Inversion with regularization for the retrieval of tropospheric aerosol parameters from multiwavelength lidar sounding

Igor Veselovskii, Alexei Kolgotin, Vadim Griaznov, Detlef Müller, Ulla Wandinger, and David N. Whiteman

We present an inversion algorithm for the retrieval of particle size distribution parameters, i.e., mean (effective) radius, number, surface area, and volume concentration, and complex refractive index from multiwavelength lidar data. In contrast to the classical Tikhonov method, which accepts only that solution for which the discrepancy reaches its global minimum, in our algorithm we perform the averaging of solutions in the vicinity of this minimum. This averaging stabilizes the underlying ill-posed inverse problem, particularly with respect to the retrieval of number concentration. Results show that, for typical tropospheric particles and 10% error in the optical data, the mean radius could be retrieved to better than 20% from a lidar on the basis of a Nd:YAG laser, which provides a combination of backscatter coefficients at 355, 532, and 1064 nm and extinction coefficients at 355 and 532 nm. The accuracy is improved if the lidar is also equipped with a hydrogen Raman shifter. In this case two additional backscatter coefficients at 416 and 683 nm are available. The combination of two extinction coefficients and five backscatter coefficients then allows one to retrieve not only averaged aerosol parameters but also the size distribution function. There was acceptable agreement between physical particle properties obtained from the evaluation of multiwavelength lidar data taken during the Lindenberg Aerosol Characterization Experiment in 1998 (LACE 98) and *in situ* data, which were taken aboard aircraft. © 2002 Optical Society of America

OCIS codes: 280.1100, 280.1310, 280.3640, 290.1090, 010.3920, 010.1110.

1. Introduction

Tropospheric aerosols play an important role in the Earth's radiation budget. Increased aerosol loading directly affects incident short-wave radiation through increased scattering and absorption. In addition, an increase in aerosol number density can influence the probability of and composition of water clouds in the troposphere by creating more nucleation sites for water droplets. This indirect effect of aerosols[1] can lead to clouds of higher albedo and greater persistence. These changes in cloud properties also have a strong effect on the Earth's radiation budget. Quantifying these aerosol effects is difficult, however. In the most recent Intergovernmental Panel on Climate Change (IPCC) report,[2] large uncertainties were attributed to both the direct and the indirect effect of aerosols. Depending on the amount of absorption by the aerosol, the direct effect on atmospheric radiation can be either positive or negative whereas the indirect effect can be as large as -2 W/m^2, which by itself could be responsible for as much as a 1–2 K cooling at the surface.

Experimental techniques that could permit a remote quantification of aerosol properties, particularly if cloud measurements could be made simultaneously and in the same volume, would provide valuable new tools for studying the influence of aerosols on climate. Here improvements we present in the retrieval algorithms for the remote determination of aerosol properties using one such technique based on multiwavelength Raman lidar.

During the past 30 years numerous inversion techniques have been proposed for the retrieval of physical

I. Veselovskii (katyv@orc.ru), A. Kolgotin, and V. Griaznov are with the Physics Instrumentation Center of the General Physics Institute, Troitsk, Moscow Region 142190, Russia. D. Müller (detlef@tropos.de) and U. Wandinger (ulla@tropos.de) are with the Institute for Tropospheric Research, Permoserstrasse 15, 04318 Leipzig, Germany. D. Whiteman (david.whiteman@gsfc.nasa.gov) is with the NASA Goddard Space Flight Center, Greenbelt, Maryland 20771-0001.

particle parameters from measurements of respective optical properties at multiple wavelengths.[3–9] These techniques are generally used in passive remote sensing, as, for example, in sunphotometry, which provides columnar-integrated aerosol properties. In contrast, active remote-sensing instruments, e.g., lidar systems, offer the opportunity to perform measurements of optical parameters with high vertical resolution. For this reason there have been numerous previous attempts to transfer the experience accumulated in passive remote sensing to the field of lidar remote sensing.[10]

Unfortunately, the results of these attempts are not satisfying. The main obstacles in the application of these methods to lidar remote sensing was the poor accuracy of the retrieved particle backscatter and extinction coefficients[11] compared with passive instruments, and the small number of radiated wavelengths. Considerable accuracy of the input optical parameters is required for the retrieval of physical parameters by use of the underlying mathematical models, which are called ill-posed inverse problems and are characterized by a nonunique solution space for a given optical data set.[12–15]

Attempts were made to overcome these problems by use of a large number of *a priori* assumptions about the investigated solution space. Particle complex refractive index and shape of the size distribution were assumed to be known. Basically the developed techniques consisted of forward schemes for which optical data were calculated with the above-mentioned assumptions and then compared with the experimental data. It became possible to derive, e.g., the particle mean radius.[16–20] Unfortunately, the highly stringent *a priori* assumptions allowed only stratospheric particle parameters to be derived, because in this case the particle refractive index and the shape of the size distribution are rather well known.[19] However, most of the atmospheric particles are concentrated in the troposphere, where they exhibit large temporal and spatial variability. In addition, the chemical and physical properties are highly diverse. Consequently, it remained impossible to provide an accurate assessment of their influence on, e.g., the Earth's radiation budget, which is one of the major issues in climate research.[2]

It was not until the end of the 1990s that a new approach proved to be successful. A new highly sophisticated scanning six-wavelength, eleven-channel aerosol lidar,[21] which makes use of the powerful Raman lidar technique, was set up at the Institute for Tropospheric Research (ITR). The independent measurement of signals that describe aerosol backscatter and extinction[23–26] on the basis of the detection of Raman signals permits the retrieval of highly accurate profiles of the particle backscatter coefficients at 355, 400, 532, 710, 800, and 1064 nm and of particle extinction coefficients at 355 and 532 nm. A specifically designed inversion scheme,[13,15,27] developed at the ITR, and which is based on the concept of Tikhonov's regularization,[3,4] then allowed us to process this optical information into physical proper-

ties. This method of inversion with regularization followed from previous studies, dating back to the late 1970s,[28–32] in which it could be shown that the introduction of basis functions and physical constraints, such as positivity and smoothness of the derived size distribution allowed the effective radius, volume, surface area, and number concentration to be derived.[32] The constraints are formulated as so-called penalty terms. The goal of the inversion is to minimize these terms, which stabilize the underlying ill-posed problem.[3,4]

A major breakthrough in particle characterization was obtained with the ITR algorithm when it was shown that the combined use of particle backscatter and extinction coefficients permitted not only the mean and integral parameters of particle size distributions to be derived, but also their mean complex refractive index. Consequently it became possible to calculate the single-scattering albedo,[15,33,34] which is one of the most important parameters in climate impact studies. Because the retrieval scheme requires high accuracy of the input optical data, preferably in the range of better than 20%,[14,15] the Raman lidar technique again proved to be the key to a successful particle characterization.

However several obstacles remain to be resolved. From an instrumental point of view, despite the impressive capabilities of the ITR system, multiwavelength lidars are still expensive and complicated to operate. Therefore it is desirable to examine the possibility of obtaining an acceptable retrieval accuracy from simplified lidar systems.[35,36] A lidar that uses only a triple Nd:YAG laser already provides backscatter coefficients at 355, 532, and 1064 nm, and extinction coefficients at 355 and 532 nm. A few studies meanwhile outline the potential of such a combination of data.[14,35] The number of laser wavelengths can be increased while a still inexpensive laser source can be maintained if a hydrogen Raman shifter is added. Calculations have shown that the use of two additional backscatter coefficients, which are provided by the Raman shifter, improves the accuracy of the retrieved parameters.

From an inversion point of view it has been repeatedly demonstrated that, in general, particle number concentrations are difficult to derive.[14,15,33,37] Depending on the specific measurement situation, the inversion into other parameters in some cases also showed large uncertainties that require further stabilizing procedure. Use of the above-mentioned reduced number of three backscatter and two extinction coefficients destabilizes the inverse problem and thus also calls for the introduction of additional stabilizing tools.[14,35]

Here we present a modified version of the Tikhonov approach that was developed at the Physics Instrumentation Center (PIC). In the modified version we average the solutions in the vicinity of the minimum penalty functions, in contrast with the classical Tikhonov method[4] that accepts only that solution for which the penalty function reaches its global minimum. We observed both a significant improvement

in retrieval accuracy of the particle parameters, as well as the fact that the approach works for a small number of emitted laser wavelengths.

In Section 2 we present the methodology. In Section 3 we show the results of simulations with synthetic data. In Section 4 we present the application to experimental data and make a comparison with airborne *in situ* data and two other computer codes that were used to process the same optical data. A summary is presented in Section 5.

2. Methodology

The inverse problem of particle size distribution retrieval is formulated in the form of Fredholm integral equations. The following relations give the aerosol backscatter and extinction coefficients:

$$\beta_i = \int_0^\infty K_\beta(m, r, \lambda_i) f(r) dr, \tag{1}$$

$$\alpha_i = \int_0^\infty K_\alpha(m, r, \lambda_i) f(r) dr. \tag{2}$$

where r is the particle radius; $m = m_R - im_I$ denotes the complex particle refractive index, with real part m_R and imaginary part m_I; λ_i is the wavelength, $K_\beta(m, r, \lambda_i)$ and $K_\alpha(m, r, \lambda_i)$ are, respectively, the backscatter and the extinction kernel functions, which can be calculated from Mie theory in the case of spherical particles.[38] The term $f(r)$ is the particle size distribution expressed as the number of particles per unit volume between r and $r + dr$. Equations (1) and (2) can be rewritten as a generalized integral equation:

$$g_j(\lambda_i) = \int_{r_{min}}^{r_{max}} K_j(m, r, \lambda_i) f(r) dr, \tag{3}$$

where j is α or β, $g_j(\lambda_i)$ are the optical data at wavelength λ_i, and r_{min} and r_{max} denote the lower and the upper limits of the particle radii.

Equation (3) has no analytical solution. A detailed description of the following method of inversion with regularization is given elsewhere.[3,4] In this approach, Eq. (3) can be rewritten in the following form:

$$g_p^{\delta} = \int_{r_{min}}^{r_{max}} K_p(m, r) f^{\delta}(r) dr, \tag{4}$$

where g_p^{δ} are the distorted optical data (α or β), $f^{\delta}(r)$ is the corresponding function that describes the particle size distribution, and $p = j, \lambda$. The integral of Eq. (4) can then be approximated by a sum that consists of the following superposition of base functions $B_j(r)$:

$$f^{\delta}(r) = \tilde{f}^{\delta}(r) + \varepsilon = \sum_j C_j B_j(r) dr + \varepsilon. \tag{5}$$

The term $\tilde{f}^{\delta}(r)$ is an approximated solution of Eq. (5), ε is the error in the solution, and C_j are constants or so-called weight coefficients. $B_j(r)$ are B-spline functions, which in our case have a triangular shape. These functions are given by the expression

$$B_j(r) = \begin{cases} 0; & r < r_{j-1} \\ 1 - \dfrac{r_j - r}{r_j - r_{j-1}}; & r_{j-1} < r \le r_j \\ & \qquad\qquad j = 1 \cdots N, \\ 1 - \dfrac{r - r_j}{r_{j+1} - r_j}; & r_j < r \le r_{j+1} \\ 0; & r > r_{j+1} \end{cases} \tag{6}$$

where r_0 and r_{N+1} limit the size range within which the inversion is performed. In the following this range is denoted as an inversion window, and N is the number of B-spline functions. Points r_j are equally separated inside the interval $[r_0, r_{N+1}]$. In our algorithm N always coincides with the number of optical data. Using Eqs. (4) and (5) we can write the optical data as a linear combination:

$$g_p^{\delta} = \sum_{j=1}^N A_{pj}(m) C_j + \varepsilon^{\nabla}, \tag{7}$$

where A_{pj} and ε^{∇} are calculated from the kernel functions, the base functions, and the errors as

$$A_{pj}(m) = \int_{r_{min}}^{r_{max}} K_p(m, r) B_j(r) dr, \tag{8}$$

$$\varepsilon_p^{\nabla} = \int_{r_{min}}^{r_{max}} K_p(m, r) \varepsilon(r) dr. \tag{9}$$

By writing the optical data as a vector $\mathbf{g}^{\delta} = [g_p^{\delta}]$, the weight coefficients as a vector $\mathbf{C} = [C_j]$, the errors as a vector $\boldsymbol{\varepsilon}^{\nabla} = [\varepsilon_p^{\nabla}]$, one can rewrite Eq. (7) in the following vector–matrix form:

$$\mathbf{g}^{\delta} = A\mathbf{C} + \boldsymbol{\varepsilon}^{\nabla}, \tag{10}$$

where the matrix $A = [A_{pj}]$ is the so-called weight matrix, the elements of which are calculated from Eq. (8). The solution of Eq. (10) then gives the vector of weight coefficients as

$$\mathbf{C} = A^{-1}\mathbf{g}^{\delta} + \boldsymbol{\varepsilon}, \tag{11}$$

where $\boldsymbol{\varepsilon} = -A^{-1}\boldsymbol{\varepsilon}^{\nabla}$ is the error vector and A^{-1} is the inverse of matrix A. Hence the problem of inversion becomes a problem of determination of weight coefficients C_j.

As outlined in Section 1, the simple solution of Eq. (11) is in general unstable. As a consequence, it is not possible to reconstruct the exact solution f. This instability could be suppressed by regularization. The idea behind it is illustrated when we rewrite Eq. (3) in the following operator form:

$$\hat{A}f = g. \tag{12}$$

If instead of the exact value of g we know the approximate value of g^δ, which means that the norm in Euclidean space

$$\|g - g^\delta\|_E \le \delta, \qquad (13)$$

the essence of the regularization technique lies in the construction of an algorithm that determines element $\tilde{f}^\delta$ appropriate to the pair (g^δ, δ), such that the convergence $\tilde{f}^\delta \xrightarrow{\delta \to 0} f$ is realized. According to Tikhonov and Arsenin,[4] we define the function $M^\gamma[f, g^\delta]$ as

$$M^\gamma[f, g^\delta] = \|\hat{A}f - g^\delta\|_E^2 + \gamma\Gamma(f), \qquad (14)$$

where $\rho' = \|\hat{A}f - g^\delta\|_E^2$ is the discrepancy, γ is a nonnegative regularization parameter or so-called Lagrange multiplier, and $\Gamma(f)$ is a penalty term that stabilizes the inverse problem.

The solution of Eq. (12) can then be found from the minimization of $M^\gamma[\tilde{f}_\gamma^\delta, g^\delta] \to \min$. In matrix–vector form it is written as[3]

$$A^T A\mathbf{C} - A^T g^\delta + \gamma H\mathbf{C} = 0. \qquad (15)$$

So finally the weight coefficients can be derived from the following relation[3]:

$$\mathbf{C} = (A^T A + \gamma H)^{-1} A^T g^\delta. \qquad (16)$$

The main difference between Eq. (11) and Eq. (16) is the so-called smoothing matrix H, which for eight B-spline functions has the following form[3]:

$$H = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ -2 & 5 & -4 & 1 & 0 & 0 & 0 & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 & 0 & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & 0 \\ 0 & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\ 0 & 0 & 0 & 1 & -4 & 6 & -4 & 1 \\ 0 & 0 & 0 & 0 & 1 & -4 & 5 & -2 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \end{pmatrix}.$$

This smoothing matrix describes the physical constraint that size distributions do not show large oscillations within a small particle size range. The specific form of H influences the maximum difference between the weight factors of successive base functions.[3] The above-given H describes the change in weight factors of three successive base functions, which is equivalent to smoothness of the solution in the second derivative.[9]

The Lagrange multiplier γ, which can take values from 0 to ∞, determines the degree of smoothing, i.e., the strength of H. There are different methods for the choice of Lagrange multiplier, the most common of which are the maximum-likelihood method, the Bayesian approach, the generalized cross-validation method, and the minimum discrepancy principle.[39–41] For our study we chose the minimum discrepancy principle, because the criteria are the simplest and most natural. The main

problem in the application of minimum discrepancy is the necessity to know *a priori* the expected error in the data. Otherwise the smallest discrepancy is observed for $\gamma \to 0$, and in this case the solutions oscillate from $-\infty$ to $+\infty$. For our situation this method can be modified. We can use the fact that the solutions have to be positive. In our modified discrepancy ρ only modules $|f^\delta|$ are used so f^δ, for which $\rho = \||A|f^\delta| - g^\delta\|$ minimizes the function in Eq. (14), are accepted as solutions. For small γ the modified discrepancy $\rho \to \infty$, with the increase of γ the value of ρ decreases, and the solution $f^\delta \to f$. For large γ the modified discrepancy ρ coincides with the classical discrepancy ρ' introduced in Eq. (14). It should be pointed out that the found solutions f^δ could contain negative oscillations, because the constraint of nonnegativity is applied only to discrepancy, but not to solutions themselves.

To consider the relative errors of optical data, we normalize the discrepancy to[9]

$$\rho \equiv \frac{1}{N} \sum_i \frac{\|g_i^\delta - \hat{A}|f^\delta|\|}{g_i^\delta}.$$

The value of γ depends on the inversion interval $[r_{\min}, r_{\max}]$, the optical data errors δ, and the complex refractive index of the particles. The choice of solution in our approach includes the following steps:

The values of $r_{\min}$, $r_{\max}$ and the complex refractive index $m = m_R - im_I$ are varied in the intervals $\Delta r_{\min}$, $\Delta r_{\max}$, Δm_R, and Δm_I.

$R_{\min}$ ranges from 0.05 to 0.5 μm. $R_{\max}$ usually ranges from 0.1 to 1 μm, but for the simulation of big particle retrieval this range could be increased up to 10 μm. At the preliminary stage of calculation the step widths of the variation of $R_{\min}$ and $R_{\max}$ are 0.05 and 0.1 μm, respectively. When the approximate locations of $R_{\min}$ and $R_{\max}$ are established, the calculations are repeated with a step width of 0.01–0.02 μm within a smaller interval. In general, several hundred inversion windows are considered. As a consequence a better resolution of the investigated particle size distribution compared to the rather limited set of 50 windows used by Müller *et al.*[13] is achieved. The real part, m_R, ranges from 1.1 to 1.8 and the imaginary part, m_I, from 0 to 0.1. After preliminary estimation the range is decreased and the calculations are performed with a step width of 0.01.

For every value of $r_{\min}$, $r_{\max}$, m_R, m_I within the chosen intervals and for the set of γ, the program calculates the solution $\tilde{f}_\gamma^\delta$ with Eqs. (16) and (5). The set of γ is determined as $\gamma = 2^K 10^{-b}$, where $K = 1, 2 \ldots 25$ is the calculation number. Parameter b is chosen between 20 and 28 to achieve the minimization of the function in Eq. (14) inside the calculation interval. The integration step in Eq. (8) is usually chosen to be 0.002 μm.

For each solution we determined the discrepancy $\rho(\gamma)$. The $\tilde{f}_\gamma^\delta$ that corresponds to the minimum of $\rho(\gamma)$ is considered to be the solution of Eq. (4).
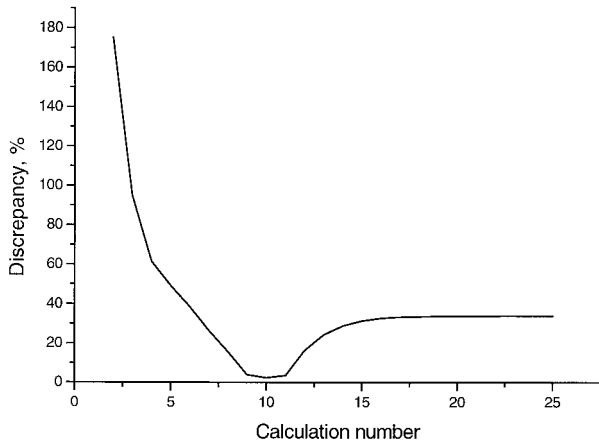
Fig. 1. Discrepancy ρ versus Lagrange multiplier γ. Calculation number K is related to the Lagrange multiplier as $\gamma = 2^K\, 10^{-23}$.

The use of modified discrepancy allows one to determine the Lagrange multiplier without initial guesses about data errors. An example is shown in Fig. 1, where for synthetic optical data the discrepancy ρ is plotted versus the calculation step K. If the Lagrange multiplier is chosen too large, the solution is oversmooth, and for small γ it oscillates. The corresponding illustrations can be found in Ref. 13.

3. Simulations

A. Methodology

We performed computer simulations to estimate the optimum number of laser wavelengths, the proportions in which the backscatter and extinction coefficients should be combined, and the realistic accuracy of the retrieved particle size parameters. We considered the laser wavelengths within the 355-nm interval, which describes the triple harmonic, and within 1064 nm, which describes the fundamental wavelength of a Nd:YAG laser. Data obtained at shorter wavelengths could be affected by atmospheric ozone absorption, and the use of longer wavelengths significantly complicates the lidar system and the retrieval of optical data. Particle backscatter and extinction coefficients were calculated[38] under the assumption of wavelength- and size-independent complex refractive indices. The initial particle size distributions, which were needed for these calculations, were assumed to be log normal:

$$\frac{\partial n(r)}{\partial \ln r} = \frac{n_t}{(2\pi)^{1/2}\ln\sigma}\exp\left[-\frac{(\ln r - \ln r_0)^2}{2(\ln\sigma)^2}\right],$$

where $n(r)$ denotes the investigated particle number concentration distribution, n_t is the total number of particles, and σ describes the width of the distribution. In the retrieval procedure we generally used a combination of particle backscatter and extinction coefficients. We used this approach because previous findings showed that this combination allows the retrieval not only of particle size parameters but also the complex refractive index.[13,15]

As mentioned before, accurate values for the extinction coefficient can be obtained with the Raman method, which demands high-power laser radiation. One has to observe that the Raman scattering cross section is proportional to λ^{-4}. For this reason, lidar systems based on a Nd:YAG laser now provide extinction coefficients only at 355 and 532 nm but not at 1064 nm. Therefore, we restricted ourselves to extinction coefficients at 355 and 532 nm for the simulations. We consider three sets of optical data. Set 1 ($2\alpha + 3\beta$) consists of two extinction and three backscatter coefficients. Set 2 ($2\alpha + 5\beta$) contains two additional backscatter coefficients at 416 and 683 nm, which could be generated by Raman shifting the double and triple laser fundamental frequency in hydrogen. Set 3 ($2\alpha + 6\beta$) consists of backscatter coefficients at 355, 400, 532, 710, 800, and 1064 nm, which are the operational wavelengths of the six-wavelength aerosol lidar of the ITR.[21]

In the simulations we retrieve total number (n_t), total surface area (s_t), and total volume concentrations (v_t), as well as mean (r_{mean}) and effective (r_{eff}) radius. The latter are defined as

$$r_{\mathrm{mean}} = \frac{\displaystyle\int_{r_{\min}}^{r_{\max}} r f(r)\partial r}{\displaystyle\int_{r_{\min}}^{r_{\max}} f(r)\partial r}, \qquad r_{\mathrm{eff}} = \frac{\displaystyle\int_{r_{\min}}^{r_{\max}} r^3 f(r)\partial r}{\displaystyle\int_{r_{\min}}^{r_{\max}} r^2 f(r)\partial r}.$$

Equation (3) may be written for number (NK), surface (SK), and volume (VK) kernal functions.[28] In general, volume or surface kernel functions are preferred.[13,14,29] In this study we applied all three types of kernel function to the same data set in order to test whether one should be preferred over another.

Errors in the optical data were introduced randomly. If we discuss error ε in optical data, it means that error in g_i varies randomly from 0 to ε. For each radius $r_j = r_{\min} + j[(r_{\max} - r_{\min})/N]$, we calculate the average deviation of the retrieved size distribution from the correct one by

$$\sigma(r_j) = \left\{\frac{1}{N_{\mathrm{sol}}}\sum_{i=1}^{N_{\mathrm{sol}}}\left[f_i(r_j) - f_{\mathrm{mean}}(r_j)\right]^2\right\}^{1/2},$$

where N_{sol} is the number of solutions. The mean, maximum, and minimum solutions are then defined as

$$f_{\mathrm{mean}} = \frac{1}{N_{\mathrm{sol}}}\sum_{i=1}^{N_{\mathrm{sol}}} f_i(r_j),$$

$$f_{\max}(r_j) = f_{\mathrm{mean}}(r_j) + \sigma(r_j),$$

$$f_{\min}(r_j) = f_{\mathrm{mean}}(r_j) - \sigma(r_j).$$

The functions $f_{\max}$ and $f_{\min}$ allow one to estimate the scatter of the solutions and therefore the error of the retrieval.
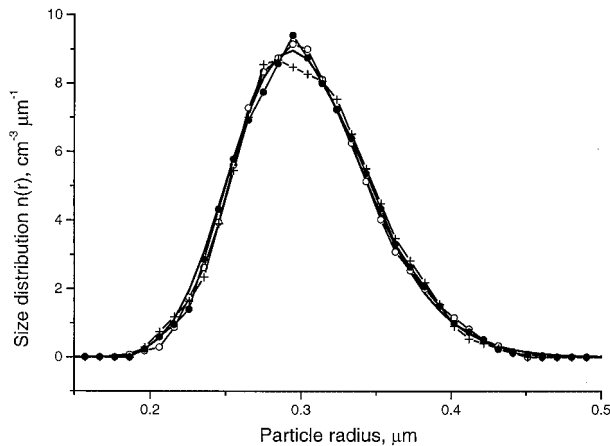
Fig. 2. Size distribution retrieved from set 1 ($2\alpha + 3\beta$, crosses), set 2 ($2\alpha + 5\beta$, filled circles), and set 3 ($2\alpha + 6\beta$, open circles). The data were assumed to be free of error ($\delta = 0$). The solid curve shows the initial log normal distribution with $r_0 = 0.3$ μm, $\ln \sigma = 0.15$ μm, and $n_t = 1$. The initial refractive index $m = 1.4 - i0.04$ was assumed to be unknown; the inversion was performed with NK density functions.

B.   Simulation Results

Figure 2 illustrates the efficiency of the minimum discrepancy criteria. For retrieval of the initial log normal distribution with $r_0 = 0.3$ μm, $\ln \sigma = 0.15$ μm, and $n_t = 1$, we used three sets of optical data. The initial refractive index was $m = 1.4 - i0.04$. In the inversion it was assumed to be unknown. Simulations have shown that, in the absence of errors in the optical data, the retrieval leads to good results for any combination of α and β, and the more coefficients that were used, the better the precision of retrieval became. All three data sets permitted the retrieval of the aerosol mean radius and number concentration to within an accuracy of better than 1%. For volume and surface area concentration the accuracy was even better. The retrieved complex refractive index coincided with the initial refractive index.

The situation changed dramatically when we included distorted optical data. Figure 3 shows the retrieval of the same initial distribution as in Fig. 2, but includes a distortion of 20% for the optical data. The errors in the data were introduced randomly, so the retrieved distribution differs for each run. We performed the procedure ten times. Figure 3 presents the run that showed the largest deviation from the initial distribution. The errors of the average parameters are the largest for data set 3, when we used six backscatter and two extinction coefficients. This means that the inversion instability cannot be suppressed just by increasing the number of used coefficients. The error of the mean radius is ~23%, for number concentration it is ~73%, for surface area concentration it is ~10%, and for volume concentration it is ~8.5%. To improve the stability of the inversion we suggest an averaging procedure, the main features of which are outlined below.
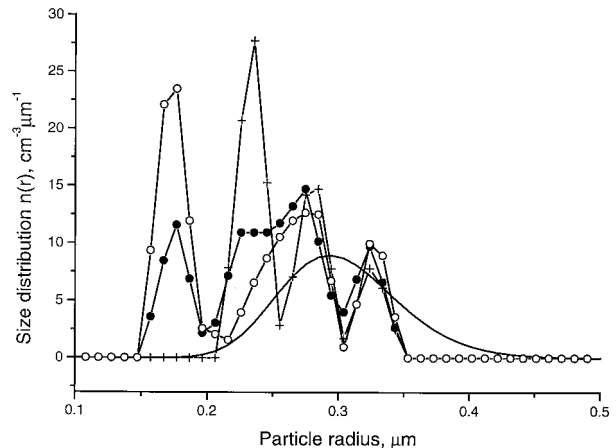

Fig. 3. Size distribution retrieved from set 1 (crosses), set 2 (filled circles), and set 3 (open circles). The data errors were assumed to be 20%. The solid curve shows the initial log normal distribution with the same parameters as in Fig. 2.

C.   Averaging of Solutions

The solution that corresponds to the minimum discrepancy can still present significant oscillations (the solution is insufficiently smoothed), which leads to errors particularly in the number concentration estimation. To minimize these oscillations, it is reasonable to consider the complete set of solutions in the vicinity of the minimum discrepancy. Such consideration should be correct, because in ill-posed problems the solutions can be numerous, so any element f^δ from the set of quasi-solutions that satisfy $\rho(f^\delta, g^\delta) \leq \delta$ should be considered as an approximate solution of the inverse problem. So we are not searching for a single solution, but for a class of solutions, which would reproduce the observations with measurement accuracy. The mean of this class can be considered as a solution (in a certain sense), and, as such, the regularization approach described in Section 2 should be considered as a tool to determine the class of solutions.

Figure 4 demonstrates the advantage of an averaging procedure for the example of six backscatter and two extinction coefficients. The plot shows the solution that corresponds to the minimum discrepancy f^δ (ρ_{min}) and the solution f^δ (ρ_{aver}) averaged over the interval [ρ_{min}, $\rho_{max} = 10\%$]. This interval includes approximately 500 individual solutions. The initial distribution was log normal with $r_0 = 0.4$ μm, $\ln \sigma = 0.3$ μm, and $n_t = 1$. The complex refractive index $m = 1.45 - i0.02$ was assumed to be unknown in the inversion. The mean distortion of the optical data was 10%. If the averaging procedure is not included, the solution f^δ (ρ_{min}) differs significantly from the initial log normal distribution, but after averaging the retrieved distribution f^δ (ρ_{aver}) comes close to the initial one. The dashed curves show the mean deviation of individual solutions (f_{max} and f_{min}) from the average value f_{mean}. The averaging procedure takes advantage of the fact that different solutions can have oscillations of opposite sign, and that,
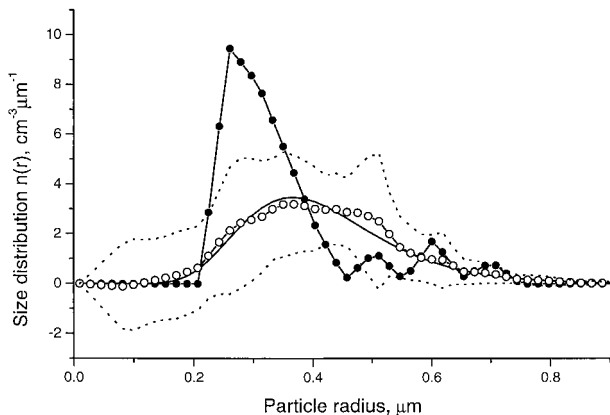
Fig. 4. Solutions that correspond to the minimum discrepancy (filled circles) and averaged over the interval of $2\% < \rho < 10\%$ (open circles). The distortion of the optical data is 10%. The solid curve shows the initial log normal distribution with $r_0 = 0.4$ μm, ln $\sigma = 0.3$ μm, and $n_t = 1$. Set three of optical data were used. The initial refractive index $m = 1.45 - i0.02$ was assumed to be unknown. The inversion was performed with VK functions. The dashed curves indicate the mean-square deviation of individual solutions from the average value.

after averaging, the mean oscillation can become much smaller than the oscillations of individual solutions. It should be mentioned that, even for a small number of basis functions, the obtained distribution is smooth, because for every value of r_{min} and r_{max} the program defines the new position of the B-spline functions. Taking the average of many individual solutions smooths the mean solution.

The criterion of the choice of averaging interval is illustrated in Fig. 5, where the number of solutions N_{sol} inside the interval is plotted versus ρ_{max}. The interval [4%, 10%] is characterized by the highest solution density and contains approximately 10% of the total number of solutions. On the basis of this criterion the choice of averaging interval does not demand any *a priori* information. Even more im-
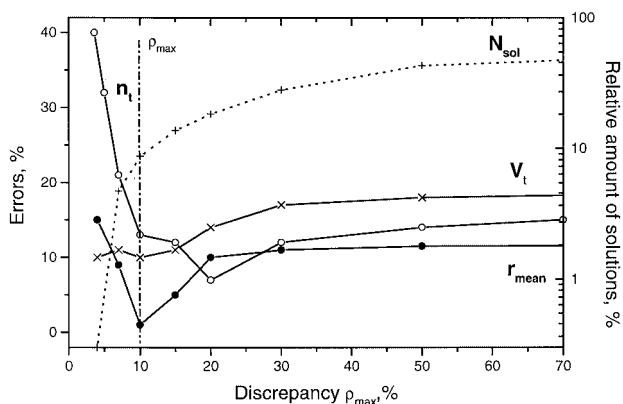

Fig. 5. Errors of number (n_t) and volume concentration (v_t) and of mean radius r_{mean} versus the averaging interval ρ_{max}. The dotted curve shows the relative number of solutions N_{sol} inside the averaging interval. The dash–dot vertical line represents the boundary of the averaging interval used for the retrieval.

portant, this procedure is stable, i.e., the result obtained does not depend strongly on the interval size. The averaging over the interval [4%, 20%] leads to practically the same size distribution. The interval size depends on data errors, number of basis functions, and kernel type. Usually we choose ρ_{max} such that approximately 10% of the total number of solutions are inside the averaging interval. Figure 5 also shows the uncertainty of number concentration n_t, volume concentration v_t, and mean radius r_{mean} as a function of ρ_{max}. The error of r_{mean} is approximately 15% for $\rho_{min} = 4\%$ and decreases to 1% when the averaging is done for the [4%, 10%] interval. For large ρ the error increases to only 10%, which shows that the mean radius does not depend strongly on ρ_{max}. The averaging procedure is especially important for the evaluation of particle number concentration n_t, which in previous studies has been shown to be affected with the largest errors.[13] The solution f^δ (ρ_{min}) leads to an error of 40% in number concentration, but after averaging this value decreases to 12%. In fact, the number concentration also shows good stability with respect to the choice of ρ_{max}. Particle surface area and volume concentrations are usually more stable compared with number concentration. The behavior of v_t and s_t for the given example is similar, and the uncertainty in both cases is approximately 10%.

D. Refractive-Index Retrieval

One of the most remarkable features of this regularization approach is that it does not need the *a priori* knowledge of the complex refractive index. On the contrary, it allows the determination of this parameter. The possibility of refractive-index retrieval from multiwavelength lidar data was first outlined in Refs. 13 and 15. In this retrieval, the combination of particle extinction and backscatter coefficients becomes especially important. When the retrieval is performed for the known refractive index, the application of only backscattering coefficients leads to a better result. But as soon as the variation of refractive index is considered, the retrieval with only backscattering coefficients becomes unstable. The stabilizing role of extinction coefficients is illustrated in Fig. 6, where the initial log normal distribution with $r_0 = 0.15$ μm, ln $\sigma = 0.3$ μm, and $n_t = 1$ is retrieved from different sets of optical data. In the simulations the number of backscattering coefficients was set to 3, 6, or 8, and the number of extinction coefficients was 0, 1, 2, 3, or 4. The average data error was 10%, and for the initial distribution evaluation we applied the averaging technique described above. The refractive index $m = 1.45 - i0.02$ was unknown in the retrieval. The errors of the aerosol parameters retrieved are summarized in Table 1. The error of number concentration estimation was the largest when only backscattering coefficients were used, and it decreased from 92% to 38% when one extinction coefficient was added. The optimal number of extinction coefficients for this configuration is 2–3 and the corresponding error of number
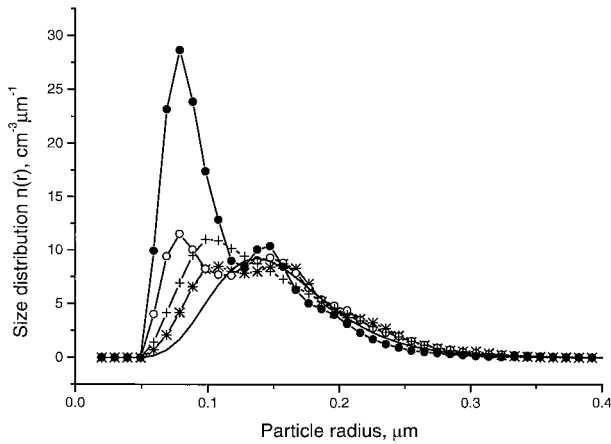
Fig. 6. Illustration of the stabilizing role of extinction coefficients. The retrieval is performed for 8β (filled circles), $8\beta + \alpha$ (open circles), $8\beta + 2\alpha$ (stars), $8\beta + 4\alpha$ (crosses). Average data distortions are 10%. The solid curve represents the initial log normal distribution with $r_0 = 0.15$ μm, ln $\sigma = 0.3$ μm and $n_t = 1$. The initial refractive index $m = 1.45 - i0.02$ was assumed to be unknown. The inversion was performed with VK functions.
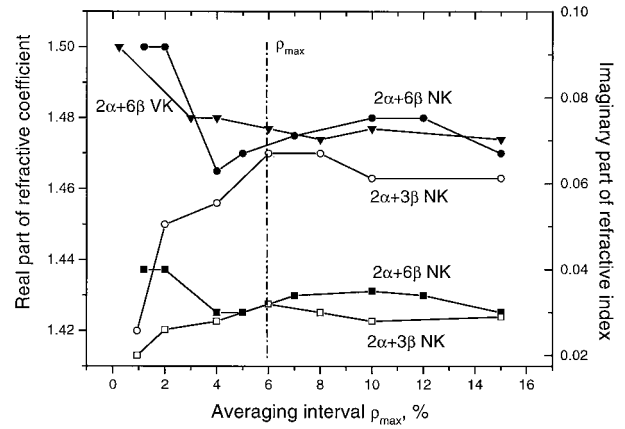

Fig. 7. Retrieval of particle complex refractive index. The real (circles and triangles) and imaginary (squares) parts are plotted versus the averaging interval ρ_{max}. For the simulation the correct refractive index $m = 1.45 - i0.02$ was used. The distortion of the optical data was 10%. For the retrieval the VK functions were applied to $2\alpha + 3\beta$ (open symbols) and to $2\alpha + 6\beta$ (filled symbols) data sets. The dash–dot vertical line represents the boundary of the averaging interval used for the retrieval. The initial size distribution was the same as in Fig. 4.

density estimation is decreased to 16%. During modeling we experimented with extinction coefficients at different wavelengths from 355 to 700 nm, and we found a stabilizing effect for each. So the combination of backscatter and extinction coefficients allows a compromise to be achieved. The results of our simulation lead to similar findings as already given in Refs. 13 and 15. The number of backscatter coefficients in the retrieval procedure should exceed the number of extinction coefficients by a factor of 2–3.

For retrieval of the complex refractive index, we used the averaging procedure in the same way as for the other aerosol parameters. But for the refractive-index retrieval we cannot use a too-large averaging interval, because for large ρ_{max} the refractive index $m \rightarrow (m_{max} + m_{min})/2$. In this case all possible solutions are averaged, and the averaging procedure loses its meaning. The averaging interval that contains 10% of solutions is usually a good compromise for both size distribution and refractive-index retrieval. Simulation results for the refractive-index retrieval are presented in Fig. 7. The correct complex refractive index in the simulation was $m = 1.45 - i0.02$. The real and imaginary parts of the refractive index are plotted as a function of ρ_{max}. For the simulation we used three sets of optical data ($2\alpha + 3\beta$, $2\alpha + 5\beta$, $2\alpha + 6\beta$) and all three types of kernel function. For

this example ρ_{min} was 1% and ρ_{max} should be taken around 7%. Approximately 10% of the total number of solutions is concentrated in this interval. In these simulations all three types of kernel function lead to similar results. Furthermore Fig. 7 shows that the retrieved refractive index for the number and the volume kernel functions is $1.47 - i0.03$. This figure again emphasizes the advantage of averaging: the solutions that correspond to the minimum discrepancy can differ for different data sets, but the averaging creates similar results.

E. Accuracy of Retrieval

To estimate the uncertainty of the retrieved parameters from experimental data we can consider the average square deviation of individual solutions. We should however keep in mind that this overestimates the uncertainty, and, as shown in Fig. 4, the actual accuracy is better after averaging. Another possibility is to perform the simulations for different mean radii and optical parameter combinations. We performed such simulations for the three sets of optical data, which are the focus of our study, i.e., $2\alpha + 3\beta$, $2\alpha + 5\beta$, $2\alpha + 6\beta$. The uncertainty depends on the mean radius. It is minimal in the center of the interval $\lambda_{min} < r_{mean} < \lambda_{max}$ and it rises

Table 1. Errors of Aerosol Parameter Retrieval for Different Combinations of Aerosol Backscatter and Extinction Coefficients

Parameter	8β (%)	8α (%)	$8\beta + \alpha$ (%)	$8\beta + 2\alpha$ (%)	$8\beta + 4\alpha$ (%)	$6\beta + 2\alpha$ (%)	$3\beta + 2\alpha$ (%)
r_{mean}	18	13	2	1	8	1	1
r_{eff}	14	1	1.4	2	1	3	3
n_t	92	58	38	16	29	18	13
S_t	11	29	20	14	15	16	18
V_t	6	29	20	14	14	19	20

Table 2. Errors ε of the Particle Parameter Estimation for Two Nd:YAG Lasers[a]

r_0 (μm)	$\varepsilon_{R\mathrm{mean}}$ (%)	$\varepsilon_{R\mathrm{eff}}$ (%)	ε_n (%)	ε_s (%)	ε_v (%)	ε_{mR}	ε_{mI} (%)
0.1	20/20	20/10	70/50	55/40	50/35	±0.08/±0.07	50/50
0.5	20/15	15/15	45/30	10/10	20/20	±0.07/±0.05	50/50
1	40/30	60/55	60/50	10/10	80/60	±0.05/±0.05	50/50

[a]One laser was a triple Nd:YAG and the second was combined with a hydrogen Raman shifter (3β + 2α/5β + 2α). Retrieval is performed with NK density functions.

near the interval boundaries. The 10% errors were introduced into the optical data. Simulation was performed for initial log normal distribution with ln σ = 0.3 μm, so $r_{\mathrm{mean}} \approx 1.1\, r_0$, and the distortion of optical data was 10%. To estimate the retrieval accuracy, the procedure was repeated twenty times for each mean radius, and the maximum deviation of the retrieved parameters from the theoretical ones was used as the measure of accuracy. The results of the simulations for different values of particle mean radii for data sets 1 and 2 are summarized in Table 2. The results obtained for set 3 are close to the results for set 2. All three types of kernel function lead to similar uncertainties. The results of the simulations shown in Table 2 can be summarized as follows:

The chosen set of wavelengths permits the retrieval of aerosol parameters for the radii 0.1 μm < r_{mean} < 1 μm.

The application of three, five, or six wavelengths, corresponding to five, seven, and eight coefficients leads to a similar uncertainty level of particle parameter estimation. But for an evaluation of the shape of the size distribution three wavelengths are insufficient. The solutions found present strong oscillations. On the other hand, an increase in the number of backscatter coefficients from six to eight did not improve the accuracy of retrieval. These findings strongly support the requirement for data combination made previously.[13–15] The retrieval of the particle complex refractive index is unstable for small particles, and for r_{mean} < 0.1 μm it becomes unbelievable. In contrast, for large particles the results are quite stable and the accuracy of the real part estimation is ±0.05. The accuracy of the retrieved imaginary part is better than 50%; see also the findings reported in Ref. 15.

The most stable parameter in the retrieval is the particle surface area concentration. The uncertainty is approximately 10% for r_{mean} > 0.1 μm.

The different kinds of kernel function lead to similar results in mean parameter estimations and size distribution shape retrieval, so we cannot conclude that one type of kernel function is preferable to another. This observation sheds new light on previous approaches in which the volume and surface area density kernels were preferred over number density kernels, because these kernels increase the sensitivity toward retrieved parameters. Simulations here indicate that the additional averaging procedure introduced sufficiently stabilizes the inverse ill-posed problem so that the number density kernel can be

used as well. However the complexity of the dependency of the solutions on individual input parameters and their uncertainties does not currently provide a final answer on this issue but rather requires further sensitivity studies.

When the 20% optical data error was considered, the uncertainty level was the same in the middle of the 0.1 μm < r_{mean} < 1 μm interval, but near the boundaries the uncertainty almost doubled. Still we believe that the Raman technique allows one to achieve a 10% accuracy in the evaluation of the optical coefficients, so the uncertainties listed in Table 2 are realistic.

4. Application to Experimental Data

The PIC developed algorithm was applied to six-wavelength lidar data obtained within the framework of the Lindenberg Aerosol Characterization Experiment 98 (LACE 98).[42] It was conducted at the Meteorological Observatory in Lindenberg (52.2 °N, 14.1 °E) during July and August 1998. Several additional surface-based and airborne lidar systems as well as characterization of particle properties from airborne platforms allowed high-quality data to be retrieved. These data were used for the validation of two other inversion schemes, i.e., from the ITR[13] and from the Institute of Mathematics (IM)[14] at the University of Potsdam, Germany. Detailed descriptions of these results are given in Refs. 37 and 43. All the lidar measurements were performed at night.

For the testing, we used two measurement cases. The first describes the measurement from 9 August 1998. A particle layer was observed in the free troposphere in the altitude range from 3 to 6 km. This layer resulted from intense biomass burning in northwestern Canada, approximately six days prior to the lidar observations. For the inversion we chose two altitude ranges, i.e., from 4200 to 5400 m and from 3500 to 4000 m. The latter describes the conditions within the center of the pollution plume. The second case was the measurement taken during the evening of 11 August 1998. In this case we observed a polluted boundary layer that resulted from advection of air masses from within the European continent.

For both measurements the results could be compared to airborne *in situ* observations of particle properties. On 9 August 1999, an aircraft flew during the time of the lidar measurement.[37,43] Because the relative humidity in this particle layer was below
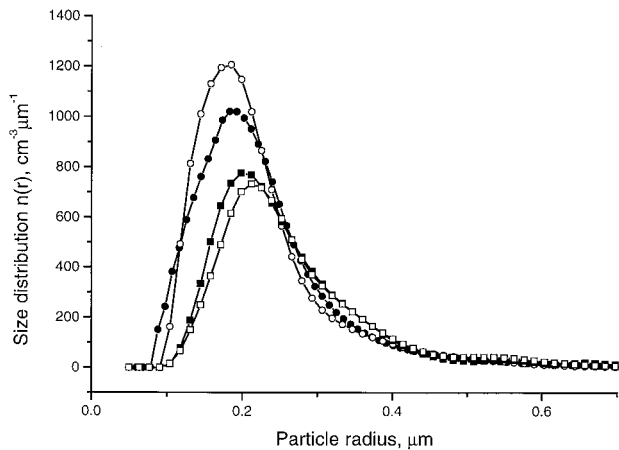
Fig. 8. Particle size distribution retrieved from the ITR multi-wavelength lidar data on 9 August 1998 for the height interval from 3500 to 4000 m. The distributions were obtained by the use of NK (squares) and VK density (circles) functions. For the inversion we used $2\alpha + 6\beta$ (filled symbols) and $2\alpha + 3\beta$ (open symbols) data sets.

50%, it was not necessary to apply correction for hygroscopic growth. On 11 August 1998, aircraft measurements were performed several hours prior to the lidar observations.[37] Because of the different measurement times backward trajectory analysis was used to identify those height ranges of the *in situ* observations, which were assumed to be comparable to the height ranges for which the lidar observations were used in the inversion. In addition the correction for hygroscopic growth, which was necessary for this measurement case, introduced additional uncertainty. A detailed outline of the error analysis is given in Ref. 37.

In the following we focus on the height range from 3500 to 4000 m of the particle layer observed on 9 August 1998. The ITR and the IM algorithms had been tested for the complete data set, which consisted of six backscatter and two extinction coefficients ($2\alpha + 6\beta$). Another test was done for the ITR algorithm with the reduced data set of three backscatter and two extinction coefficients ($2\alpha + 3\beta$), the results of which are described in detail in Ref. 32.

Figure 8 shows the particle size distribution re-

trieved for the lidar measurements by the use of NK and VK density functions. The results obtained with the surface kernel density functions are always between the NK and the VK functions. The results of the retrieval are summarized in Table 3. Figure 9 illustrates the estimation of the particle parameters. The particle effective radius, number, surface area, and volume concentration are plotted versus the size of the averaging interval. With respect to the effective radius the minimum discrepancy for the NK for the $2\alpha + 6\beta$ set is around $\rho_{min} = 1.2\%$, and the highest solution density is inside the [1.2%, 8%] interval. For the $2\alpha + 3\beta$ set $\rho_{min} < 0.1\%$ for all kernel types. Inside the averaging interval the value of r_{eff} calculated with the VKs for a full data set smoothly varies from 0.26 to 0.31 µm so the value of r_{eff} must be taken in the middle of this interval with the uncertainty of retrieval covering the maximum and minimum values, thus giving $r_{eff} = 0.28 \pm 0.04$ µm. For NK the variation of r_{eff} inside the averaging interval is larger, from 0.27 to 0.35 µm, so $r_{eff} = 0.31 \pm 0.04$ µm. The VKs created a higher number concentration (185 ± 90 cm^{-3}) in comparison with the NK results (160 ± 80 cm^{-3}). Still the difference is within the uncertainty, which we list in Table 2 for this radius. The surface area concentration varies from 118 to 125 µm^{2}cm^{-3} for the VK and from 117 to 125 µm^{2}cm^{-3} for the NK. This parameter is the most stable one in the retrieval. The volume concentration takes values between 10.5 and 14.5 µm^{3}cm^{-3} for the different kernel functions. The real part of the refractive index changes from 1.62 to 1.52 for VKs and from 1.61 to 1.48 for NK. Although the mean values obtained with both types of kernel are close, the retrieval with the VKs is again more stable, so the uncertainty is less. The imaginary part of the refractive index is approximately 0.03 for all the kernels and data sets. The inversion of the reduced data set leads to similar parameter values. The retrieved aerosol parameters are in reasonable agreement with the results of the ITR and IM codes and *in situ* measurements, although the PIC code provides lower values of aerosol number concentration.

The observed difference in the inversion with the density of the NK and the VK was unexpected for us, because during the simulation with synthetic data

Table 3. Physical Particle Parameters[a]

| | ITR Algorithm | | IM Algorithm | PIC Algorithm | | | | Falcon, *in situ* |
| | Full Set VK | $3\beta + 2\alpha$ VK | Full Set VK | Full Set VK | $3\beta + 2\alpha$ VK | Full Set NK | $3\beta + 2\alpha$ NK | Particles with |
Parameter								$r > 50$ nm
r_{eff} (µm)	0.27 ± 0.04	0.27 ± 0.04	0.24 ± 0.01	0.28 ± 0.04	0.28 ± 0.04	0.31 ± 0.04	0.31 ± 0.04	0.25 ± 0.07
n_t (cm^{-3})	291 ± 70	305 ± 120	506 ± 131	185 ± 90	200 ± 100	160 ± 80	155 ± 80	271 ± 74
v_t (µm^3 cm^{-3})	13 ± 2	13 ± 2	11 ± 1	12 ± 3	12 ± 3	12.5 ± 3	13.5 ± 4	8 ± 5
s_t (µm^2 cm^{-3})	139 ± 7	145 ± 8	136 ± 5	122 ± 20	122 ± 20	125 ± 20	126 ± 20	95 ± 55
m_R	1.64 ± 0.09	1.63 ± 03	1.66 ± 0.02	1.57 ± 0.05	1.58 ± 0.06	1.55 ± 0.07	1.52 ± 0.07	1.56
m_I	0.05 ± 0.02	0.05 ± 0.02	0.053 ± 0.004	0.03 ± 0.015	0.03 ± 0.015	0.03 ± 0.015	0.035 ± 0.015	0.07

[a]From the inversion of lidar data obtained from measurements in the height region from 3500 to 4000 m on 9 August 1998, 2200–2400 UTC, and from *in situ* measurements of particle size distributions from 3400 to 3900 m aboard an aircraft. The NK and VK density functions are applied to the full and reduced ($3\beta + 2\alpha$) data sets, respectively.
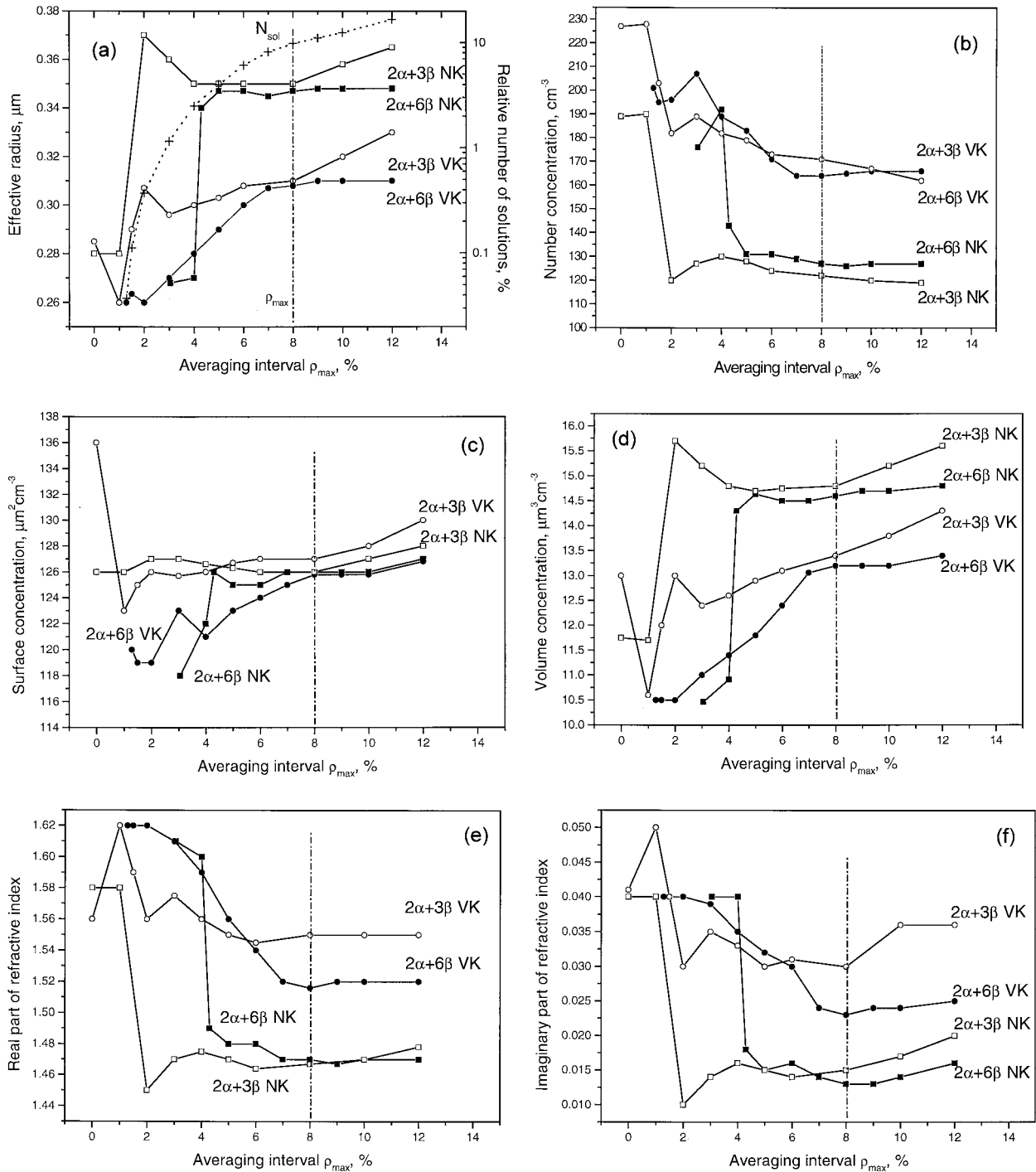
Fig. 9. Retrieval of particle parameters from six-wavelength observations on 9 August 1998 for the height interval from 3500 to 4000 m. (a) Effective radius, (b) number, (c) surface area, and (d) volume concentrations, and (e) real and (f) imaginary parts of the refractive index versus the averaging interval for solutions obtained with NK, SK, and VK density functions.   The NK applied to the full set of data (filled symbols) and to data obtained from a triple Nd:YAG laser ($2\alpha + 3\beta$, open symbols).   The dotted curve in (a) shows the relative amount of solutions N_{sol} inside the averaging interval for the NK functions.   The dash–dot vertical line represents the boundary of the averaging interval used for the retrieval.

both kernel types led to similar results. To illustrate this we have chosen the initial log normal distribution with $r_0 = 0.2$ μm, which is close to the distributions presented in Fig. 8. Figure 10 shows the size distributions retrieved with the NK and the

VK and for 10% data errors. Both types of kernel lead to similar concentrations, and the density of the NKs for this example reproduce the initial distribution even better than the VKs. One possible explanation of the observed difference in the case of the
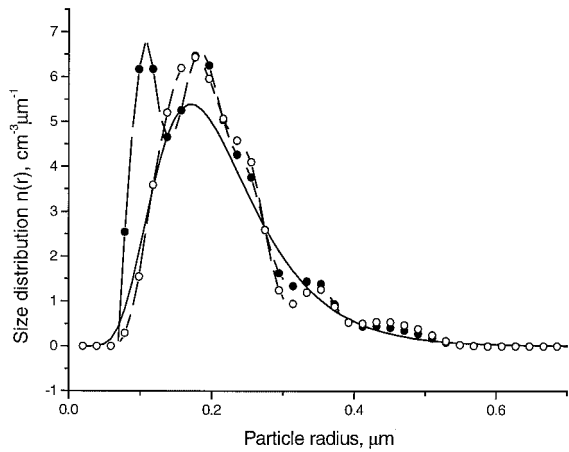
Fig. 10. Size distribution retrieved with VK (filled circles) and NK (open circles) density functions. The solid curve shows the initial log normal distribution with $r_0 = 0.2$ μm, $\ln \sigma = 0.4$ μm and $n_t = 1$. Optical data distortion is 10%, the refractive index $m = 1.54 - i0.02$ was suggested to be unknown during the inversion.

inversion of experimental data could be that we performed simulations for a single type of particle, whereas the real aerosol is a mixture of particles with different refractive indices. In fact, detailed analysis of the *in situ* data pointed toward an external rather than an internal mixture of particle types.[37] Still the observed difference is smaller than the estimated uncertainty of the method. It should also be noted, that the absence of spikes in the distributions presented in Fig. 8 and the concentration of solutions in the interval with $\rho_{\max} < 8\%$ indicates that the mean error for these measurements is less than 10%.

The results obtained from the reduced data set for both the PIC and the ITR codes do not differ significantly from those obtained with the complete data set. However we note a slight increase of the uncertainty of the retrieved parameters in each case. It is worth mentioning that the same result, i.e., an increase of the uncertainty, has been found from the ITR algorithm. On the one hand the results for the experimental data further strengthen the argument that the use of a Raman lidar, which uses only one Nd:YAG laser, might be sufficient to provide important particle parameters.[35] On the other hand the increased uncertainty levels call for further investi-

gation of the general applicability of the inversion schemes to the reduced data sets.

Table 4 summarizes the results obtained for the height range from 4400 to 5000 m on 9 August 1998. Again NK and VK were used for inversion of the complete and the reduced data set. The respective parameters obtained from the different kernel types and data combinations differ insignificantly within the uncertainties. The effective radius obtained with the NK is approximately 0.112 μm. With the VK it is approximately 0.110 μm. The number concentration shows a larger difference. It is 850 cm^{-3} for the VKs and 700 cm^{-3} for the NKs. The surface area concentration varies between 95 and 100 μm^{2}cm^{-3}. The volume concentration takes values between 3.5 and 3.7 μm^{3}cm^{-3} for the different kernel functions. The real part of the refractive index changes from 1.58 to 1.60, the imaginary part presents significant uncertainty, and we estimate it to be 0.01 ± 0.01.

With respect to the size parameters all three algorithms show comparable values within the uncertainty for the lower height range on 9 August 1998. The largest scatter was found for number concentration for which the IM algorithm gives considerably higher values than the PIC algorithm. The ITR values lie between the two. The ITR and IM algorithms give the same complex refractive index. The PIC algorithm underestimates this parameter, particularly with respect to the imaginary part. The *in situ* data were restricted to particle sizes above 50 nm in radius, which is at the lower end of the optically active size range. In general the *in situ* size parameters and the real part of the complex refractive index are lower than the ones from the inversion. Yet there is acceptable agreement. In contrast, a larger imaginary part was derived from the *in situ* observations. In this context we should again mention that the complex refractive index from the *in situ* observations was derived on the assumption of an internal mixture of particles. In fact, the observations showed an external mixture. Further details can be found in Ref. 37. The mean deviation of all the parameters derived with the PIC algorithm for the *in situ* values is approximately 30–35%. The lowest deviation of 27% is given by the ITR algorithm for the lower height range. It is 34% for the IM algorithm.

Table 4. Same as Table 3 but for Lidar Measurements from 4200 to 5400 m and *in situ* Data from 4400 to 5000 m

| Parameter | ITR Algorithm Full Set VK | IM Algorithm Full Set VK | PIC Algorithm | | | Falcon, *in situ* Particles with $r > 50$ nm |
			Full Set NK	$3\beta + 2\alpha$ NK	Full Set VK	
r_{eff} (μm)	0.23 ± 0.02	0.16 ± 0.01	0.11 ± 0.02	0.11 ± 0.03	0.11 ± 0.02	0.18 ± 0.05
n_t (cm^{-3})	663 ± 224	973 ± 316	700 ± 350	750 ± 400	850 ± 400	354 ± 74
v_t (μm^{3}cm^{-3})	3.4 ± 0.4	2.3 ± 0.2	3.5 ± 1	3.7 ± 1	3.7 ± 1	6 ± 3
s_t (μm^{2}cm^{-3})	44 ± 1	44 ± 3	95 ± 40	100 ± 50	100 ± 40	91 ± 53
m_R	1.76 ± 0.05	1.77 ± 0.02	1.60 ± 0.07	1.58 ± 0.08	1.59 ± 0.07	—
m_I	0.046 ± 0.005	0.043 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	—

| Parameter | ITR Algorithm Full Set VK | IM Algorithm Full Set VK | PIC Algorithm | | | Partenavia *in situ* $r > 50$ nm | |
			Full Set NK	$3\beta + 2\alpha$ NK	Full Set VK	Dry	75% Humidity
r_{eff} (μm)	0.12 ± 0.02	0.11 ± 0.002	0.19 ± 0.02	0.2 ± 0.02	0.17 ± 0.02	0.17 ± 0.03	0.24 ± 0.05
n_t (cm^{-3})	—	—	1860 ± 900	1430 ± 700	2940 ± 1500	858 ± 56	858 ± 56
v_t (μm^{3}cm^{-3})	35 ± 7	23 ± 1	34 ± 8	34 ± 8	33 ± 8	11 ± 3	32 ± 8
s_t (μm^{2}cm^{-3})	921 ± 221	624 ± 30	535 ± 200	515 ± 200	590 ± 200	194 ± 54	400 ± 130
m_R	1.58 ± 0.1	1.74 ± 0.02	1.51 ± 05	1.52 ± 0.05	1.61 ± 0.07	1.53	
m_I	0.01 ± 0.02	0.031 ± 0.0001	0.01 ± 0.01	0.01 ± 0.01	0.03 ± 0.015	0.02	

[a]From inversion of lidar data obtained from measurement in the height region from 1420 to 1700 m on 11 August 1998, 2020–2220 UTC, and from *in situ* measurements of particle size distributions from 2000 to 2400 m at 1230 UTC.

With respect to the upper height range large spatial inhomogeneities of the volume concentration distributions were found in Ref. 42. Time–height plots of backscatter coefficients from lidar also show large inhomogeneities at this height range.[37] Complex refractive indices from *in situ* data were not provided for this height range. So a comparison for this parameter is possible only among the three inversion results. Large differences between the three algorithms were found for effective radius and surface area concentration and for the imaginary part of the complex refractive index. Again the PIC algorithm underestimates the complex refractive indices provided by the ITR and the IM. The mean deviation for the *in situ* values increases to 46–57% for the PIC algorithm. The deviation for the ITR algorithm increases to 53%, and the deviation for the IM increases to 75%.

Table 5 shows the results for the height range from 1420 to 1700 m of the measurement from 11 August 1998. Comparison with the *in situ* parameters is complicated by the fact that a correction for hygroscopic growth had to be applied. If we consider the results only for 75% relative humidity, the parameters of the PIC algorithm deviate by less than 50% for effective radius, volume, and surface area concentration, and the real part of the complex refractive index. The imaginary part is a factor of 2 lower in the case of inversion with NK density, but 50% larger if VKs are used. The number concentration is a factor of 2–3 larger than the *in situ* values if the two kernel types are used in the inversion.

The PIC algorithm provided the lowest mean devi-ation of approximately 23–27% for the *in situ* values. The ITR algorithm showed a mean deviation of 49%, and the IM algorithm deviated by approximately 41%. For calculation of the mean deviation, number concentration was not included because reasonable number concentrations could not be provided by the other two algorithms. The uncertainties were larger than the respective mean values that reached several times 10,000 cm^{-3}, indicating a failure of the regularization.[37] Only the PIC algorithm provided somewhat reasonable values for this parameter.

Table 6 shows the results for the lower height range on 11 August 1998. The PIC algorithm gives lower volume, surface area, and number concentrations than the ITR and the IM algorithms. Number concentration for this example is 1 order of magnitude lower. The real part is between the ITR and the IM values. The imaginary parts in all three algorithms are close. The best mean performance for this height range was achieved with the PIC algorithm. The number concentration was especially close to the *in situ* data. The mean deviation was 53–57%. The number concentration was again omitted in these calculations for the reasons given above.

5. Conclusion

We have developed a rather simple modification of a regularization algorithm, which is based on the minimum discrepancy criterion. The code retrieves particle size parameters from lidar observations at multiple wavelengths. Computer simulations demonstrate that application of the algorithm even to a

Table 6. Same as Table 5 but for Lidar Measurements from 900 to 1000 m and *in situ* Measurements from 500 to 900 m

| Parameter | ITR Algorithm Full Set VK | IM Algorithm Full Set VK | PIC Algorithm | | | Partenavia *in situ* $r > 50$ nm | |
			Full Set NK	$3\beta + 2\alpha$ NK	Full Set VK	Dry	40% Humidity
r_{eff} (μm)	0.13 ± 0.02	0.15 ± 0.02	0.14 ± 0.02	0.14 ± 0.02	0.14 ± 0.02	0.14 ± 0.01	0.16 ± 0.01
n_t (cm^{-3})	11193 ± 6609	18455 ± 5312	1450 ± 700	1440 ± 700	1740 ± 800	607 ± 89	607 ± 89
v_t (μm^{3}cm^{-3})	29 ± 3	18 ± 1	13 ± 4	12 ± 4	13.5 ± 5	5 ± 1	7 ± 1
s_t (μm^2 cm^{-3})	696 ± 38	616 ± 45	280 ± 50	250 ± 50	300 ± 60	100 ± 19	131 ± 20
m_R	1.48 ± 0.03	1.62 ± 0.02	1.56 ± 05	1.6 ± 0.07	1.56 ± 0.05	1.53	
m_I	0.008 ± 0.003	0.008 ± 0.003	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.02	

simplified set of lidar data obtained from a triple Nd:YAG laser, which might provide backscatter coefficients at 355, 532, and 1064 nm and extinction coefficients at 355 and 532 nm, allows the estimation of important microphysical particle parameters, i.e., effective radius, volume, surface area, and number concentration, and the complex refractive index within the radius range from 0.1 to 1 μm. The development of such a relatively inexpensive lidar equipment with the capability to estimate these particle parameters significantly extends the area of lidar application. The retrieval accuracy might be improved if we combine a triple Nd:YAG laser with a hydrogen Raman shifter, which would add two more backscatter coefficients at 416 and 683 nm.

The common approach in the inversion with regularization is to select only those solutions for which the regularization or minimum discrepancy term takes its minimum. In the modification developed here we average all individual solutions that are within a certain range around this minimum solution. As a consequence, the stability of the inversion, particularly with respect to number concentration, could be increased significantly compared with current approaches. In the simulations we also tested the influence of the number, surface area, and VK functions on the inversion results. These functions are input parameters for the algorithm. They describe backscatter or extinction for single particles. In general we observed no significant differences of the kernel functions for the different data combinations used in this study. Still we applied both NK and VK functions to each data set, and the uncertainty of retrieval covered the difference in the results that we obtained.

We compared the results of this algorithm with those of two other regularization codes, which are in routine operational use, and with results of aircraft *in situ* measurements. The optical data sets, which were obtained within the framework of the LACE 98, describe a biomass-burning particle layer and an anthropogenically polluted boundary layer. The quite different particle properties provided a benchmark test for the performance capabilities of the inversion. For the biomass-burning layer the presented algorithm showed good agreement with the results from the other two algorithms. Mean deviation to the *in situ* observation ranged from 30 to 57% and thus was within the range given by the other two algorithms. For the polluted boundary layer we found larger discrepancies than were found with the other two inversion schemes. The average deviation from the *in situ* observations was 23–57% and thus outperformed the results of the other two schemes. This was particularly the case for number concentration and might be the result of the additional averaging of solutions. However, we also noted a systematic underestimation of the complex refractive index. In general, the imaginary part was lower by a factor of 2 compared with the *in situ* data. We observed a

maximum deviation of a factor of 4 compared with the results from the other two inversion schemes.

Our main goal was to establish the utility of the averaging procedure to determine the final solution, which was done for a monomodal size. We also performed preliminary retrievals for bimodally distributed aerosols indicating that the integral parameters, such as effective radius and the densities (number, volume, and surface), could be estimated. However, the accurate retrieval of a bimodal size distribution places a strong requirement on the quality and number of the input data and is the subject of future research.

The results presented here strengthen the technique for remote quantification of aerosols by use of lidar. These measurements provide new information about aerosol properties that would be useful for the study of the influence of aerosols on climate. It is interesting to point out as well that cloud liquid water measurements have also been made by use of Raman lidar.[44,45] A retrieval technique has been demonstrated that uses Raman lidar measurements to quantify remotely cloud liquid water, cloud droplet radius, and cloud droplet number density.[46] The combination of these cloud measurements with those of the aerosols demonstrated here could permit all essential parameters that are necessary for the study of the indirect effect of aerosols to be measured with Raman lidar.

References

1. S. Twomey, "Influence of pollution on shortwave albedo of clouds," J. Atmos. Sci. **34,** 1149–1152 (1977).
2. J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, and D. Xiaosu, eds., "Third Assessment Report of Working Group I of the Intergovernmental Panel on Climate Change" (Cambridge University, Cambridge, England, 2001).
3. S. Twomey, *Introduction to the Mathematics of Inversion in Remote Sensing and Direct Measurements* (Elsevier, New York, 1977).
4. A. N. Tikhonov and V. Y. Arsenin, eds., *Solution of Ill-Posed Problems* (Wiley, New York, 1977).
5. C. D. Rodgers, "Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation," Rev. Geophys. Space Phys. **14,** 609–624 (1976).
6. M. D. King, "Sensitivity of constrained linear inversions to the selection of the Lagrange multiplier," J. Atmos. Sci. **39,** 1356–1369 (1982).
7. A. Tarantola, *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation* (Elsevier, Amsterdam, 1987).
8. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN; the Art of Scientific Computing* (Cambridge University, Cambridge, England, 1992).
9. O. Dubovik and M. King, "A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements," J. Geophys. Res. **105,** D16, 20673–20696 (2000).
10. V. E. Zuev and I. E. Naats, eds., *Inverse Problems of Lidar Sensing of the Atmosphere* (Springer-Verlag, Berlin, 1983).
11. J. D. Klett, "Stable analytical inversion solution for processing lidar returns," Appl. Opt. **20,** 211–220 (1981).
12. J. Hadamard, "Sur les problémes aux derivees parielies et leur signification physique," Bull. Univ. Princeton 49–52 (1929).

13. D. Müller, U. Wandinger, and A. Ansmann, "Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: theory," Appl. Opt. **38,** 2346–2357 (1999).

14. C. Böckmann, "Hybrid regularization method for the ill-posed inversion of multiwavelength lidar data in the retrieval of aerosol size distributions," Appl. Opt. **40,** 1329–1342 (2001).

15. D. Müller, U. Wandinger, and A. Ansmann, "Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: simulation," Appl. Opt. **38,** 2358–2368 (1999).

16. G. Beyerle, R. Neuber, O. Schrems, F. Wittrock, and B. Knudsen, "Multiwavelength lidar measurements of stratospheric aerosols above Spitsbergen during winter 1992/93," Geophys. Res. Lett. **21,** 57–60 (1994).

17. G. Feingold and C. J. Grund, "Feasibility of using multiwavelength lidar measurements to measure cloud condensation nuclei," J. Atmos. Oceanic Technol. **11,** 543–1558 (1994).

18. B. Stein, M. Del Guasta, J. Kolenda, M. Morandi, P. Rairoux, L. Stefanutti, J. P. Wolf, and L. Wöste, "Stratospheric aerosol size distribution from multispectral lidar measurements at Sodankylä during EASOE," Geophys. Res. Lett. **21,** 1311–1314 (1994).

19. U. Wandinger, A. Ansmann, J. Reichardt, and T. Deshler, "Determination of stratospheric aerosol microphysical properties from independent extinction and backscattering measurements with a Raman lidar," Appl. Opt. **34,** 8315–8329 (1995).

20. M. J. Post, "A graphical technique for retrieving size distribution parameters from multiple measurements: visualization and error analysis," J. Atmos. Oceanic Technol. **13,** 863–873 (1996).

21. D. Althausen, D. Müller, A. Ansmann, U. Wandinger, H. Hube, E. Clauder, and S. Zörner, "Scanning 6-wavelength 11-channel aerosol lidar," J. Atmos. Oceanic Technol. **17,** 1469–1482 (2000).

22. D. N. Whiteman, S. H. Melfi, and R. A. Ferrare, "Raman lidar system for measurement of water vapor and aerosols in the Earth's atmosphere," Appl. Opt. **31,** 3068–3082 (1992).

23. A. Ansmann, M. Riebesell, U. Wandinger, C. Weitkamp, E. Voss, W. Lahmann, and W. Michaelis, "Combined Raman elastic-backscatter lidar for vertical profiling of moisture, aerosols extinction, backscatter, and lidar ratio," Appl. Phys. B **55,** 18–28 (1992).

24. R. A. Ferrare, S. H. Melfi, D. N. Whiteman, K. D. Evans, M. Poellot, and Y. J. Kaufman, "Raman lidar measurements of aerosol extinction and backscattering. Derivation of aerosol real refractive index, single-scattering albedo, and humidification factor using Raman lidar and aircraft size distribution measurements," J. Geophys. Res. **103,** D16, 19673–19689 (1998).

25. A. Ansmann, U. Wandinger, M. Riebesell, C. Weitkamp, and W. Michaelis, "Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar," Appl. Opt. **31,** 7113–7131 (1992).

26. A. Ansmann, M. Riebesell, and C. Weitkamp, "Measurement of atmospheric aerosol extinction profiles with a Raman lidar," Opt. Lett. **15,** 746–748 (1990).

27. D. Müller, F. Wagner, U. Wandinger, A. Ansmann, M. Wendisch, D. Althausen, and W. von Hoyningen-Huene, "Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: experiment," Appl. Opt. **39,** 1879–1892 (2000).

28. J. Heintzenberg, H. Müller, H. Quenzel, and E. Thomalla, "Information content of optical data with respect to aerosol properties: numerical studies with a randomized minimization-search-technique inversion algorithm," Appl. Opt. **20,** 1308–1315 (1981).

29. D. P. Donovan and A. I. Carswell, "Principal component analysis applied to multiwavelength lidar aerosol backscatter and extinction measurements," Appl. Opt. **36,** 9406–9424 (1997).

30. V. V. Veretennikov, V. S. Kozlov, I. E. Naats, and V. Ya. Fadeev, "Optical studies of smoke aerosols: an inversion method and its applications," Opt. Lett. **4,** 411–413 (1979).

31. A. P. Ivanov, F. P. Osipenko, A. P. Chaykovskiy, and V. N. Shcherbakov, "Study of the aerosol optical properties and microstructure by the method of multiwave sounding," Izv. Acad. Sci. USSR Atmos. Oceanic Phys. **22,** 633–639 (1986).

32. P. Qing, H. Nakane, Y. Sasano, and S. Kitamura, "Numerical simulation of the retrieval of aerosol size distribution from multiwavelength laser radar measurements," Appl. Opt. **28,** 5259–5265 (1989).

33. D. Müller, F. Wagner, D. Althausen, U. Wandinger, and A. Ansmann, "Physical properties of the Indian aerosol plume derived from six-wavelength lidar observation on 25 March 1999 of the Indian Ocean Experiment," Geophys. Res. Lett. **27,** 1403–1406 (2000).

34. A. Ansmann, D. Althausen, U. Wandinger, K. Franke, D. Müller, F. Wagner, and J. Heintzenberg, "Vertical profiling of Indian aerosol plume with six-wavelength lidar during INDOEX: a first case study," Geophys. Res. Lett. **27,** 963–966 (2000).

35. D. Müller, U. Wandinger, D. Althausen, and M. Fiebig, "Comprehensive particle characterization from three-wavelength Raman-lidar observations: case study," Appl. Opt. **34,** 4863–4869 (2001).

36. J. Bösenberg, A. Ansmann, J. M. Baldasano, D. Balis, C. Böckmann, B. Calpini, A. Chaikovsky, P. Flamant, A. Hågård, V. Mitev, A. Papayannis, J. Pelon, D. Resendes, J. Schneider, N. Spinelli, T. Trickl, G. Vaughan, G. Visconti, and M. Wiegner, "EARLINET: a European Aerosol Research Lidar Network," in *Laser Remote Sensing of the Atmosphere. Selected papers of the 20th International Laser Radar Conference, Vichy, France*, A. Dabas, C. Loth, and J. Pelon, eds. (Ecole Polytechnique, Paris, France, 2001).

37. U. Wandinger, D. Müller, C. Böckmann, D. Althausen, V. Matthias, J. Bösenberg, V. Weiss, M. Fiebig, M. Wendisch, A. Stohl, and A. Ansmann, "Optical and microphysical characterization of biomass-burning and industrial-pollution aerosols from multiwavelength lidar and aircraft measurements," J. Geophys. Res., accepted for publication.

38. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983).

39. G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," Technometrics **21,** 215–223 (1979).

40. F. O'Sullivan, "A statistical perspective on ill-posed inverse problems," Stat. Sci. **1,** 502–527 (1986).

41. P. C. Sabatier, "Basic concepts and methods of inverse problems," in *Basic Methods of Tomography and Inverse Problems*, P. C. Sabatier, ed. (Hilger, London, 1987).

42. A. Ansmann, U. Wandinger, A. Wiedensohler, and U. Leiterer, "Lindenberg Aerosol Characterization Experiment 1998 (LACE 98): overview," J. Geophys. Res., accepted for publication.

43. M. Fiebig, A. Petzold, U. Wandinger, M. Wendisch, C. Kiemle, A. Stifter, M. Ebert, T. Rother, and U. Leiterer, "Optical closure for an aerosol column: method, accuracy, and inferable properties applied to a biomass-burning aerosol and its radiative forcing," J. Geophys. Res., accepted for publication.

44. S. H. Melfi, K. D. Evans, J. Li, D. Whiteman, R. Ferrare, and G. Schwemmer, "Observation of Raman scattering by cloud droplets in the atmosphere," Appl. Opt. **36,** 3551–3559 (1997).

45. I. A. Veselovskii, H. K. Cha, D. H. Kim, S. C. Choi, and J. M. Lee, "Raman lidar for the study of liquid water and water vapor in the troposphere," Appl. Phys. B **71,** 113–117 (2000).

46. D. N. Whiteman and S. H. Melfi, "Cloud liquid water, mean droplet radius and number density measurements using a Raman lidar," J. Geophys. Res. **104,** D24, 31411–31419 (1999).