At the most basic level, the data acquisition system forms a discrete return profile by sampling the PMTs at a user-defined rate (roughly 60 ns per bin or 17 MHz, corresponding to 9 m raw resolution). The system then sums together the profiles from 200 successive laser shots to produce a single average profile which is then saved as an ASCII file.

Beginning offline data analysis, the background photon count is subtracted from the signal, estimated by the average of the last 100 bins of the data. The maximum height at which data can then be considered reliable (or the MRH) is determined from the signal-to-noise ratio. The magnitude of the noise is estimated as the standard deviation of the top 20 km of data<sup>7</sup>. A signal is considered significant if it's magnitude is at least three times that of the noise (around the 95% confidence level). An example of this height is overplotted on figure 8, demonstrating a typical day-night cycle with MRH varying from 2 – 7 km for the elastic channel. The Chilbolton group applies a noise filter to their data which neglects all measurements below this  $3\sigma$  threshold, producing ‘smoother’ profiles with height than RACHEL.

Next, the transmission of the atmosphere due to Rayleigh scattering is estimated. This currently uses the US Standard Atmosphere [Office, 1976], which assumes a constant lapse rate of  $-6.5$  K  $\text{km}^{-1}$  across the lowest 11 km of the atmosphere, though this will eventually be replaced or enhanced with radiosonde measurements. Under these conditions, for a vertical beam,

$$\mathcal{T}(\lambda_0, z) = \exp \left[ - \int_0^z \sigma(\lambda_0) N(z') dz' \right], \quad (55)$$

where  $\sigma$  is the Rayleigh scattering cross-section (see eqn. 4) and  $N(z)$  is the number density of air. For an ideal gas with molar mass  $M$  described by pressure  $p$ , temperature  $T$ , and density  $\rho$ ,

$$\int_0^z \sigma(\lambda_0) N(z') dz' = \frac{\sigma(\lambda_0)}{M} \int_0^z N_A \rho(z') dz', \quad (56)$$

$$= \frac{\sigma(\lambda_0) N_A}{M} \int_{p(0)}^{p(z)} - \frac{dp'}{g}, \quad (57)$$

$$= \frac{\sigma(\lambda_0) N_A}{Mg} [p(0) - p(z)]. \quad (58)$$

<sup>7</sup>It has been recently noticed that a more appropriate estimate would be to use a Poisson distribution for the noise.

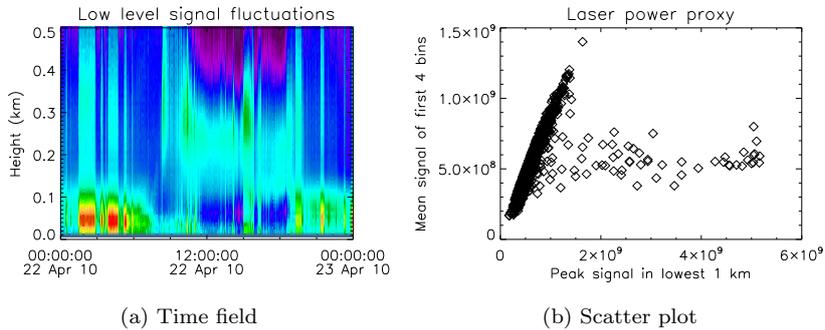


Figure 7: Illustration of the low level proxy for laser power. The field shows the range and overlap-corrected elastic backscatter. At the very lowest levels, where the overlap function is strongly varying, the signal appears to vary in an unphysical manner, most likely related to laser power fluctuations. The scatter plot shows the correlation between the peak value of the range and overlap-corrected elastic backscatter and the mean value of the return in the first four bins.

Assuming hydrostatic balance for an atmosphere with constant lapse rate  $\Gamma$ ,

$$\int_0^z \sigma(\lambda_0) N(z') dz' = \frac{\sigma(\lambda_0) N_A}{Mg} p(0) \left[ 1 - \left( 1 - \frac{\Gamma z}{T(0)} \right)^{\frac{gM}{R\Gamma}} \right], \quad (59)$$

where  $N_A$  is Avagadro's constant,  $g$  is the gravitational acceleration, and  $R$  is the molar gas constant.

A current failing of the RACHEL system is the lack of any monitoring of the laser power. The manufacturer's estimate of average power has been used in any work in this study, but shot-to-shot power will vary both randomly in time about this value (which should be less important due to averaging over multiple shots) and as a function of the laser head temperature. During the work, it was observed that the laser would randomly switch between a high and low powered mode of operation. Hence, some proxy for the laser power is needed from within the signal itself.

An obvious choice would be the peak value of each profile. However, the magnitude of the peak is highly dependent on the properties of the atmosphere (for example, the peak of a profile often corresponds to a cloud rather than the 'natural' peak of  $A(r)/r^2$ ). It was observed that, when considering range and overlap-corrected elastic backscatter profiles (using a measured overlap function), the lowest levels exhibited unusual behaviour. Though these could be caused by the surface boundary layer, the variations of these signals in time are not consistent with natural variations, as shown in figure 7a. Further, it was found that the mean signal from the lowest few bins of the profile was strongly correlated with the peak value of each profile (fig. 7b). Hence, ABC plots are normalised by the mean value of the first four bins of each profile<sup>8</sup>.

Following from section 2.6, the dead time of the detector of RACHEL's elastic channel is given by the manufacturer as 10 ns, such that photon fluxes of

<sup>8</sup>For clarity, these are the first four bins after the seven neglected from each profile, as described in section 3.1

$10^7 \text{ s}^{-1}$  will produce a 10% deviation from a linear response (known as the count linearity). Our studies generally use 60 ns bin durations such that the count linearity is  $0.6 \text{ counts bin}^{-1}$ . Hence, it is fairly important to correct for the detector’s response. However, as the dead time has not yet been measured independently to sufficient accuracy, no correction has been applied to any results presented in this report.

The dead time could be determined in the field using the methods described in Whiteman [2003]. Here, one acquires two profiles in quick succession — once with a neutral density filter attenuating the elastic channel and once without. The nonparalyzable correction (eqn. 52) is then applied to both profiles and the dead time is taken as the value for which the ratio between the signals is most constant with range. This has not been performed as the optical system should not be opened in the field to prevent irreparable soiling of the filters. In addition, changing a neutral density filter would take tens of minutes, over which time the atmosphere is likely to evolve in a significant way. The impact of installing a filter wheel into the optical fibre to facilitate such measurements is under evaluation.

In addition, to extend the dynamic range of the elastic channel, the detector uses a prescaler to reduce the count rate by a factor of four. This will introduce an uncertainty into the time measurement as the prescaler is not time-gated (i.e. if three photons are observed in a particular bin and one in the next, the detector will produce a single pulse in the second bin). However, the low fluxes that are most affected by this error should mostly be observed above the MRH and would be neglected by the SNR constraint.

Finally, the detector will have a certain (generally) constant probability of reacting to an incident photon, known as the quantum efficiency. The data sheet gives the ‘count sensitivity’ (the number of counts produced per second by the detector per picowatt incident from a standard lamp) as roughly  $1.5 \times 10^5 \text{ counts s}^{-1} \text{ pW}^{-1}$  at 350 nm. Knowing that the energy of a photon is  $hc/\lambda$ , it can be shown that this corresponds to a quantum efficiency of  $0.085 \text{ counts photon}^{-1}$ . This is considered part of the efficiency of the instrument ( $\eta$  in equation 6), but its variation with wavelength may need to be remembered for the Raman channels.

### 3.4 The Eyjafjallajökull eruption

The RACHEL system remained at the Chilbolton Observatory during the volcanic eruption of April 14th, 2010 and was operated from the 15th to 25th to monitor the appearance and evolution of the Eyjafjallajökull ash cloud over southern England. The data is summarised in figure 8. A detailed analysis will be performed once the optimal estimation scheme is completed. For this report, a brief qualitative review of the data is given.

Two distinct periods can be observed. The ash cloud first appears after midday on April 16th as a curved feature at 3 km that descends towards the planetary boundary layer (PBL), highlighted in figure 9a. The feature was identified as the volcanic ash cloud due to the large depolarisation ratio exhibited by the returns recorded by the EZ-Lidar operated on the site by Reading University (data not currently available). A large backscattered depolarisation ratio is a signature of highly unspherical particles, such as would be expected from ash.

The ash is then observed within the mixed layer of the PBL. Higher than

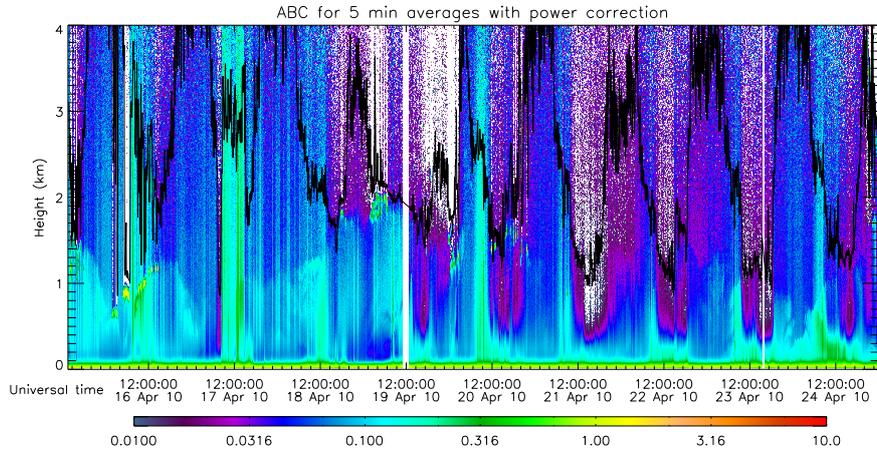


Figure 8: The effects of the volcanic ash plume over the Chilbolton Observatory as seen by RACHEL, presented as ABC on an arbitrary scale. The height at which the SNR first falls below a value of 3, which can be taken as the maximum height at which the data remains reliable, is overplotted in black.

usual PBL heights of 1.5 – 2 km (rather than 0.5 – 1.5 km) persist after the cloud passes, as is clear in fig. 8 by comparing the region before April 19th to that after<sup>9</sup>. In particular, the increased aerosol concentration allows the instrument to resolve the formation of layers in the night-time stratified mixed layer (figure 9b). In addition, the PBL exhibits an unusually large depolarisation ratio during this time. After the passing of the frontal system, the PBL returns to a more typical cycle.

### 3.5 Water vapour

There are three functional channels in RACHEL<sup>10</sup>. The plots of the previous parts of this section were produced exclusively from the elastic backscatter. These could be enhanced through use of the nitrogen Raman channel and a standard atmosphere to estimate the extinction coefficient, though this has not yet been implemented as the forward model (see section 4) will perform this function.

The remaining channel measures the Raman backscattering of water vapour. Though the cross-section for Raman scattering is of a similar order of magnitude for most species found in the atmosphere, water vapour exists in much lower concentrations than nitrogen, making its return several orders smaller than the nitrogen return. Thus, whilst the elastic and nitrogen channels can be used to probe the boundary layer over 10 s averages and up to about 7 km over 5 minutes averages, the water vapour channel is dominated by noise by 2 km over any physically useful averaging period. To improve the height coverage, vertical

<sup>9</sup>Around midday on April 19th, a weak frontal system passed the site and though no precipitation fell, the change of behaviour indicates a change to an air mass with significantly less ash.

<sup>10</sup>The carbon dioxide signal requires hours of integration to be significant and so is not appropriate for the methods outlined in this report.

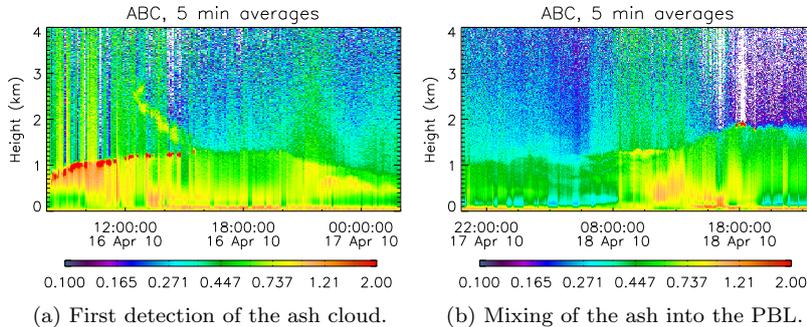


Figure 9: Details of fig. 8, showing the appearance and mixing of the ash with the PBL. Once mixed, the ash acts as a tracer revealing the formation of a stable, stratified mixed layer overnight that is again mixed with the formation of the convective mixed layer in the morning.

smoothing may also be implemented. Here a simple sliding window average over 10 bins (90 m) is used, though it is intended to evaluate alternatives, such as Gaussian or diffusive smoothing.

Fortunately, there can be significant variations in water vapour concentrations in the PBL and so the available signal is still useful. Following equation 17,

$$\frac{N_{H_2O}}{N_{N_2}} = C \frac{P_s(\lambda_{H_2O}, r)}{P_s(\lambda_{N_2}, r)}, \quad (60)$$

where the exponential has been neglected as negligible for a first approximation.

Figure 10 shows a comparison of measurements of the water vapour mixing ratio made simultaneously with the CUV and RACHEL, revealing an impressive level of agreement between the instruments. The CUV measurements are calibrated against radiosondes launched from Larkhill, approximately 30 km west of the Chilbolton site<sup>11</sup>. Although potential exists for differences in conditions between the two sites, monitoring of the calibration whenever lidar measurements are made reduces the effect of these and a reliable calibration factor can be determined [Leblanc and McDermid, 2008]. The RACHEL measurements are presented uncalibrated whilst appropriate methods are evaluated.

## 4 Data analysis

### 4.1 Optimal estimation

The eventual aim of this project is to develop an optimal estimation retrieval scheme for lidar measurements. Optimal estimation methods have been applied to satellite retrievals successfully for several years, but have not been widely applied to lidar systems (Samokhvalov [1979] and Volkov et al. [2002] are examples of linearised schemes). Considering the ad-hoc nature of the existing

<sup>11</sup>This is ‘calibration’ in the sense used by the lidar community to describe determining the constant of proportionality in a solution, such as  $C$  in equation 17, rather than an absolute calibration of the response of the system to known conditions.

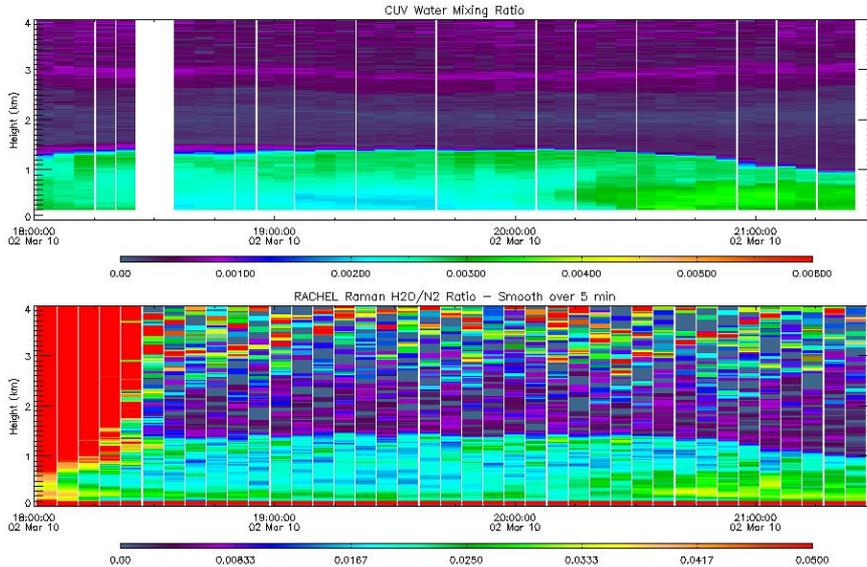


Figure 10: A comparison of the water vapour mass mixing ratio measured by the CUV and RACHEL. CUV data are calibrated against radiosonde measurements, whilst RACHEL data are presented on an arbitrary scale. Both data sets are presented as five minutes averages, with RACHEL data additionally being smoothed with a 90 m sliding-window average.

retrieval methods, the potential gain from a successful implementation could be substantial.

This work uses the methods outlined in Rodgers [2000, Chapter 7]. This defines an inverse problem,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \boldsymbol{\epsilon}, \quad (61)$$

which relates the measurement vector,  $\mathbf{y}$ , to the state vector,  $\mathbf{x}$ , through the forward model,  $\mathbf{F}$ , with an error,  $\boldsymbol{\epsilon}$ . For a lidar, the measurement vector describes the number of photons observed in each height bin for each channel. The state vector describes the unknown state of the atmosphere, here summarised by the extinction and backscatter coefficients. The forward model is a non-linear function which describes the response of the lidar to a given state vector.

The solution,  $\hat{\mathbf{x}}$ , with the maximal probability of representing the true state of the atmosphere, given the measurements (known as the maximum *a posteriori* solution) is sought. This can be found by using Bayesian statistics and assuming the probability density functions describing all quantities are Gaussian (or at least symmetric). In such circumstances,

$$-2 \ln P(\mathbf{y}|\mathbf{x}) = [\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})] + c_1, \quad (62)$$

where  $c_1$  is a constant,  $\mathbf{S}_\epsilon^{-1} = \langle [\mathbf{y} - \mathbf{F}(\mathbf{x})][\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \rangle$  is the error covariance of the measurement and  $P(\mathbf{x})$  means the probability of observing a state  $\mathbf{x}$ . The prior knowledge of the system is summarised by an *a priori* state vector,  $\mathbf{x}_a$ , which, though not necessarily realistic, is also modelled with a symmetric pdf such that,

$$-2 \ln P(\mathbf{x}) = [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a] + c_2, \quad (63)$$

where  $\mathbf{S}_a^{-1}$  is the covariance matrix of the *a priori*.

Substituting these into Bayes' Theorem,

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}, \quad (64)$$

gives the cost function,

$$-2 \ln P(\mathbf{x}|\mathbf{y}) = [\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})] + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a] + c_3, \quad (65)$$

where it has been assumed that  $P(\mathbf{y})$  is a normalising constant.

The maximum probability state,  $\hat{\mathbf{x}}$ , is found by equating the derivative of (65) to zero,

$$0 = \nabla_{\mathbf{x}} \{-2 \ln [P(\mathbf{x}|\mathbf{y})]\} = -[\nabla_{\mathbf{x}} \mathbf{F}(\mathbf{x})]^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})] + \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a]. \quad (66)$$

Defining  $\mathbf{K}(\mathbf{x}) = \nabla_{\mathbf{x}} \mathbf{F}(\mathbf{x})$ ,

$$-\mathbf{K}^T(\hat{\mathbf{x}}) \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})] + \mathbf{S}_a^{-1} [\hat{\mathbf{x}} - \mathbf{x}_a] = 0. \quad (67)$$

For a quasi-linear problem, a straightforward Newton iteration can be used. It states that for the general vector equation  $\mathbf{g}(\mathbf{x}) = 0$ , the solution can be found by the iteration,

$$\mathbf{x}_{i+1} = \mathbf{x}_i - [\nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}_i)]^{-1} \mathbf{g}(\mathbf{x}_i). \quad (68)$$

Hence, to solve equation 67,

$$\mathbf{g} = -\mathbf{K}^T(\hat{\mathbf{x}}) \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})] + \mathbf{S}_a^{-1} [\hat{\mathbf{x}} - \mathbf{x}_a], \quad (69)$$

$$\nabla_{\mathbf{x}} \mathbf{g} = \mathbf{S}_a^{-1} + \mathbf{K}^T \mathbf{S}_\epsilon^{-1} \mathbf{K} - [\nabla_{\mathbf{x}} \mathbf{K}^T] \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})]. \quad (70)$$

For the quasi-linear case the last term of (70) is neglected, giving the Gauss-Newton iteration,

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i)^{-1} [\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}_i)) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a)]. \quad (71)$$

However, as this forward model is not necessarily quasi-linear, the last term of (70) is instead approximated by  $\gamma_i \mathbf{S}_a^{-1}$ , where  $\gamma_i$  is chosen at each step to reduce the cost function (65). This is known as the Levenberg-Marquardt method,

$$\mathbf{x}_{i+1} = \mathbf{x}_i + [(1 + \gamma_i) \mathbf{S}_a^{-1} + \mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_i]^{-1} [\mathbf{K}_i^T \mathbf{S}_\epsilon^{-1} (\mathbf{y} - \mathbf{F}(\mathbf{x}_i)) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a)]. \quad (72)$$

General practice is that after an iteration, if the cost has increased,  $\gamma_i$  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.

## 4.2 Forward model

A first attempt at a forward model has been produced, summarising the most fundamental aspects of a lidar return. In order to reduce the length of the state vector of the problem at this stage, the vertical bin size has been increased by a factor of ten ( $\sim 90$  m) and limited to a maximum height of 10 km, giving a state

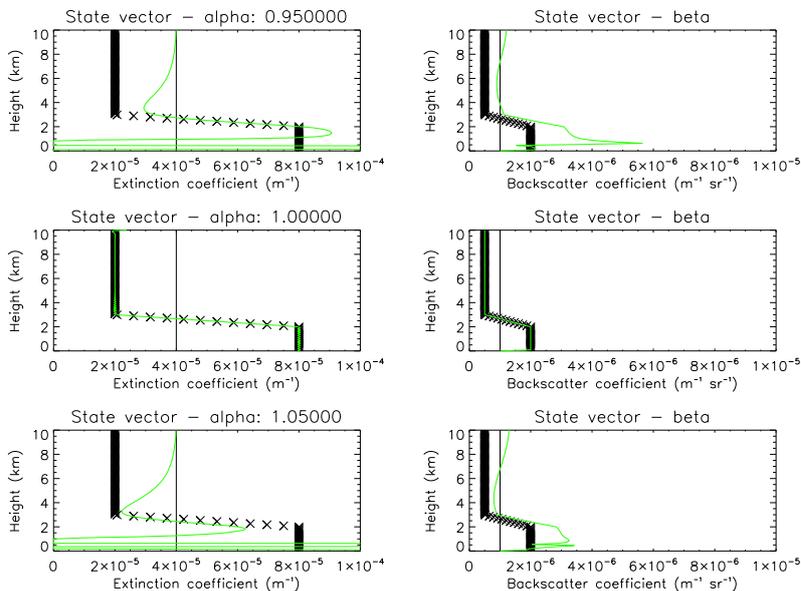


Figure 11: Examples of the retrieved profiles from the simple forward model. The left column plots the extinction coefficient and the right the backscatter coefficient. The ‘true’ answer used to produce the data are plotted in crosses and the *a priori* estimate is shown as a black line. The result of the retrieval is shown in green. The rows show retrievals produced using a system model where the parameter  $z^*$  is deviated from its true value by  $-5\%$ ,  $0\%$ , and  $+5\%$ , respectively.

vector of about 200 elements — half the aerosol extinction of each bin and the remainder the aerosol backscatter. From this, the model determines the elastic and nitrogen Raman return profiles in photon counts.

The model is split into two components — the dependence of the return profile on the parameters of the instrument and that on the state of the atmosphere. The first is summarised by, for a profile averaged over  $m$  shots,

$$v(\lambda_{N_2}, z_i) = \frac{E_L(\lambda_0)m}{hc/\lambda_0} \eta(\lambda_0, \lambda_{N_2}) \frac{c\Delta t}{2} \frac{A(z_i)}{z_i^2} s \times \exp \left[ - \int_0^{z_i} \alpha_m(\lambda_0, z') + \alpha_m(\lambda_{N_2}, z') dz' \right]. \quad (73)$$

The average pulse energy,  $E_L$ , pulse width,  $\Delta t$ , and the system channel efficiencies,  $\eta$ , are outlined in table 1. The parameters determining the effective area are determined from a fit of the theoretical model to its measured form (table 3). A scale factor  $s$  is included to account for a prescaler in the elastic channel reducing counts by a factor of four.

Further, as outlined in sections 2.1 and 3.3 (specifically equation 59), the US standard atmosphere is used with measurements of surface pressure and temperature<sup>12</sup> to determine the Rayleigh extinction and backscatter ( $\alpha_m$  and

<sup>12</sup>These measurements are first processed using a Kalman smoothing filter.

$\beta_m$ ). Radiosonde measurements were also collected from a site in the vicinity of Chilbolton and will eventually be integrated into the model to better represent the atmosphere. As nitrogen is a well-mixed gas, this also gives the backscatter coefficient of the Raman channel,

$$\beta_{N_2} = N_{N_2}(z_i) \left| \frac{d\sigma_R}{d\Omega} \right|_{\lambda_0, \pi}. \quad (74)$$

The temperature dependence of the cross-section is neglected as small relative to the uncertainty in its known value (being  $1 - 2 \times 10^{-34} \text{m}^2$ , adjusted from Fenner et al. [1973] assuming a  $\lambda^4$  dependence).

Then, the aerosol transmission is calculated from the aerosol extinction coefficient,  $\alpha_a$ , using a simple trapezium rule integration scheme,

$$\mathcal{T}_a(\lambda, z_i) = \exp \left\{ -\zeta w \left[ \frac{1}{2} [\alpha_a(\lambda, z_i) + \alpha_a(\lambda, 0)] + \sum_{j=1}^{i-1} \alpha_a(\lambda, z_j) \right] \right\}, \quad (75)$$

where  $\zeta = 1 + \lambda/\lambda_0$  and  $w$  is the length of a bin.

These steps are then combined with the aerosol backscatter coefficient,  $\beta_a$ , to give the observed profiles,

$$P_s(\lambda_0, z_i) = v(\lambda_0, z_i) \mathcal{T}_a(\lambda_0, z_i) [\beta_m(z_i) + \beta_a(z_i)], \quad (76)$$

$$P_s(\lambda_{N_2}, z_i) = v(\lambda_{N_2}, z_i) \mathcal{T}_a(\lambda_{N_2}, z_i) \beta_{N_2}(z_i). \quad (77)$$

The nonlinearity of the detectors is then modeled with the transform,

$$P_{\text{obs}} = \frac{P_s}{1 + \frac{P_s \tau}{mw}}, \quad (78)$$

where  $w$  is the integration time used with the PMTs.

Due to the relative simplicity of these equations, analytical derivatives are also calculated by the forward model to give the matrix  $\mathbf{K}(\mathbf{x})$ . The measurement errors are currently assumed to be uncorrelated and the a priori is taken as the scattering expected from a purely Rayleigh scattering atmosphere. These will be improved by processing CUV data from several weeks using either the Klett or Ansmann algorithm to obtain an estimate of the average extinction and backscatter profiles and their variance. The data will be taken from the same season as the data to be analysed and separate day and night profiles may be derived.

It is hoped that all the model parameters except the laser pulse energy, duration, and PMT dead time can be determined from fitting a theoretical function to measurement of the overlap function from off-zenith observations collected in the field. Inaccuracy in the known value of the Raman cross-section will hopefully be compensated by retrieval of the efficiency of the system in this manner.

Once a more efficient method of replacing the neutral density filter can be found, the methods outlined in section 3.3 can be used to determine the PMT dead time. The feasibility of adding a power meter to the optical system to monitor variations in the pulse energy is being investigated. For the data currently collected, the effectiveness of using the mean signal in the first few bins as a proxy for laser power will be investigated further.

Retrievals with this simple model are under evaluation. Using hypothetical observations produced with the forward model, the retrieval converges to the correct answer in about 19 iterations, indicating the convergence criteria are too rigorous. The threshold for convergence is currently the floating-point precision of the computer used, but to improve efficiency this will be updated to instead consider the error of each term. Introducing small errors into the model parameters for the retrieval do not necessarily cause the answer to diverge, but do produce unphysical features that need to be addressed, as highlighted in figure 11.

## 5 Conclusions

Lidar is a powerful tool for atmospheric profiling, presenting unprecedented temporal and spatial resolution within the troposphere. The response of a lidar is summarised by the lidar equation, (6), describing the optical properties of the atmosphere in terms of volume extinction and backscattering coefficients. As first outlined in Klett [1981], by assuming a power law relationship between these quantities, the lidar equation may be solved for a single-channel lidar system to estimate the extinction of the atmosphere.

RACHEL is a new Raman lidar system, designed to observe the inelastic Raman backscatter of a UV laser beam by nitrogen and water vapour with the unique capacity to scan the entire sky under automated operation. The Raman scattering from nitrogen allows, as first described in Ansmann et al. [1990], the calibration of the elastically backscattered return by giving an independent measurement of the extinction of the atmosphere, using of a standard atmosphere and/or radiosonde measurements to estimate the density profile of nitrogen.

An initial field campaign was performed with the RACHEL system at the Chilbolton Observatory in Hampshire, where measurements could be compared against the plethora of instrumentation operated at the site. This revealed an inaccuracy in the height binning of the RACHEL acquisition system. Whilst we await assessment to exactly specify the nature of the problem, existing data have been fit against other instruments on site to determine an appropriate bin length. The data collected was also used to estimate the instrument's overlap function, which was fit to a theoretical model to investigate the degree of misalignment in the system. Off-zenith measurements were also collected, which may be used to provide a further estimate of the overlap function without assumptions about the state of the atmosphere.

Despite being fundamentally governed by the extinction and backscatter of the atmosphere, these quantities are not necessarily determined directly from lidar returns due to the underconstrained nature of the retrieval problem. Retrieval schemes such as those mentioned above have proven successful over the years in presenting qualitative measurements of the changes in atmospheric properties and observing the evolution of turbulent motions. However, current methods fail to present rigorously derived quantitative parameterisation of these features in all circumstances. As such, an optimal estimation scheme for lidar is being developed. A forward model is in the early stages of preparation, specifying the most basic elements of the lidar as summarised by the lidar equation (eqn. 6). The model parameters will mostly be determined from fits of a theoretical overlap function to observations, though further work is needed

Year two	Planned tasks
First quarter	Installation of RACHEL in Oxford and initial vertical measurements to ensure sensible operation; Development of off-zenith measurements into a retrieval scheme for the overlap function; Repair of weather station and investigation of other calibration measurement sources
Second quarter	Initial observations of Oxford air mass through off-zenith measurements; Integration of calibration measurements into forward model; Development of forward model against simulated data
Third quarter	Intensive measurement campaign during clearer summer weather; Application of forward model to measurements; Analysis of data with other algorithms for comparison
Fourth quarter	Final Oxford measurements; Planning for validation measurements and future field work
Year three	Planned tasks
First quarter	Installation of Rachel at Chilbolton Observatory for validation measurements; Application of retrieval methods to extensive Chilbolton data set to demonstrate usage on multiple instrument platforms; Discussions with Hovemere Ltd. about future of Rachel instrument and Oxford's involvement with it
Second quarter	Analysis of Chilbolton data for validation against radar measurements and investigation of science questions (aerosol loadings in vicinity of clouds); Comparison to results of Oxford campaign; Feedback to Hovemere Ltd. for potential next generation of Rachel instrument
Third quarter	Potential additional field campaign in rural Oxford (University field site) for further validation of observations; Transfer of knowledge of retrieval algorithm to Hovemere Ltd. for usage in commercial applications; Thesis writing
Fourth quarter	Thesis writing

Table 4: Plan of work for years two and three.

to determine the accuracy to which the system's efficiency can be determined from these measurements and to monitor variations in laser pulse energy. Various vertical and temporal smoothing schemes will also need to be evaluated to reduce the length of the state vector to limit processing time to achievable levels.

Over the next year, it is hoped to develop a more robust forward model and use this to retrieve verifiable, quantitative measurements of the backscattering of the atmosphere. This will be initially validated against the instruments at the Chilbolton Observatory. The method can then hopefully be applied to two weeks of measurements collected after the Eyjafjallajökull eruption of April 2010 in order to investigate the structure of the plume, its interaction with the PBL, and, potentially, investigate the properties of the ash itself using measurements at a range of wavelengths. Once repairs are complete, the RACHEL system will be deployed in central Oxford, where we intend to complete a year-long campaign

investigating the distribution and evolution of pollution in urban and rural environments, as both can potentially be observed from this location. Looking further into the future, the data collected over the next year in Oxford could be validated by another deployment of RACHEL at the Chilbolton Observatory and, possibly, at a field site on the outskirts of Oxford. At that stage, Hovemere may be beginning development of the next version of RACHEL and it is hoped that the results of the field campaigns could be input into this process.

## A Derivation of effective area

### A.1 Telescope optics

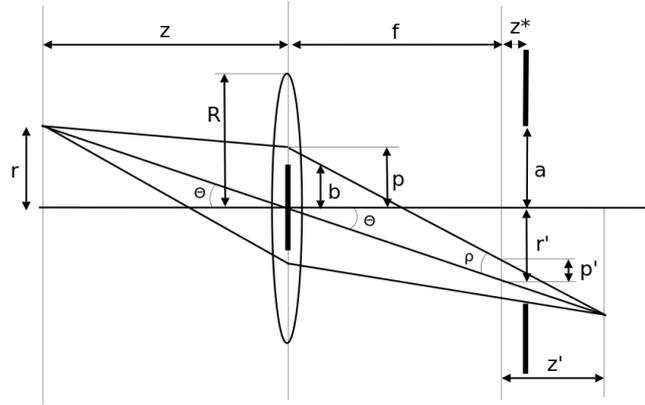


Figure 12: A Cassegrain telescope modeled by a lens, radius  $R$  and focal length  $f$ , with a central obstruction, radius  $b$ . An object at distance  $z$  from the lens plane and radial distance  $r$  from the optical axis is imaged through an aperture, radius  $a$ , in a plane a distance  $z^*$  from the focal plane. The image is focused a distance  $z'$  from the focal plane and radial distance  $r'$  from the optical axis. It produces a spot of radius  $p$  on the lens and a spot of radius  $p'$  in the image plane.

A Cassegrain telescope can be modeled by the lens system of figure 12. The radii of the lens, central obstruction, and exit aperture are taken as the radii of the telescope's primary and secondary mirrors and the optical fibre, respectively. By considering the central ray of the lens, subtending angle  $\Theta$ , an expression for  $r'$  can be derived,

$$\tan \Theta = \frac{r}{z} = \frac{r'}{f + z^*} \quad \Rightarrow \quad r' = \frac{(f + z^*)r}{z} = \gamma \frac{fr}{z}, \quad (79)$$

where  $\gamma = 1 + z^*/f$ .

A similar expression for  $p'$  can be found by considering similar triangles about the angle  $\rho$ ,

$$\frac{p'}{p} = \frac{|z' - z^*| \cos \Theta}{(f + z') \cos \Theta} = \left| \frac{z' - z^*}{f + z'} \right|. \quad (80)$$

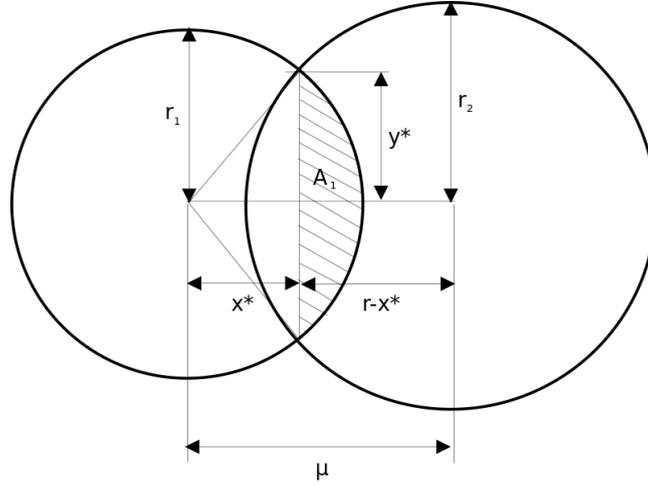


Figure 13: The overlap area of two circles, radii  $r_1$  and  $r_2$ , with a distance  $\mu$  between their centres.

However, by the lens maker's formula,

$$\frac{1}{f} = \frac{1}{z} + \frac{1}{f + z'}, \quad (81)$$

$$f + z' = \frac{fz}{z - f}, \quad (82)$$

$$z' = \frac{f^2}{z - f}. \quad (83)$$

Substituting this into (80),

$$p' = p \left( \frac{f^2}{z - f} - z^* \right) \bigg/ \frac{fz}{z - f}, \quad (84)$$

$$= \frac{fp}{z} \left| 1 + \frac{z^*}{f} - \frac{z^*z}{f^2} \right|, \quad (85)$$

$$= \nu \frac{fp}{z}, \quad (86)$$

where  $\nu = \left| \gamma - \frac{z^*z}{f^2} \right|$ .

When the aperture is in the focal plane,  $\gamma = \nu = 1$ .

## A.2 Overlap area of two circles

A brief consideration is given to the form of  $\mathcal{A}(r_1, r_2; \mu)$ . In the event the circles completely overlap or do not overlap at all, its evaluation is trivially,

$$\mathcal{A}(r_1, r_2; \mu) = \begin{cases} 0, & \mu \geq r_1 + r_2; \\ \pi \min[r_1^2, r_2^2], & \mu \leq |r_1 - r_2|. \end{cases} \quad (87)$$

Otherwise, consider figure 13. The shaded area  $A_1$  is obtained by subtracting the area of the triangle from that of the sector,

$$A_1 = r_1^2 \cos^{-1} \left( \frac{x^*}{r_1} \right) - x^* y^*. \quad (88)$$

Similarly,

$$A_2 = r_2^2 \cos^{-1} \left( \frac{r - x^*}{r_2} \right) - (r - x^*) y^*. \quad (89)$$

Adding these,

$$\mathcal{A}(r_1, r_2; \mu) = r_1^2 \cos^{-1} \left( \frac{x^*}{r_1} \right) + r_2^2 \cos^{-1} \left( \frac{\mu - x^*}{r_2} \right) - r y^*. \quad (90)$$

The values of  $x^*$  and  $y^*$  are derived by solving the simultaneous equations,

$$r_1^2 = x^{*2} + y^{*2}, \quad (91)$$

$$r_2^2 = (x^* + \mu)^2 + y^{*2}, \quad (92)$$

to give,

$$x^* = (\mu^2 + r_1^2 - r_2^2)/2\mu, \quad (93)$$

$$y^* = \sqrt{[(r_1 + r_2)^2 - \mu^2][\mu^2 - (r_1 - r_2)^2]}/2\mu. \quad (94)$$

Hence,

$$\mathcal{A}(r_1, r_2; \mu) = \begin{cases} 0, & \mu \geq r_1 + r_2; \\ \pi \min[r_1^2, r_2^2], & \mu \leq |r_1 - r_2|; \\ r_1^2 \cos^{-1} \left( \frac{\mu^2 + r_1^2 - r_2^2}{2\mu r_1} \right) + r_2^2 \cos^{-1} \left( \frac{\mu^2 + r_2^2 - r_1^2}{2\mu r_2} \right) \\ - \frac{1}{2} [(r_1 + r_2)^2 - \mu^2][\mu^2 - (r_1 - r_2)^2]^{1/2}, & \text{otherwise.} \end{cases} \quad (95)$$

### A.3 Derivative of $\mathcal{A}$

It is obvious from (32) that when  $|r_1 - r_2| \geq \mu$  or  $\mu \geq r_1 + r_2$  that  $d\mathcal{A}(r_1, r_2; \mu) = 0$ . Considering where  $|r_1 - r_2| \leq \mu \leq r_1 + r_2$  and using the identity,

$$\chi(r_1, r_2, \mu) = [(r_1 + r_2)^2 - \mu^2][\mu^2 - (r_1 - r_2)^2] = -\mu^4 + 2\mu^2(r_1^2 + r_2^2) - (r_1^2 - r_2^2)^2,$$

it can be shown that,

$$\begin{aligned} \frac{d\mathcal{A}(r_1, r_2; \mu)}{d\mu} &= -r_1^2 \left[ 1 - \left( \frac{\mu^2 + r_1^2 - r_2^2}{2\mu r_1} \right)^2 \right]^{-1/2} \left[ \frac{4\mu^2 r_1 - 2r_1(\mu^2 + r_1^2 - r_2^2)}{4r_1^2 \mu^2} \right] \\ &\quad - \mu_2^2 \left[ 1 - \left( \frac{\mu^2 + r_2^2 - r_1^2}{2\mu r_2} \right)^2 \right]^{-1/2} \left[ \frac{4\mu^2 r_2 - 2r_2(\mu^2 + r_2^2 - r_1^2)}{4r_2^2 \mu^2} \right] \\ &\quad - \frac{1}{4} [\chi(r_1, r_2, \mu)]^{-1/2} [-4\mu^3 + 4(r_1^2 + r_2^2)\mu], \quad (96) \\ &= -r_1^2 \left[ \frac{2\mu^2 r_1 - 2r_1^3 + 2r_1 r_2^2}{2\mu r_1 \sqrt{4\mu^2 r_1^2 - (\mu^2 + r_1^2 - r_2^2)^2}} \right] \end{aligned}$$

$$\begin{aligned}
& -r_2^2 \left[ \frac{2\mu^2 r_2 - 2r_2^3 + 2r_2 r_1^2}{2\mu r_2 \sqrt{4\mu^2 r_2^2 - (\mu^2 + r_2^2 - r_1^2)^2}} \right] \\
& - \frac{(r_1^2 + r_2^2)\mu - \mu^3}{\sqrt{\chi(r_1, r_2, \mu)}}, \tag{97}
\end{aligned}$$

$$\begin{aligned}
& = -r_1 \left[ \frac{2r_1(\mu^2 - r_1^2 + r_2^2)}{2\mu \sqrt{-\mu^4 + 2\mu^2(r_1^2 + r_2^2) - (r_1^2 - r_2^2)^2}} \right] \\
& - r_2 \left[ \frac{2r_2(\mu^2 - r_2^2 + r_1^2)}{2\mu \sqrt{-\mu^4 + 2\mu^2(r_1^2 + r_2^2) - (r_1^2 - r_2^2)^2}} \right] \\
& - \frac{(r_1^2 + r_2^2)\mu - \mu^3}{\sqrt{\chi(r_1, r_2, \mu)}}, \tag{98}
\end{aligned}$$

$$\begin{aligned}
& = -\frac{r_1^2(\mu^2 - r_1^2 + r_2^2)}{\mu \sqrt{\chi(r_1, r_2, \mu)}} - \frac{r_2^2(\mu^2 - r_2^2 + r_1^2)}{\mu \sqrt{\chi(r_1, r_2, \mu)}} - \frac{(r_1^2 + r_2^2)\mu - \mu^3}{\sqrt{\chi(r_1, r_2, \mu)}}, \tag{99}
\end{aligned}$$

$$\begin{aligned}
& = -\frac{2\mu^2(r_1^2 + r_2^2) - (r_1^4 + r_2^4 - 2r_1^2 r_2^2) - \mu^4}{\mu \sqrt{\chi(r_1, r_2, \mu)}}, \tag{100}
\end{aligned}$$

$$\begin{aligned}
& = -\frac{\sqrt{\chi(r_1, r_2, \mu)}}{\mu}. \tag{101}
\end{aligned}$$

#### A.4 Integral of $\sqrt{\chi}$

Expanding the integral of (41),

$$\begin{aligned}
I & = \frac{1}{2} \int_{|\alpha-\rho|}^{w(z)} \sqrt{\chi(\alpha, \rho, \mu)} \, d\mu^2, \\
& = \frac{1}{2} \int_{|\alpha-\rho|^2}^{w^2(z)} \sqrt{-\mu^4 + 2\mu^2(\alpha^2 + \rho^2) - (\alpha^2 - \rho^2)^2} \, d\mu^2. \tag{102}
\end{aligned}$$

Completing the square,

$$I = \frac{1}{2} \int_{|\alpha-\rho|^2}^{w^2(z)} \sqrt{4\alpha^2 \rho^2 - [\mu^2 - (\alpha^2 + \rho^2)]^2} \, d\mu^2. \tag{103}$$

Writing<sup>13</sup>  $\alpha^2 + \rho^2 - \mu^2 = 2\alpha\rho \cos y$ ,

$$I = \frac{1}{2} \int_{\mu=|\alpha-\rho|}^{\mu=w(z)} \sqrt{4\alpha^2 \rho^2 - 4\alpha^2 \rho^2 \cos^2 y} \, 2\alpha\rho \sin y \, dy, \tag{104}$$

$$= 2\alpha^2 \rho^2 \int_{\mu=|\alpha-\rho|}^{\mu=w(z)} \sin^2 y \, dy, \tag{105}$$

$$= \alpha^2 \rho^2 \int_{\mu=|\alpha-\rho|}^{\mu=w(z)} 1 - \cos 2y \, dy, \tag{106}$$

<sup>13</sup>Selected such that the integral equals zero for  $w(z) = |\alpha - \rho|$ .

$$= \alpha^2 \rho^2 \left[ y - \frac{1}{2} \sin 2y \right]_{r=|\alpha-\rho|}^{r=w(z)}, \quad (107)$$

$$= \alpha^2 \rho^2 \left[ \cos^{-1} \left( \frac{\alpha^2 + \rho^2 - \mu^2}{2\alpha\rho} \right) - \sin y \cos y \right]_{\mu=|\alpha-\rho|}^{\mu=w(z)}, \quad (108)$$

$$= \alpha^2 \rho^2 \left[ \cos^{-1} \left( \frac{\alpha^2 + \rho^2 - \mu^2}{2\alpha\rho} \right) - \frac{\alpha^2 + \rho^2 - \mu^2}{2\alpha\rho} \sqrt{1 - \left( \frac{\alpha^2 + \rho^2 - \mu^2}{2\alpha\rho} \right)^2} \right]_{\mu=|\alpha-\rho|}^{\mu=w(z)}, \quad (109)$$

$$= \left[ (\alpha\rho)^2 \cos^{-1} \left( \frac{\alpha^2 + \rho^2 - \mu^2}{2\alpha\rho} \right) - \frac{\alpha^2 + \rho^2 - \mu^2}{4} \sqrt{\chi(\alpha, \rho, \mu)} \right]_{\mu=|\alpha-\rho|}^{\mu=w(z)}, \quad (110)$$

$$= (\alpha\rho)^2 \cos^{-1} \left( \frac{\alpha^2 + \rho^2 - w^2(z)}{2\alpha\rho} \right) - \frac{\alpha^2 + \rho^2 - w^2(z)}{2\alpha\rho} \sqrt{\chi(\alpha, \rho, w(z))}. \quad (111)$$

## References

- Amodeo, A., Mattis, I., Boeckmann, C., D'Amico, G., Mueller, D., Osterloh, L., Chaikovsky, A., Pappalardo, G., Ansmann, A., Apituley, A., Alados-Arboledas, L., Balis, D., Comeron, A., Freudenthaler, V., Mitev, V., Nicolae, D., Papayannis, A., Perrone, M. R., Pietrueczuk, A., Pujadas, M., Putaud, J. P., Ravetta, F., Rizi, V., Simeonovs, V., Spinelli, N., Stebel, K., Stoyanov, D., Trickl, T., and Wiegner, M. (2007). Optimization of lidar data processing: A goal of the EARILINET-ASOS project - art. no. 67500f.
- Ansmann, A., Riebesell, M., and Weitkamp, C. (1990). Measurement of atmospheric aerosol extinction profiles with a Raman lidar. *Optics Letters*, 15(13):746–748.
- Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W. (1992). Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar. *Applied Optics*, 31(33):7113–7131.
- Argall, P. and Sica, R. (2003). *Lidar: Atmospheric sounding introduction*, pages 1169–1176. Academic Press, San Diego, CA.
- Bucholtz, A. (1995). Rayleigh-scattering calculations for the terrestrial atmosphere. *Applied Optics*, 34(15):2765–2773.
- Collis, R. T. H. (1966). Lidar — a new atmospheric probe. *Quarterly Journal of the Royal Meteorological Society*, 92(392):220–230.
- Council, S. (2003 – 2010). Chilbolton Facility for Atmospheric and Radio Research, [Wrench, C.L.]. British Atmospheric Data Centre.
- Fenner, W. R., Hyatt, H. A., Kellam, J. M., and Porto, S. P. S. (1973). Raman cross-section of some simple gases. *Journal of the Optical Society of America*, 63(1):73–77.

- Fernald, F., Herman, B., and Reagan, J. (1972). Determination of aerosol height distributions by lidar. *Journal of Applied Meteorology*, 11:482–489.
- Friedland, S. S., Katzenstein, J., and Zatzick, M. R. (1956). Pulsed searchlighting the atmosphere. *Journal of Geophysical Research*, 61(3):415–434.
- Gaumet, J. L., Heinrich, J. C., Cluzeau, M., Pierrard, P., and Prieur, J. (1998). Cloud-base height measurements with a single-pulse erbium-glass laser ceilometer. *Journal of Atmospheric and Oceanic Technology*, 15(1):37–45.
- Grant, W. B. (1991). Differential absorption and Raman lidar for water-vapor profile measurements — A review. *Optical Engineering*, 30(1):40–48.
- Guerrero-Rascado, J. L., Ruiz, B., Chourdakis, G., Georgoussis, G., and Alados-Arboledas, L. (2008). One year of water vapour Raman lidar measurements at the Andalusian Centre for Environmental Studies (CEAMA). *International Journal of Remote Sensing*, 30:5437–5453.
- Halldórsson, T. and Langerholc, J. (1978). Geometrical form factors for the lidar function. *Applied Optics*, 17(2):240–244.
- Hughes, H. G., Ferguson, J. A., and Stephens, D. H. (1985). Sensitivity of a lidar inversion algorithm to parameters relating atmospheric backscatter and extinction. *Applied Optics*, 24(11):1609–1613.
- Hulburt, E. O. (1937). Observations of a searchlight beam to an altitude of 28 kilometers. *Journal of the Optical Society of America*, 27(11):377–382.
- King, L. (1923). On the complex anisotropic molecule in relation to the dispersion and scattering of light. *Proceedings of the Royal Society A*, 104:333–357.
- Klett, J. D. (1981). Stable analytical inversion solution for processing lidar returns. *Applied Optics*, 20(2):211–220.
- Klett, J. D. (1985). Lidar inversion with variable backscatter extinction ratios. *Applied Optics*, 24(11):1638–1643.
- Leblanc, T. and McDermid, I. S. (2008). Accuracy of Raman lidar water vapor calibration and its applicability to long-term measurements. *Applied Optics*, 47(30):5592–5603.
- Mattis, I., Muller, D., Ansmann, A., Wandinger, U., Preissler, J., Seifert, P., and Tesche, M. (2008). Ten years of multiwavelength Raman lidar observations of free-tropospheric aerosol layers over central Europe: Geometrical properties and annual cycle. *Journal of Geophysical Research-Atmospheres*, 113(D20).
- Measures, R. M. (1992). *Lidar Remote Sensing: Fundamentals and Applications*. Krieger Publishing Company, Malabar, Florida, second edition.
- Noh, Y. M., Kim, Y. J., Choi, B. C., and Murayama, T. (2007). Aerosol lidar ratio characteristics measured by a multi-wavelength Raman lidar system at Anmyeon Island, Korea. *Atmospheric Research*, 86(1):76–87.
- Office, U. (1976). U.S. Standard Atmosphere.

- Penndorf, R. (1957). Tables of the refractive index for standard air and the Rayleigh scattering coefficient for the spectral region between 0.2 and 20.0- $\mu$  and their application to atmospheric optics. *Journal of the Optical Society of America*, 47(2):176–182.
- Press, W. H., Vetterling, W. T., Teukolsky, S. A., and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, second edition.
- Reagan, J. A., McCormick, M. P., and Spinhirne, J. D. (1989). Lidar sensing of aerosols and clouds in the troposphere and stratosphere. *Proceedings of the IEEE*, 77(3):433–448.
- Rodgers, C. D. (2000). *Inverse methods for atmospheric sounding: Theory and practice*, volume 2 of *Series on Atmospheric, Oceanic, and Planetary Physics*. World Scientific, Singapore, second edition.
- Samokhvalov, I. V. (1979). Double-scattering approximation of lidar equation for inhomogeneous atmosphere. *Optics Letters*, 4(1):12–14.
- Velotta, R., Bartoli, B., Capobianco, R., Fiorani, L., and Spinelli, N. (1998). Analysis of the receiver response in lidar measurements. *Applied Optics*, 37(30):6999–7007.
- Volkov, S. N., Kaul, B. V., and Shefontuk, D. I. (2002). Optimal method of linear regression in laser remote sensing. *Applied Optics*, 41(24):5078–5083.
- Whiteman, D. N. (2003). Examination of the traditional Raman lidar technique. I. Evaluating the temperature-dependent lidar equations. *Applied Optics*, 42(15):2571–2592.