Refractive index and integral aerosol properties retrieval from Raman lidar data using principle component analysis

Martin de Graaf[1], **David P. Donovan**[1], **Arnoud Apituley**[2]

[1]*Royal Netherlands Meteorological Institute (KNMI), Wilhelminalaan 10, 3732 GK, De Bilt, The Netherlands E-mail:Martin.de.Graaf@knmi.nl, donovan@knmi.nl*
[2]*National Institute for Public Health and the Environment (RIVM), Anthonie van Leeuwenhoeklaan 9, 3721 MA, Bilthoven, The Netherlands E-mail:Arnoud.Apituley@rivm.nl*

ABSTRACT

Principle Component Analysis (PCA) is applied to invert multiwavelength Raman lidar aerosol extinction and backscatter data. Integral properties of the aerosol size distribution, like total volume, are approximated as linear combinations of the aerosol extinction and backscatter measurements. Furthermore, PCA kernels are sensitive to complex refractive index and aerosol type information can be derived from the measurement of optical data. The refractive index information content of the kernel sets is investigated, which can be used for routine retrievals of physical aerosol parameters from ground-based and space-based lidar measurements.

1. INTRODUCTION

Raman lidar measurements have become an powerful and frequently used tool for the determination of quantitative estimates of profile of aerosol extinction and backscatter. The Raman lidar technique allows separation of aerosol and pure molecular backscatter and extinction. Inversion of the aerosol optical data, to derive microphysical aerosol quantities, has been hampered by the fact that the inversion problem is ill-posed and solutions depend non-linearly on measurement errors [1; 2]. A finite number of extinction and backscatter measurements is generally not enough to obtain accurate detailed information about the entire aerosol size distribution. Nevertheless, various integral properties of the aerosol size distribution such as aerosol total volume, surface area density and aerosol extinction and backscatter at various wavelengths can be predicted with useful accuracy, and the predictions are much less sensitive to the retrieval method employed. PCA is particularly well suited to asses the information content of multiwavelength lidar data and retrieve aerosol bulk physical and optical properties. More importantly, the sensitivity of the retrieved aerosol parameters to measurement errors can be assessed. In general, PCA allows the prediction of aerosol physical properties without any restrictive assumption about the size distribution. However, the shape and refractive index of the particles in question must still be assumed. PCA has been successfully applied to stratospheric sulfate aerosol lidar measurements to retrieve aerosol effective radius, surface area, and volume profiles [3]. This method is here extended to include various tropospheric aerosol models, including sea salt, soot, and mineral aerosol, using the sensitivity of the PCA kernels to varying refractive index A test case for this method is presented to investigate the information content of the complex refractive index in the kernel sets.

2. METHODS

2.1. Principle Component Analysis

The Mie solutions of Maxwell's equations in the far field give the extinction and backscatter of a spherical scatterer at a given height and wavelength:

$$\alpha(\lambda) = \int_0^\infty \pi r^2 Q_{\alpha,\lambda}(r) \frac{\mathrm{d}n(r)}{\mathrm{d}r}\mathrm{d}r, \tag{1}$$

and

$$\beta_\pi(\lambda) = \int_0^\infty \pi r^2 Q_{\beta_\pi,\lambda}(r) \frac{\mathrm{d}n(r)}{\mathrm{d}r}\mathrm{d}r, \tag{2}$$

respectively, where r is the particle radius, $Q_{\beta_\pi,\lambda}(r)$ is the backscatter efficiency, $Q_{\alpha,\lambda}(r)$ is the extinction efficiency, and $\mathrm{d}n(r)/\mathrm{d}r$ is the aerosol size distribution. Equations (1) and (2) can be rewritten as

$$g_i = \int_0^\infty K_i(\lambda, r, m) \frac{dV(r)}{dr} dr, \tag{3}$$

where $\mathrm{d}V(r)/\mathrm{d}r$ is the aerosol volume size distribution, g_i is either the backscatter or extinction measurement and K_i is the appropriate extinction or backscatter volume kernel for wavelength λ_i:

$$\begin{aligned} K_i &= 3Q_{\alpha,\lambda_i}(r)/4r \quad &\text{(extinction)} \\ &= 3Q_{\beta_\pi,\lambda_i}(r)/4r \quad &\text{(backscatter)} \end{aligned}$$

The volume size distribution $\mathrm{d}V(r)/\mathrm{d}r$ can be written as

$$\frac{dV(r)}{dr} = \sum_{i=1}^{N} y_i K_i(r) + \psi(r) \tag{4}$$

where $\psi(r)$is that part of $\mathrm{d}V(r)/\mathrm{d}r$ that is orthogonal to K_i and thus inaccessible with the given set of measurements. Substitution of (3) yields

$$g_i = \int_0^\infty \sum_{j=1}^{j=N} K_i(r) K_j(r) y_j(r) dr + \int_0^\infty K_i(r) \psi(r) dr \tag{5}$$

which can be written as (since $\int_0^\infty K_i(r)\,\psi(r)dr = 0$ for $i = 1,.., N$):

$$\bar{\mathbf{g}} = \mathbf{C}\bar{\mathbf{y}} \tag{6}$$

Here, the elements of the $N \times N$ matrix $\mathbf{C}$ are defined so that $C_{ij} = \sum_k K_{ik} K_{jk}$ and $\bar{\mathbf{g}}$ and $\bar{\mathbf{y}}$ are column vectors. Since $\mathbf{C}$ is real, symmetric and positive, its eigenvalues are all real and positive and $\mathbf{C}$ can be expanded in terms of its eigenvalues and eigenvectors

$$\mathbf{C} = \mathbf{U}\mathbf{L}\mathbf{U}^{\mathbf{t}}, \tag{7}$$

where $\mathbf{L}$ is the diagonal matrix containing the eigenvalues of $\mathbf{C}$ in descending order and $\mathbf{U}$ is the matrix containing the corresponding eigenvectors as its columns. This can now easily be inverted to give

$$\bar{\mathbf{y}} = \mathbf{C}^{-\mathbf{1}}\bar{\mathbf{g}} = \mathbf{U}\mathbf{L}^{-\mathbf{1}}\mathbf{U}^{\mathbf{t}}\bar{\mathbf{g}}. \tag{8}$$

If the volume size distribution (equation 4) is discretised, equation 8 can be inserted to give

$$\begin{aligned} \bar{\mathbf{v}} &= \mathbf{K}^{\mathbf{t}}\bar{\mathbf{y}} \\ &= \mathbf{K}^{\mathbf{t}}\mathbf{U}\mathbf{L}^{-\mathbf{1}}\mathbf{U}^{\mathbf{t}}\bar{\mathbf{g}}. \end{aligned} \tag{9}$$

Any other integral property of the aerosol can now be retrieved using

$$P = \bar{\mathbf{w}}^{\mathbf{t}}\mathbf{K}^{\mathbf{t}}\mathbf{U}\mathbf{L}^{-\mathbf{1}}\mathbf{U}^{\mathbf{t}} \cdot \bar{\mathbf{g}} + \bar{\mathbf{w}}^{\mathbf{t}}\bar{\psi} \tag{10}$$

and appropriate weighting functions $\bar{\mathbf{w}}$. For volume kernels these are given by:

$$\begin{array}{ll} w_j = 1 & \text{(volume)} \\ w_j = 3/r_j & \text{(surface)} \\ w_j = 3/4\pi r_j^3 & \text{(number density)} \\ w_j = 3Q_{\alpha,\lambda}(r_j)/4r_j & \text{(aerosol extinction)} \\ w_j = 3Q_{\beta_\pi,\lambda}(r_j)/4r_j. & \text{(aerosol backscatter)} \end{array} \tag{11}$$

$\bar{\psi}$ contains the error in the solution due to the incompleteness of the kernels.

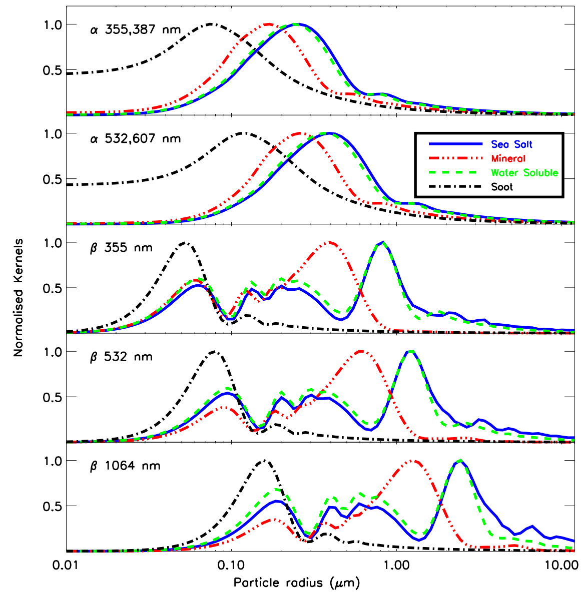

Figure 1: Normalized volume kernels (extinction and backscatter coefficients) for four aerosol types, soot (black), mineral dust aerosol (red), water soluble aerosol (green) and sea salt (blue). The top two panels show the extinction kernels averaged over the emitted and Raman shifted wavelength, the bottom three panels show the backscatter kernels at the emitted wavelengths.

2.2. Refractive index sensitivity of kernels

Since the kernels are sensitive to the refractive index of the aerosol model, this can be used to retrieve information about the refractive index of the aerosols. A sensitivity study was performed to asses the possibility of refractive index retrieval from Raman lidar measurements. At RIVM in Bilthoven a Raman lidar is deployed using a Nd:YAG laser emitting at a wavelength of 1064 nm and its doubled (532 nm) and tripled harmonics (355 nm) [5]. Therefore, in the study the backscatter coefficients at these wavelengths were considered. For the extinction coefficients the average of the extinction at the emitted and the Raman shifted wavelength is considered. This configuration is now becoming more and more popular in the aerosol lidar community, e.g. EARLINET.

Figure 1 shows an example of volume kernels (normalized extinction and backscatter coefficients) for four aerosol types: sea salt (at 95% humidity) and general water soluble aerosols (at 95% humidity), mineral aerosols, and soot. These aerosol models represent a large range in complex refractive index [4], pure soot being highly absorbing at all wavelengths and having a very large real part of the refractive index of 1.75, mineral aerosols being absorbing in the UV and having a moderate value of the real part of the refractive index of around 1.5, while sea salt and water soluble aerosols are highly scattering and have refractive indices near that of water.

Since the kernels are different at different parts of the par-

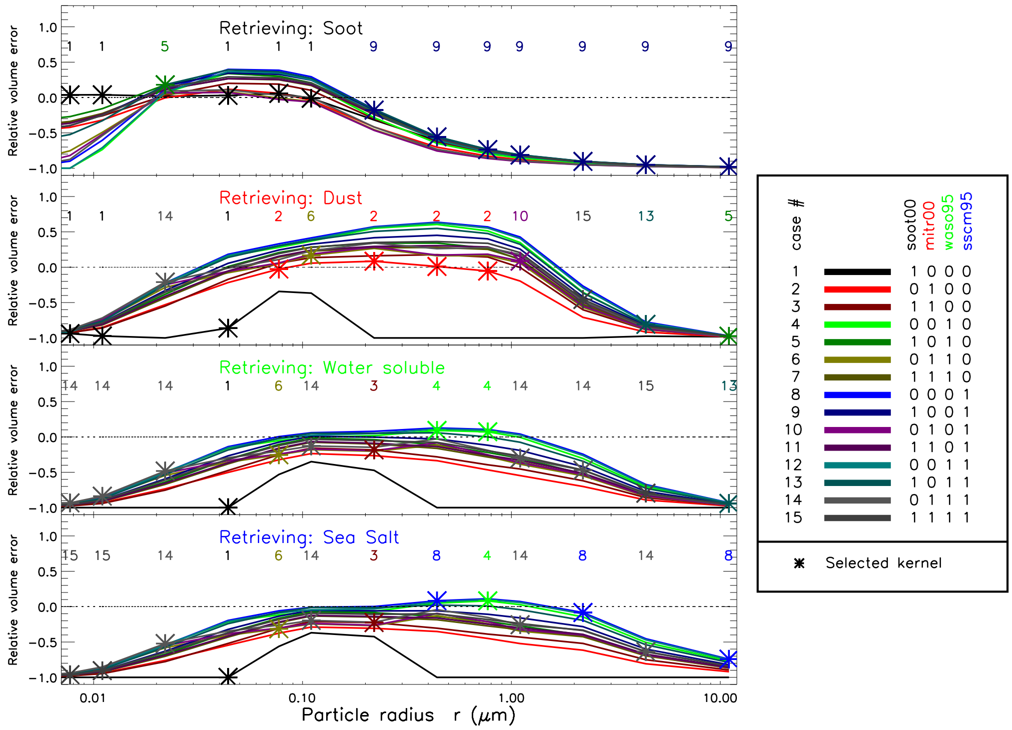

Figure 2: Retrieval of volume and refractive index for 15 mixtures of kernels: Each mixture is indicated by color and a case number, which relates the (bitwise) added kernels; 1 is soot, 2 is mineral aerosol, 4 is water soluble and 8 is sea salt. The volume retrieval for each mixture is given by the solid line in the corresponding color. The selected kernel mixture is indicated by an asterisk and the corresponding case number in the plot. The top panel shows the results when soot was used as input ('measurement'), the second panel shows the result for mineral aerosol, the third water soluble and the bottom panel sea salt.

ticle size interval, they can be used to distinguish between aerosol type (or refractive index). Clearly, no distinction will be possible between sea salt and water soluble aerosols, which have nearly the same kernel sensitivity, due to their similar refractive index.

2.3. Refractive index retrieval

The refractive index sensitive kernels can be used to retrieve both the refractive index and integral properties of the aerosol. In this case total volume is retrieved together with refractive index. In order to deal with situations where several types of aerosols may be simultaneously present, fifteen mixtures of the four different kernel (per wavelength) were constructed which consisted of averages of one or several of the original kernels:

$$K_j = \frac{\sum_{i=1}^{4} a_{i,j} K_i}{\sum_{i=1}^{4} (a_{i,j})},\ j = 1,..,15, \tag{12}$$

where $a_{i,j}$ is either 0 or 1.

All these kernels are used for a volume retrieval (using a weight $w_j = 1$) and a subsequent retrieval of the aerosol extinction and aerosol backscatter ($w_j = 3Q_{\gamma,\lambda}(r_j)/4r_j$), see equation 11. For each wavelength, the retrieval of aerosol extinction and backscatter is exact if the extinction or backscatter at the wavelength is used, i.e. the measurement of α_{λ_i} or β_{λ_i} is selected as the solution of the inversion.

The extinction and backscatter coefficients at each wavelength can also be estimated from the solutions at other wavelengths. So, instead of using the kernels at all five wavelengths, one is omitted and the extinction and backscatter at that wavelength are estimated using the remaining kernels. Then, the mixture of kernels is selected which has with the smallest difference ϵ between the α or β retrieved in this way and the measurement $\bar{\mathbf{g}}$

$$\epsilon = |\bar{\mathbf{g}} - \{\bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}}\}^{\mathbf{t}}|, \tag{13}$$

with

$$\{\bar{\alpha}, \bar{\boldsymbol{\beta}}\}^{\mathbf{t}} = \bar{\mathbf{w}}^{\mathbf{t}}\mathbf{K}^{\mathbf{t}}\mathbf{U}\mathbf{L}^{-\mathbf{1}}\mathbf{U}^{\mathbf{t}} \cdot \bar{\mathbf{g}} \tag{14}$$

where $\mathbf{U}\mathbf{L}^{-\mathbf{1}}\mathbf{U}^{\mathbf{t}}$ is a $N{-}1 \times N{-}1$ matrix. Once the optimum mixture is selected this set is then used to predict the aerosol physical properties.

3. RESULTS

Figure 2 shows the result of the refractive index retrieval for four cases. The measurement sets in these cases are Mie computations for aerosol types that are equal to the four base mixtures, i.e. soot, mineral aerosol water soluble aerosol and sea salt. Mie calculations were performed with different particle size radii (x-axis) for all four aerosol types, and the volume was retrieved with PCA using all 15 mixtures. The relative volume error $|(V_{\rm PCA} - V_{\rm true})/V_{\rm true}|$ was plotted as function of particle radius for each mixture in Figure 2 as colored lines. The legend explains the color codes, which are linked to a case number since the mixtures and colors are bitwise mixed. In addition α and β are also retrieved, as explained above. To select 'the best' volume estimate, the kernel mixture with the lowest error in α and β is selected. This selection is indicated with a asterisk in Figure 2 and the according case number printed above it.

For example, in the top panel of the figure, for particle radius 0.007 to 0.1 μm, the smallest error in α and β are produced when only soot is selected in the kernel (except for a particle radius 0.02 μm, when a mixture of soot and water soluble apparently produces a smaller error than soot alone). For larger particle radii the smallest error is produced by a mixture of soot and sea salt (case number 9). These choices of mixtures also produce the smallest or close to the smallest error in volume, which is to be expected, since the retrieved aerosol type is soot. According to Figure 1 the soot kernels are not sensitive for particle radii larger than 1.0 μm, so no correct retrieval of refractive index can be expected beyond this particle radius. The same holds for the volume retrieval, which error increases rapidly for larger particle radii.

For desert dust the retrieval is very good in the range of particle radius where the mineral kernels are sensitive, 0.06 to 1 μm, and the selection of mineral kernels also ensures the smallest error in volume.

The retrieval of water soluble aerosols and sea salt produce nearly similar results, since the kernels are more or less the same. The kernels are sensitive to a large part of the particle radius space (0.05 to 3 μm), but only distinctive at a larger particle radius. Therefore, a fairly good volume estimate can be obtained for a particle radii of about 0.08 to 2 μm, but only at larger particle radius the retrieval of refractive index is reasonable (although no distinction can be made between the selection of sea salt or water soluble aerosol). At intermediate particle radius (0.08 to 2 μm) all kernels are sensitive and good volume and α and β retrievals can be achieved with any mixture, but no distinction can be made between the mixtures.

4. CONCLUSION

Microphysical quantities can be retrieved from inversion of multiwavelength Raman lidar data using PCA, accounting for aerosol extinction and backscatter, without making *a piori* assumptions of the aerosol size distribution. Integral properties of the size distribution can be retrieved in this way if the kernels are valid, which means the aerosols are spherical. In previous studies in which PCA was used to invert lidar data, also refractive index needed to be assumed. Here, the sensitivity of the kernels to complex refractive index is demonstrated and the possibility of retrieving the refractive index from inversion of the data.

The sensitivity of the kernels to refractive index is dependent on the size distribution and can only be retrieved with useful accuracy in a part of the particle radius space. For the wavelengths considered here, this implies that only soot particles smaller smaller than about 1 μm can be retrieved and mineral aerosols between 0.06 and 1 μm. Accurate results for total volume or other integral properties can be retrieved for sea salt and water soluble aerosols between 0.08 to 2 μm, but the refractive index is only accurate at the largest particle radii.

ACKNOWLEDGMENTS

This work was financed by the Netherlands Institute for Space Research, project number GO-2005/075.

REFERENCES

1. Veselovskii, I., A. Kolgotin, V. Griaznov, D. Müller, U. Wandinger and D. N. Whiteman, Inversion with regularization for the retrieval of tropospheric aerosol parameters from multiwavelength lidar sounding, *Appl. Opt.*, *41*(18), 3685–3699, 2002.
2. Veselovskii, I., A. Kolgotin, V. Griaznov, D. Müller, K. Franke and D. N. Whiteman, Inversion of multiwavelength Raman lidar data for retrieval of bimodal aerosol size distribution, *Appl. Opt.*, *43*(5), 1180–1195, 2004.
3. Donovan, D. P. and A. I. Carswell, Principle component analysis applied to multiwavelength lidar aerosol backscatter and extinction measurements, *Appl. Opt.*, *36*(36), 9406–9424, 1997.
4. Hess, M., P. Koepke and I. Schult, Optical Properties of Aerosols and Clouds: The software package OPAC, *Bull. Am. Met. Soc.*, *79*, 831–844, 1998.
5. Wilson, K., A. Apituley and C. Potma, Caeli - a Raman lidar for the diurnal observation of clouds, aerosol and water vapour, *This conference*, 2008.