Atmospheric
Measurement
Techniques

Microphysical particle properties derived from inversion algorithms developed in the framework of EARLINET

Detlef Müller[1], Christine Böckmann[2], Alexei Kolgotin[3], Lars Schneidenbach[a], Eduard Chemyakin[4], Julia Rosemann[2], Pavel Znak[5], and Anton Romanov[6]

[1]School of Physics, Astronomy and Mathematics, University of Hertfordshire, Hatfield, Hertfordshire, UK
[2]Institute of Mathematics, University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany
[3]Physics Instrumentation Center, Troitsk, Russia
[4]Science Systems and Applications, Inc., NASA Langley Research Center, Hampton VA, USA
[5]V. A. Fock Institute of Physics, St. Petersburg University, Ulyanovskaya 1, 198504 St. Petersburg, Russia
[6]The National University of Science and Technology, Moscow, Russia
[a]formerly at: Institute for Computer Science, University of Potsdam, Am Neuen Palais 10,
14469 Potsdam, Germany

Correspondence to: Detlef Müller (d.mueller@herts.ac.uk)

Abstract. We present a summary on the current status of two inversion algorithms that are used in EARLINET (European Aerosol Research Lidar Network) for the inversion of data collected with EARLINET multiwavelength Raman lidars. These instruments measure backscatter coefficients at 355, 532, and 1064 nm, and extinction coefficients at 355 and 532 nm. Development of these two algorithms started in 2000 when EARLINET was founded. The algorithms are based on a manually controlled inversion of optical data which allows for detailed sensitivity studies. The algorithms allow us to derive particle effective radius as well as volume and surface area concentration with comparably high confidence. The retrieval of the real and imaginary parts of the complex refractive index still is a challenge in view of the accuracy required for these parameters in climate change studies in which light absorption needs to be known with high accuracy. It is an extreme challenge to retrieve the real part with an accuracy better than 0.05 and the imaginary part with accuracy better than 0.005–0.1 or ±50 %. Single-scattering albedo can be computed from the retrieved microphysical parameters and allows us to categorize aerosols into high- and low-absorbing aerosols.

On the basis of a few exemplary simulations with synthetic optical data we discuss the current status of these manually operated algorithms, the potentially achievable accuracy of

data products, and the goals for future work. One algorithm was used with the purpose of testing how well microphysical parameters can be derived if the real part of the complex refractive index is known to at least 0.05 or 0.1. The other algorithm was used to find out how well microphysical parameters can be derived if this constraint for the real part is not applied.

The optical data used in our study cover a range of Ångström exponents and extinction-to-backscatter (lidar) ratios that are found from lidar measurements of various aerosol types. We also tested aerosol scenarios that are considered highly unlikely, e.g. the lidar ratios fall outside the commonly accepted range of values measured with Raman lidar, even though the underlying microphysical particle properties are not uncommon. The goal of this part of the study is to test the robustness of the algorithms towards their ability to identify aerosol types that have not been measured so far, but cannot be ruled out based on our current knowledge of aerosol physics.

We computed the optical data from monomodal logarithmic particle size distributions, i.e. we explicitly excluded the more complicated case of bimodal particle size distributions which is a topic of ongoing research work. Another constraint is that we only considered particles of spherical shape in our simulations. We considered particle radii as large as

7–10 µm in our simulations where the Potsdam algorithm is limited to the lower value. We considered optical-data errors of 15 % in the simulation studies. We target 50 % uncertainty as a reasonable threshold for our data products, though we attempt to obtain data products with less uncertainty in future work.

1   Introduction

The start of EARLINET (European Aerosol Research Lidar Network) in 2000 marked the beginning of the development of inversion algorithms that can be used for the retrieval of aerosol microphysical properties from Raman lidar observations. Based on exploratory work (Qing et al., 1989) it seemed feasible that Raman lidar measurements of backscatter coefficients and extinction coefficients could provide variables such as particle effective radius and complex refractive index from which, under favourable measurement conditions, single-scattering albedo could be derived.

We followed two conceptual approaches. One methodology was developed at the Leibniz Institute for Tropospheric Research (TROPOS), Leipzig, Germany. The development of the methodology continues at the University of Hertfordshire (UH). The second method was developed at the University of Potsdam (UP) (Potsdam, Germany). Both methods in part follow the same basic mathematical concepts in the sense that they are true inversion algorithms. True inversion algorithms mean the following: (a) the underlying mathematical equations that connect the microphysical particle properties (which are the variables we are interested in) and the optical properties (which are the variables that are measured with lidar) are solved explicitly; (b) therefore we do not carry out forward computations, which are commonly referred to as Mie-scattering computations; (c) we do not use traditional look-up tables that contain an array of microphysical aerosol properties and the optical properties that belong to these microphysical properties; and (d) our approach neglects constraints that are used in forward computations, for example the need to prescribe the shape of the particle size distribution as input.

The advantage of our approach is that we can identify the share of fine-mode and coarse-mode particles in particle size distributions, as the inversion algorithms allow us to find approximate solutions of the underlying particle size distributions. As with any other method, there exists plenty of disadvantages, for example measurement errors have a direct impact on the quality of the retrieval results. If measurement errors become too large, the inversion algorithms will not be able to find reasonable solutions. The inversion algorithms also respond strongly to systematic errors of the optical input data. This means that if calibrations of the optical profiles are not done carefully, or if optical data are faulty because of the incomplete overlap between laser beam and field-of-view

of the receiver telescope (Wandinger and Ansmann, 2002), the inversion algorithms will deliver the wrong microphysical particle properties.

1.1   Work at TROPOS/UH

With regard to work at TROPOS/UH, our algorithm development began on the basis of data we obtained from the first truly operational multiwavelength Raman lidar (Althausen et al., 2000) in the 1990s. This instrument provides backscatter coefficients measured at 355, 400, 532, 710, 800, and 1064 nm and extinction coefficients at 355 and 532 nm. The system uses four lasers that operate simultaneously. The high costs of this system and the labour-intensive measurement and data analysis work showed us fairly early that a wider use of the inversion technology in the lidar community would require reducing the optical data needed for data inversion. Based on exploratory work (Müller et al., 2001b; Veselovskii et al., 2002; Böckmann et al., 2005), it seemed feasible that a Raman lidar consisting of only one Nd:YG laser could still fulfil the minimum number of optical data such that successful data inversion could be carried out. This minimum requirement is measurements of backscatter coefficients at 355, 532, and 1064 nm and extinction coefficients at 355 and 532 nm.

The algorithm that was initially developed at TROPOS follows the concept of Tikhonov's inversion with regularization (Tikhonov and Arsenin, 1977). The algorithm is based on a data-operator-controlled software environment. That makes this algorithm highly labour-intensive and the output with regard to inversion results is rather low. However, this approach allows us to carry out detailed sensitivity studies in order to test the quality of the data inversion products and to push the envelope in what can theoretically be achieved in terms of aerosol characterization with state-of-the-art multiwavelength Raman lidar. Several sensitivity studies dealt with the ability of the algorithm to retrieve effective radius, and surface area and volume concentration, and the complex refractive index (CRI) (Müller et al., 1999a, b, 2001b; Veselovskii et al., 2002, 2004). Most sensitivity studies dealt with monomodal particle size distributions, although some work has also been done in the context of bimodal particle size distributions (Veselovskii et al., 2004).

The retrieval of the CRI remains a major challenge in our work. The accuracy requirements for the imaginary part of the CRI are considerable if we want to obtain useful values (high accuracy and precision) of the single-scattering albedo (SSA) which is one of the key parameters in climate change studies. Mishchenko et al. (2004) mention that an accuracy of 0.03 and a precision of 0.02 of the SSA in the wavelength range between 350 and 2500 nm is needed in order to meet the requirements for sensible impact studies of particulate pollution on radiative forcing.

Nowadays we manage to obtain meaningful values of the SSA, but at the expense of time-intensive, data-operator-

controlled data analysis which involves a careful evaluation of the inversion results (particularly of the CRI). We learnt to retrieve the imaginary part to its correct order of magnitude if it is less than $0.01i$. This value describes moderately light-absorbing aerosol. We learnt to keep the uncertainty bars to approximately $\pm 50\%$ if the imaginary part is larger than $0.01i$. Such values describe strongly light-absorbing particles. We noticed that a systematic bias of the imaginary part may occur if its value is close to 0. This bias is naturally introduced as the imaginary part has a minimum value of 0. This lower limit leads to an underestimation of the SSA.

We analysed plenty of different aerosol types in the course of more than 15 years of measurements with multiwavelength Raman lidar. Still, a statistically significant set of results for each aerosol type has not been achieved because of the labour-intensive manual data analysis. Examples of aerosol types we analysed involve urban/industrial pollution over Europe (Müller et al., 2005), East Asia (Noh et al., 2011), and South Asia (Müller et al., 2001a). We investigated properties of fresh and aged biomass burning aerosols produced in North America (Müller et al., 2005), Europe (Alados-Arboledas et al., 2011), and West Africa (Tesche et al., 2011). In all these cases we assumed in our data analysis that particles have spherical shape, i.e. we used Mie theory in our simulations.

We remain cautious with regard to the inversion of optical data that describe non-spherical (mineral dust) particles, as to date we do not have a reliable light-scattering model that allows us to describe light-scattering at 180°. We carried out some limited studies in which we used the T-matrix algorithm for non-spherical (spheroid) particles as a test on how far we can use this model which has not been designed for describing particle backscatter coefficients. We conclude from our limited set of studies that we may be able to infer particle effective radius of dust particles and the ratio of spherical-to-non-spherical particles (Veselovskii et al., 2010; Müller et al., 2013), though doubt regarding the trustworthiness of the results remains. We fail in retrieving sensible results with regard to the complex refractive index.

1.2 Work at UP

The Potsdam algorithm that was developed at UP is based on the concept of using truncated singular value decomposition (TSVD) as regularization method, (e.g. Hansen (2010)). This method was adapted to work for the retrieval of the particle size distribution function (PSD) and is called hybrid regularization technique (Böckmann, 2001) since it is using a triple of regularization parameters. Simulation studies demonstrated that the hybrid method can be used to invert monomodal and bimodal PSDs (Böckmann, 2001; Böckmann et al., 2005). The minimum requirement for meaningful inversion results are optical input data obtained from measurements of backscatter coefficients at 355, 532, and 1064 nm and extinction coefficients at 355 and 532 nm. How-

ever, the software can also be used with a smaller or larger number of optical coefficients. The inversion of an ill-posed problem, such as the retrieval of the PSD, is always a challenging task because small measurement errors and tiny rounding errors often will be hugely amplified during the solution process unless an appropriate regularization method is used. The use of a regularization method requires the careful determination of the associated regularization parameters (e.g. Doicu et al. (2010); Hansen (2010); Rodgers (2000); Vogel (2002)). Therefore, in one of the next stages of our development work we decided to use two regularization techniques in parallel for comparison purposes. The second method we used for comparison is an iterative regularization method based on Páde iteration (Kirsche and Böckmann, 2006; Böckmann and Pornsawad, 2008). Here, the number of iteration steps serves as the regularization parameter. This method was adapted and tested successfully to invert the PSD (Böckmann and Kirsche, 2006). The approximated PSD is a linear combination of B splines with appropriately determined weights. The B splines of order d are polynomials of degree $d-1$. Since the distribution of the B-spline nodes also plays a critical role, we modified the first algorithm TSVD such that non-equidistant node grids can be used too (Böckmann, 2001). The iterative method was equipped with an adaptive non-equidistant node grid by Osterloh et al. (2011).

There is another challenge that needs to be considered. The CRI is actually also unknown. In order to solve this problem, a grid of viable options for the CRI (all combinations of real parts of CRI (RPCRI) and imaginary parts of CRI (IPCRI)) is assumed. If this is not done, one additionally has to deal with a non-linear problem. The CRI grid technique is very time consuming, i.e. the computer runtime of the inversion process is very large. Therefore, we developed a semi-automated software for spherical particles which is able to run on a parallel processor machine (Osterloh et al., 2009) (see Fig. 1). The software is adapted with a tool that reads NetCDF files from the EARLINET database (The EARLINET publishing group, 2014).

From a mathematical point of view, it is also very important (as a selection criteria for an appropriate regularization method) to investigate the degree of ill-posedness of the problem. Investigations show a moderate ill-posedness (Böckmann et al., 2005; Osterloh et al., 2013) which is not the worst case. Furthermore, first investigations with regard to non-spherical particle shape were made. Simulation studies show that the particle shape is an important factor (Böckmann and Wauer, 2001a, b). We carried out retrievals with two-dimensional B splines in which we use a two-dimensional spheroidal model. This model allows us to consider two-dimensional PSDs in the sense that such PSDs depend on particle radius and aspect ratio. Results are presented in Böckmann and Osterloh (2014). Numerical simulations show that the Páde type regularization method is able to retrieve two-dimensional PSDs from the use of backscatter
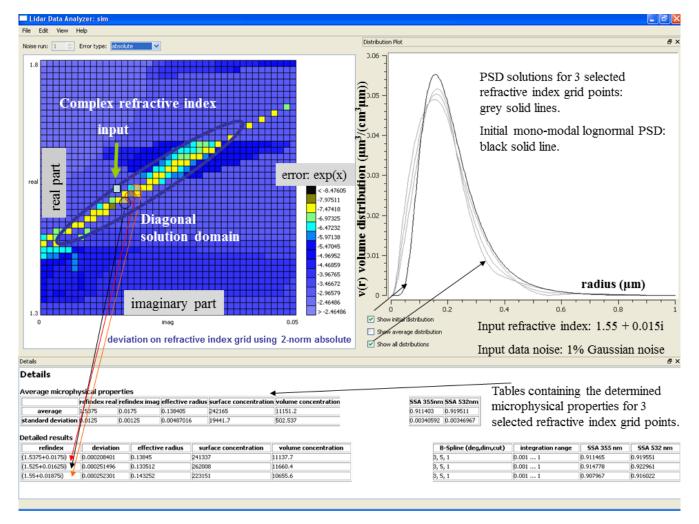
Figure 1. Screen shot of the developed software and explanation of the post-processing procedure: selection of suitable grid points of CRI clustering along the diagonal domain (left), corresponding PSD and initial PDS (right), and tables with retrieved microphysical parameters for each selected CRI and mean values with standard deviation (bottom). The residual errors on the right hand side of the grid appear in ascending order from top to bottom on a logarithmic scale.

and extinction coefficient profiles, as well as depolarization ratio information.

The semi-automatic software was applied to measurement cases. We analysed Raman lidar measurement data with three backscatter coefficients at 355, 532, and 1064 nm and two extinction coefficients at 355 and 532 nm. From these optical particle variables we retrieved microphysical particle properties of different aerosol types. Successful retrievals have been made for biomass burning and industrial pollution aerosols over Germany (Wandinger et al., 2002), continental aerosols (Eixmann et al., 2002), Arctic haze aerosols (Hoffmann et al., 2012), biomass burning aerosols over Romania (Osterloh et al., 2013) and Athens (Papayannis et al., 2008). Moreover, the two-dimensional spheroidal model was applied to a measurement scenario of Saharan dust observed over Barbados (13.16° N, 59.44° W). We obtained results for the two-dimensional fine-mode PSD. We found promising

results, in particular, when we used additional depolarization ratio profiles (Böckmann et al., 2012). Finally, Samaras et al. (2015) show a direct quantitative comparison of the retrieved microphysical properties to measurements from a Compact Time of Flight Aerosol Mass Spectrometer (AMS) which allowed us to validate the results from our algorithm for the case of spherical particles. With regard to the fine-mode fraction of the investigated PSD, we observed remarkable similarities between the retrieved PSD and the one measured by the AMS. Additionally, microphysical retrievals performed with Sun photometer data were also used to explore the results of biomass burning aerosols.

We did not perform a direct comparison of the performance of the two algorithms (TROPOS/UH and UP) in this contribution as we are mainly interested in learning how we can meet demands for highly accurate single-scattering albedos. This demand will likely require a coordinated use

of both algorithms in future rather than selecting one algorithm over the other algorithm in EARLINET. We carried out simulations with the main purpose of investigating what maximum accuracy and precision of microphysical parameters can be achieved (UP algorithm). The complex refractive index is the key to retrieving accurate values of single-scattering albedo and profiles of absorption coefficients. We therefore investigated if an accurate knowledge of the real part of the refractive index would significantly improve the retrieval of the imaginary part which then would feed into improved values of single-scattering albedo. We used the UH/TROPOS algorithm for this part of the study.

Section 2 presents main features of the two methodologies. Section 3 presents some simulation studies that illustrate the current status. Section 4 closes with a summary and outlook.

2 Methodology

The microphysical properties are derived from solving Fredholm integral equations of the first kind (Müller et al., 1999a; Veselovskii et al., 2002; Böckmann, 2001; Böckmann et al., 2005; Ansmann and Müller, 2005). These equations can be written in general terms as

$$g_l^*(\lambda_i) = \int_0^\infty K_l(m, r, \lambda_i, s) f^*(r) \, dr. \tag{1}$$

The term $g_l^*(\lambda_i)$ denotes the exact optical data at the measurement wavelength λ_i. The optical data usually have an uncertainty as we use experimental data. The subscript l denotes the type of optical data, i.e. β denotes particle backscatter coefficients and α denotes particle extinction coefficients.

The backscatter and extinction kernel functions are denoted by $K_\beta(m, r, \lambda, s)$ and $K_\alpha(m, r, \lambda, s)$. The kernel functions depend on the complex refractive index m of the particles, i.e. $m = m_R - m_I i$. The term m_R denotes the real part. The term $m_I i$ denotes the imaginary part of the CRI. The radius of the particles is denoted as r. The shape properties s of the particles determine their backscatter and extinction properties. We only consider spherical shape of the particles and therefore drop this parameter in the following discussion.

The kernel functions $K_l(m, r, \lambda_i)$ are calculated from Mie-scattering theory (Bohren and Huffman, 1983). The term $f^*(r)$ describes the exact (true) particle size distribution which is described as the number of particles per unit volume between particle radius r and $r + dr$.

Equation (1) can be written as

$$g_p^* = \int_{r_{\min}}^{r_{\max}} K_p(m, r) f^*(r) \, dr + \epsilon_p^{\text{limits}}. \tag{2}$$

The term g_p^* describes the exact values of the optical data. The subscript $p = (l, \lambda_i)$ describes the kind of optical input

data at a given wavelength. The lower and upper integration limits of the particle radii within which the inversion is performed are denoted by $r_{\min}$ and $r_{\max}$. We write this radius interval as $[r_{\min}, r_{\max}]$. The term $\epsilon_p^{\text{limits}}$ represents the model error that results from the fact that the integration limits are not from 0 to ∞.

Disregarding the model error term we obtain

$$g_p^* = \int_{r_{\min}}^{r_{\max}} K_p(m, r) f_M(r) \, dr. \tag{3}$$

The investigated PSD $f_M(r)$ is obtained from the numerical solution of Eq. (3) (Twomey, 1965; Zuev and Naats, 1983; Engl et al., 1996; Hansen, 2010). The numerical solution is not an easy task since this compact Fredholm operator is ill posed. Firstly, the expression $f_M(r)$ is approximated by a sum that consists of the linear superposition of base functions:

$$f_M(r) = f(r) + \epsilon(r) = \sum_{j=1}^{N_B} f_j B_j(r) + \epsilon(r). \tag{4}$$

The term $f(r)$ is an approximation of the solution of Eq. (3). The expression f_j describes weight coefficients. $B_j(r)$ are B-spline base functions. N_B is the number of base functions that are used in the inversion. The term $\epsilon(r)$ is the discretization error. This error is caused by discretizing the PSD with the linear combination of base functions.

Neglecting the discretization error term we can express the measured optical particle properties g_p by an approximation of a linear combination of base functions according to Eqs. (3) and (4):

$$g_p = \sum_{j=1}^{N_B} A_{pj}(m) f_j. \tag{5}$$

The matrix elements A_{pj} are calculated from the kernel functions and the base functions as

$$A_{pj} = \int_{r_{\min}}^{r_{\max}} K_p(m, r) B_j(r) \, dr. \tag{6}$$

If we write the optical data as vector $\boldsymbol{g} = [g_p]$ and the weight coefficients as vector $\boldsymbol{f} = [f_j]$ we can write Eq. (5) as a vector–matrix expression:

$$\boldsymbol{g} = \mathbf{A} \boldsymbol{f}. \tag{7}$$

The matrix $\mathbf{A} = [A_{pj}]$ is rectangular. It is called weight matrix. The elements of this matrix are calculated from Eq. (6).

Since the operator is ill posed, as mentioned above, the resulting linear equation system Eq. (7) is ill conditioned. Therefore, the system cannot be solved by standard routines. We need specific regularization techniques to solve such systems.

The unknown vector of the weight coefficients f is connected by the matrix operator $\mathbf{A}$ with the optical data vector g. We note that the data vector g contains measurement errors. For ill-conditioned systems this means that small data errors may be amplified strongly during a standard solution process. Finally, the total error contains the measurement error and the mathematical approximation errors in an additive way.

The TROPOS/UH algorithm and the UP algorithm use different regularization methods. The TROPOS/UH algorithm uses Tikhonov regularization (Phillips, 1962; Tikhonov and Arsenin, 1977). The UP algorithm uses regularization on the basis of truncated singular value decomposition (TSVD) (Engl et al., 1996; Hansen, 2010).

With respect to the subsequently realized simulation study in which we apply simulated noise to the optical data, the following issue is considered. On the one hand, since noise or measurement errors are randomly distributed to the optical data one needs a reasonable sample size for the simulation study for the case of noisy data. On the other hand, the simulation studies are very time consuming. Thus, a compromise between reasonable sample size and runtime is needed. Therefore, we decided on 8–10 runs as sample size. Details of both algorithms are described in the next subsections.

2.1 TROPOS/UH algorithm

2.1.1 Solving the Fredholm equations

Detailed descriptions of solving the modified version of Tikhonov's inversion algorithm is explained in detail by Müller et al. (1999a) and Veselovskii et al. (2002).

The number of base functions, N_B, is equal to or exceeds the number of optical data. Veselovskii et al. (2002, 2004) prefer to keep the number of base functions nearly equal to the number of input optical data.

In the case of the algorithm developed at TROPOS, the base functions have triangular shape on a logarithmic radius scale (Müller et al., 1999a, b). We tested two other shapes of base functions, i.e. histogram columns (Heintzenberg et al., 1981) and monomodal logarithmic-normal distributions (Amato et al., 1995). We did not find significant improvement of our data products when we used these shapes of base functions. We believe that one main cause for this result is the fact that the main error sources in data inversion are incorrect optical input data, unknown complex refractive index and uncertainties caused by the regularization procedure. All these errors outweigh the potential improvements that could be obtained by using a more suitable description of the investigated PSDs, e.g. base functions of logarithmic-normal shape. The base functions of second order, i.e. first degree that we use, have also shown to work sufficiently well for the reconstruction of PSDs of bimodal shape (Veselovskii et al., 2004).

We can solve Eq. (7) by introducing the transposed matrix $\mathbf{A}^T$. Assuming its existence we use $\left(\mathbf{A}^T\mathbf{A}\right)^{-1}$ as the inverse of matrix $\mathbf{A}^T\mathbf{A}$. In that way we obtain the simple normal solution of Eq. (7) for the weight factors:

$$f = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T g. \tag{8}$$

The solution of Eq. (8) usually leads to physically useless results because the mathematical problem is ill posed. The reason for it is that the matrix-operator $\mathbf{A}$ is ill conditioned since it is the discretized representative of an infinite-dimensional ill-posed operator. Details on this property can be found in, e.g. Twomey (1977). Explanations can also be found in, e.g. Müller et al. (1999a), who provide details on how to overcome this problem. Briefly, we introduce the equation (Twomey, 1977; Tikhonov and Arsenin, 1977)

$$\left(\mathbf{A}^T\mathbf{A} + \gamma\mathbf{H}\right)f = \mathbf{A}^T g. \tag{9}$$

The expression $\mathbf{H}$ is the smoothing matrix operator (Twomey, 1977). This operator describes the physical constraint that the retrieved PSD does not show "large" oscillations in the size range for which the PSD is retrieved. This size range is determined by the inversion window.

Details on the appropriate choice of $\mathbf{H}$ can be found in Twomey (1977). $\mathbf{H}$ influences the maximum difference between the weight factors of successive base functions. γ is the Lagrange multiplier. It is the non-negative regularization parameter that determines the degree of smoothing of the investigated PSDs, i.e. the strength of the operator $\mathbf{H}$.

2.1.2 Identification of the solution space

Finding the solution requires the application of several constraints. These constraints stabilize the inversion problem and help us reject mathematical solutions that are physically not reasonable. We use the simplifying assumption of a wavelength- and size-independent complex refractive index of the aerosol particles.

The rationale for using a gliding inversion window is given by Müller et al. (1999a, b). We used to apply 50 inversion windows (Müller et al., 1999a) but subsequently moved to 400 windows in this manually operated algorithm. We find that the quality of the retrieval can be improved if we use more inversion windows on a smaller radius search grid. Figure 1 in Müller et al. (1999a) shows how the inversion windows are formed.

We also discretize the CRI search space. In this contribution the real part m_R varied from 1.325 to 1.8 with step 0.025. The imaginary part, $m_I i$, varies from 0 to $0.05i$ with step $0.003i$.

In that way we obtain k individual mathematical solutions for a given optical data set. Each solution number k is characterized by an inversion window $\left[r_{min}^{(k)}, r_{max}^{(k)}\right]$ and a CRI value $m^{(k)} = m_R^{(k)} - m_I^{(k)}i$, which defines the vector $f^{(k)}$ of

the weight coefficients and the vector $g^{(k)}$ of the optical data that can be backcalculated from the inversion results.

$$\mathbf{A}^{(k)} f^{(k)} = g^{(k)} \tag{10}$$

We calculate the discrepancy vector, as introduced by Veselovskii et al. (2002):

$$\boldsymbol{\rho}^{(k)} = \mathbf{A}^{(k)} \left| f^{(k)} \right| - g \tag{11}$$

and the normalized discrepancy value

$$\rho^{(k)} = \frac{1}{N_O} \sum_{j=1}^{N_O} \frac{\rho_j^{(k)}}{g_j} 100\%. \tag{12}$$

The symbol $|\cdot|$ means that every element of the vector $f^{(k)}$ is converted into its absolute value. The number of optical data is N_O. The term $\rho_j^{(k)}$ denotes the jth component of vector $\boldsymbol{\rho}^{(k)}$. The jth component of vector g is g_j.

The simulations were carried out with synthetic data and with uncertainties added to the data. The main purpose of the simulation with erroneous data was to learn by how much the inversion products could deviate from the correct results for various error levels. We tried to answer this question by distorting the optical data such that extreme changes (distortions) of the spectra of the backscatter coefficients (at 355, 532, and 1064 nm) and the spectra of the extinction coefficients (at 355 and 532 nm) could be achieved. We assumed an uncertainty of 5, 10, and 15 % for the data points.

For example in the case of 15 % error, we added 15 % to the extinction coefficient measured at 355 nm and we subtracted 15 % from the extinction coefficient at 532 nm. We did the same for the backscatter coefficients at 355, 532, and 1064 nm. In that case six combinations of +15 % and −15 % error are possible; however two scenarios were not considered as they merely lead to an overall shift of the backscatter spectrum to +15 % or −15 %. In combination with the two possible distortions of the extinction spectrum we obtain eight possible error scenarios. The inversion was carried out for each of the eight distorted optical data set, and in this way we obtained eight different solutions. We then averaged these eight solutions, which provides us with a mean value of the inversion of the erroneous data set and an uncertainty bar in terms of 1 standard deviation. An important point in that approach is that we did not apply any smoothing constraints that would reduce these extreme slopes of the extinction and backscatter spectra. Thus we also create extreme values of the extinction-, backscatter-, and lidar-ratio-related Ångström exponents.

2.1.3 Particle size distributions: examples

Table 1 summarizes the parameters of the PSDs that were used in the simulation studies. Additional explanations are given in Sect. 3.1.

Table 1. Input parameters of the particle size distributions used in the simulation studies. We used monomodal PSDs with mean radius 100 nm. We used PSDs normalized to one particle per cm^3, gsd (σ) denotes the geometric standard deviation (mode width), er is the effective radius, sc is the surface area concentration, and vc the volume concentration.

gsd	er (µm)	sc (µm² µm⁻³)	vc (µm² µm⁻³)
1.5	0.15	0.18	0.0088
1.7	0.2	0.22	0.015
1.9	0.28	0.29	0.027
2.1	0.4	0.38	0.05
2.3	0.57	0.5	0.095
real part: 1.4, 1.5, 1.6			
imaginary part: 0, 0.005, 0.01, 0.03, 0.05			

Figure 2 shows examples of PSDs retrieved with the TRO-POS/UH algorithm. The task of deriving the shape of the particle size distribution is challenging as we are dealing with a small number of optical data, i.e. the information content of the set of optical data is low. However, during our development work we noticed that we can obtain some information on the particle size distribution too.

The panel shows the results for moderately absorbing aerosols, i.e. the imaginary part is $0.01i$. The first row (a–c) shows the results for error-free data and the assumption that the real part can be derived to an uncertainty of 0.1. The second row (d–f) shows the results for error-free data and the ideal case that we can derive the real part to 0.05 accuracy. The rationale for using these uncertainties of 0.1 and 0.05 will be given in Sect. 3.1. The third row (g–i) shows the results for a measurement error of 15 % and an uncertainty of 0.1 of the real part. The fourth row (j–l) shows the results for the measurement error of 15 % and an uncertainty of 0.05 of the real part.

The left plot of each row (a, d, g, j) shows the results for the mean radius of 60 nm. The middle plot in each row (b, e, h, k) shows the results for the mean radius of 140 nm. The right plot of each row (c, f, i, l) shows the results for the mean radius of 300 nm. These three mean radii are equivalent to effective radii of 0.23, 0.55, and 1.2 µm. The numbers describe 1 PSD (in volume concentration presentation) for which particles are mainly in the fine-mode fraction (left column), 1 PSD for which particles are in the fine- and coarse-mode fraction (centre column), and 1 PSD for which the particles mainly are in the coarse-mode fraction (right column).

We see that the shape of the particle size distributions can be derived to some degree. The individual solutions in general exhibit similar features. The panel shows that we may not be able to derive the exact shape of the particle size distributions. The individual PSDs show peaks that are not exactly at the position of the peak of the true PSD. Nevertheless, if
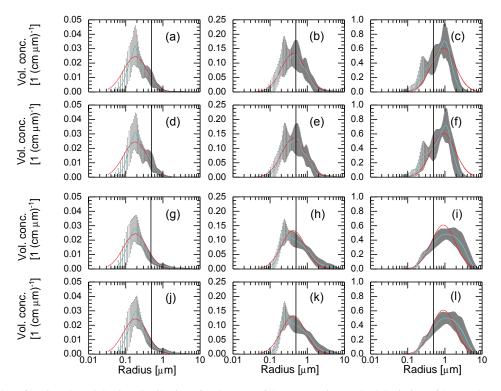
Figure 2. Examples of retrieved particle size distributions for the case of the geometric standard deviation of 2.1, real part of 1.5, imaginary part of $0.01i$ and for mean radius of **(a, d, g, j)** 60 nm, **(b, e, h, k)** 140 nm, and **(c, f, i, l)** 300 nm. The true PSD is shown as red curve. The retrieved mean PSDs are shown as solid light blue lines and error bars (grey) in terms of 1σ standard deviation. The error bars are plotted on a dense grid and follow from the averaging of several hundred individual inversion solutions for each of the mean PSDs. The vertical lines are positioned at 500 nm particle radius. We define 500 nm particle radius as separation between fine-mode and coarse-mode particles. Row **(a–c)** represents the results from using the constraint that the real part is known to 0.1 in the data inversion and that the optical data are error free. Row **(d–f)** shows the results for error-free data and that the real part is known to 0.05. Row **(g–i)** shows the results if the real part is known to 0.1 and that the optical data have an extreme error of 15 %. Row **(j–l)** shows the results if the real part is known to 0.05 and that the optical data have an extreme error of 15 %.

we average all individual PSDs, the mean solution is comparably close to the true PSD.

2.2 UP algorithm

2.2.1 Solving the Fredholm equations and description of the software

The Potsdam algorithm (UP) uses TSVD as a hybrid regularization method and collocation with B splines $B_j(r)$, $j = 1, \ldots, n$, of variable order d. The discretization technique of the Fredholm integral Eq. (3) itself follows the same rules as the TROPOS/UH algorithm (see Eqs. 6, 7). For more details about the TSVD we refer to Engl et al. (1996), Hansen (2010), and Böckmann (2001). In contrast to the TROPOS/UH algorithm, the UP method uses the number n and the order d of the B splines as the first two regularization parameters. The TROPOS/UH algorithm uses the fixed order 2, i.e. B splines with triangular shape. We found that if we use TSVD, the number and shape of the B splines has a large influence on the accuracy of the inverted PSD. Our

experience shows that $n = 3, \ldots, 8$ and $d = 3, 4$ are the most appropriate parameters.

It is a well-known fact (e.g. Hansen (2010)) that the discretization dimension (number n of used B splines) has regularization properties. The matrix $\mathbf{A}$ is decomposed uniquely into $\mathbf{A} = \mathbf{V}^{\mathrm{T}} \mathbf{D} \, \mathbf{U}$ by orthonormal matrices $\mathbf{U}$ and $\mathbf{V}$. $\mathbf{D}$ is a rectangular diagonal matrix. This matrix contains non-negative singular values that cluster to zero. Since small singular values amplify the data noise (measurement errors) and generate oscillations in the solution, namely the PSD, it is necessary to truncate them. The truncation level of the singular values $k = 0, \ldots, \min(5, n) - 1$ is the third regularization parameter of the hybrid method. Thus we obtain a triple (d, n, k) of parameters.

The linear equation system $\mathbf{A} f = (\mathbf{V}^{\mathrm{T}} \mathbf{D} \, \mathbf{U}) f = g$ has to be solved for each parameter triple, e.g. $n = 3, \ldots, 8$, $d = 3, 4$ and $k = 0, \ldots, \min(5, n) - 1$; we note that matrix $\mathbf{A}$ is derived in Eq. (6). The PSD $f(r)$ is determined from a linear combination of a particular set of B splines (number and order) as $f(r) = \sum_{j=1}^{n} f_j B_j(r)$. Details on the non-negative

constraint that is included in addition in the retrieval of the PSD is explained by Böckmann (2001).

The spline number n and order d are parameters that are not independent of each other. These two parameters are related to each other through the number of B-spline nodes. Our algorithm can use an equidistant or a non-equidistant node grid. For the latter option we employ a non-equidistant grid that uses the Chebyshev polynomial roots for these nodes. These nodes are transformed into the interval $[r_{min}, r_{max}]$ to avoid frequently observed characteristics (e.g. oscillations) of the PSD.

Figure 1 shows a screenshot of the developed software. The workflow of the software is split into a set-up, computation, and evaluation step. The set-up step allows the user to specify input data, define the parameter space that will be searched, configure a simulation, and select retrieval methods. The goal of this step is to create a job description that is fed into the computation step.

For the computation step, we have developed a parallel software that allows us either to cope with the vast parameter space or to enable us to carry out a more refined search for a solution. This part of the software is designed such that it runs separately from the interactive set-up and evaluation step. In this way it can be used for parallelized execution on a supercomputer or a computer cluster. A master process splits the work task into small units and delegates the calculations to any available worker process. Once a worker has completed its task, it returns the results to the master. The current search algorithm allows for what we describe as embarrassingly parallel processing (i.e. it does not require any interaction between the workers and therefore scales to a large number of workers (Osterloh et al., 2009).

The screenshot in Fig. 1 was taken from the result-evaluation step after completion of one of the inversion computations. We use the Qt cross-platform application framework. The Qt-based front-end allows the data operator to interactively explore the results and plot further details (right box) for selected coordinates (grid points) (left box).

It is obvious that including the wavelength- and size-independent complex refractive index grid, (Fig. 1; left box), the solution space of the algorithm is quite huge. It contains $|n| \times |d| \times |k| \times |\mathrm{RPCRI}| \times |\mathrm{IPCRI}|$ solutions in total. The term $|\cdot|$ denotes the number of different values of the specific parameter. The solution space is restricted in the following way: for every specific refractive index the best triple (d, n, k) is picked in terms of least residual error. In contrast to the TROPOS/UH algorithm that uses Eq. (7), we find this least residual error from Eq. (3) and forward calculation. In the next step, the best triple at each grid point is used. The associated least residual error is presented on a logarithmical colour scale which displays the error magnitude (Fig. 1). This visual representation is very convenient for the post-processing procedure. At this point we are able to evaluate three different coloured refractive index grids with respect to three different mathematical norms: the Euclidian norm of the absolute and of the relative error, and the maximum norm of the relative error.

As already mentioned in the introduction, a second regularization technique has been included in the software for the purpose of comparison. This regularization technique solves the linear equation system $\mathbf{A} f = g$ in contrast to TSVD by using an iterative regularization method. Kirsche and Böckmann (2006) and Böckmann and Kirsche (2006) developed a whole family of Páde iterations that are used to solve linear equation systems by means of regularization. The well-known Landweber iteration is a member of this family (e.g. Hansen (2010)). We deal again with a triple of parameters (d, n, j), where the third one is now the number of iteration steps j. The number of iteration steps depends on the data noise level ε according to $j = \lfloor \varepsilon^{-1} \rfloor$. The term $\lfloor \varepsilon^{-1} \rfloor$ denotes the integer part of the real number $1/\varepsilon$. Simulations have shown that $j = \lfloor \varepsilon^{-1} \rfloor$ is an appropriate choice. In case of an unknown noise level ε, $j = 30$ is a good choice, e.g. for moderate absorption. To account for the non-negative restriction of the PSD we use a projected iteration (Osterloh et al., 2011).

As noted above there is a strong connection between the distribution of the B-spline nodes and the quality of the reconstructed PSD. To take this connection into consideration, the Páde algorithm adapts the nodes according to certain rules. Indeed, it is easy to see how the PSD is strongly smoothed out in areas in which only a few nodes exist. Strong slopes and curvatures of the PSD require many nodes in their vicinity for an accurate reconstruction. In order to account for this behaviour, the nodes automatically slide towards radius intervals that have larger weight in the PSD during the iteration process, in contrast to fixed non-equidistant Chebyshev nodes. For more details we refer to Osterloh et al. (2011). If we use either TSVD or Páde regularization, we obtain a logarithmically coloured refractive index grid that indicates the error magnitude, as explained above (Figs. 1, 3).

Forward calculations provide us with backscatter and extinction coefficients, which are subsequently used in the data inversion. Even in the noiseless case in which these input data are computed without explicitly considering measurement errors, we still obtain uncertainties because approximation and rounding errors are added to the coefficients, i.e. the input data that are used in the inversion are not truly noiseless. Even those small errors can be harmful for an ill-posed inversion problem.

For the post-processing procedure we manually select the best complex refractive indices depending on the best PSDs. In Fig. 3 we show a few examples for 1.5 of the real part of the CRI, except (c) and (d) with real part 1.4. We use noiseless input data in the retrieval. We select gsd $\sigma = 1.5$ (first and second column in Fig. 3) for the fine-mode particles and gsd $\sigma = 2.1$ (third and fourth column in Fig. 3) for the fine-coarse-mode particles. The rows indicate non-absorbing and absorbing particles (from top to bottom in Fig. 3). The imaginary parts of the CRI are $0i$, $0.005i$, $0.01i$, and $0.05i$. Fig-
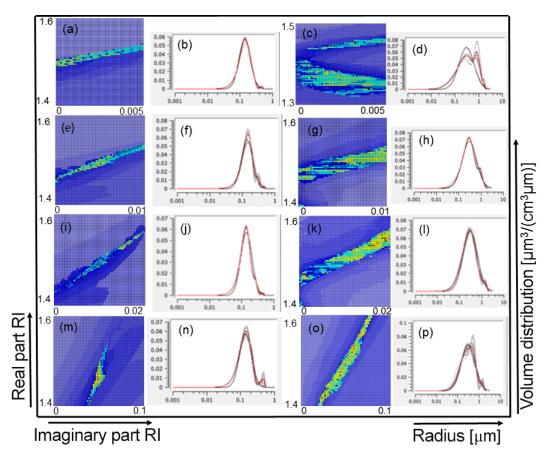
Figure 3. Examples of colour-coded refractive index grids and PSDs for noiseless input data. The rows correspond to different imaginary parts of the CRI: 0, 0.005, 0.01 and 0.05. The first two columns contain fine-mode particles with gsd $\sigma = 1.5$ and real part 1.5. We used equidistant B-spline nodes. The third and fourth columns correspond to intermediate-mode particles with gsd $\sigma = 2.1$, real part of CRI 1.5 and non-equidistant B-spline nodes, except **(c)**, **(d)** with real part 1.4 and equidistant nodes. The mean PSD (the solution) is drawn as red single line.

ure 3 for the most part shows the maximum norm except for (c), (g), and (i) where the Euclidean norm of the absolute error is shown.

The selection procedure is easier for fine-mode particles (Fig. 3, first column) than for fine-coarse- and coarse-mode particles (not shown here) since the best CRIs are located along a very narrow diagonal, in particular with regard to real part 1.5 as shown in Fig. 3a, e, i, and m. We assume that this diagonal structure indicates a lack of information, which keeps us from determining the refractive index uniquely. However, we will give an estimation of the CRI by using the mean of approximately 10 to 20 best values; details are given below.

For the real part 1.4 the diagonal structure more or less disappears for fine-coarse- (Fig. 3c) and coarse-mode particles (not shown here) except for strongly light-absorbing particles of imaginary part 0.05i. We observe a similar behaviour for the same particle types if the real part is 1.6 with non-absorbing or weakly absorbing particles (imaginary part 0i and 0.005i), not shown here. Columns 1 and 3 of Fig. 3 show

that the selection procedure of the CRI becomes easier with increasing imaginary part. We note that we did not use larger coarse-mode particles, e.g. the case of $\sigma = 2.5$, in our simulation study because of the ill-posedness from a mathematical point of view, i.e. the smoothness of the kernel function in Eq. (1). Our investigations show that the Potsdam algorithm should not be used for radii larger than 5–7 µm depending on the refractive index (rule of thumb) (see Osterloh et al. (2013) and Samaras et al. (2015)).

2.2.2 Identification of the solution space for UP algorithm

In this section we explain the selection procedure of CRI and PSD. The main selection criteria are based on our knowledge of working with simulated and experimental data for the case of a grid mesh of 30×30 to 60×60 grid points (see Figs. 1 and 3).

First, if grid points of the CRI are located along a diagonal they can be collected into one cluster as long as this selec-

tion results in a good representation of the PSD (Fig. 3a, e, i, m). Isolated grid points which are not located in a cluster inside the diagonal region or in a cluster inside an arbitrarily shaped region (see Fig. 3c) should be removed even in the case that a good mathematical approximation of the PSD exists. But usually the approximation of the PSD that results from isolated grid points is not very good. Physically meaningful PSDs appear in clusters.

Second, if grid points form a thin vertical or horizontal line and additionally provide a bad approximation of the solution, i.e. of the PSD, all data points along the vertical or horizontal line should be removed. Grid points that appear isolated at the end of a diagonal or an arbitrarily shaped region (we denote them as boundary points) may be removed.

Third, as a rule of thumb one should select at least 10 and at most 20 grid points. If meaningful, one can use accumulative two or three mathematical norm grids of the refractive index (as previously explained). If more than one cluster exists and if these clusters have the same number of solutions it is difficult to decide which cluster should be preferred. Changing the mathematical norm grid may help in the decision making. For more details see Samaras et al. (2015).

Finally, for fine-mode particles ($\sigma = 1.5, 1.7$, i.e. narrow mono-lognormal PSD) we suggest using an equidistant B-spline node grid. However, our experience shows that better results can be achieved for coarse-mode particles ($\sigma = 2.3$, expanded mono-lognormal PSD) with a non-equidistant grid of left-hand Chebyshev nodes.

Furthermore, the use of the Páde iteration as regularization method often leads to very good results with respect to the pattern of the obtained refractive index grid. Therefore, the manually controlled selection process of the CRI, as described above, often can be performed more easily with Páde iteration than with TSVD regularization. This is especially true for the PSD examples $\sigma = 1.5, 1.7, 1.9$ (fine- and intermediate-mode particles), independent of the real and imaginary part of the CRI. In all cases, the selection process of the CRI grid points that are associated with the best PSD solutions is very easy to do since the best solutions in most CRI grids form distinctive diagonals (not shown here).

In general grid points are not isolated, i.e. all points are located in clusters. If a diagonal structure is absent, we know from experience that meaningful solutions can be identified in arbitrarily shaped clusters too (Samaras et al., 2015). Only broader modes, $\sigma = 2.1, 2.3$, seem to be occasionally problematic in terms of seeking a solution cluster, which could lead to oscillatory PSDs. Therefore, the use of the Páde iteration is a huge improvement. Nevertheless, it should be mentioned that the retrieved mean PSD is less accurate in the case of small particle radii compared to the PSD obtained with TSVD. Therefore, the Páde iteration was not used in this simulation study. It is an ongoing work to combine the methods in an appropriate way to make use of the different advantages associated with both methods.

2.2.3 Particle size distribution: examples

The UP algorithm was tested with the same examples (Table 1) as the TROPOS/UH algorithm. The selection of the best CRI grid points, i.e. the mean CRI, is always strongly connected with the corresponding PSD, i.e. we look for similar shapes of the PSD. The retrieved mean PSD solution, Figs. 3, 4 (solid red line), is the average of 10 to 20 PSDs (solid grey lines) corresponding to the selected 10 to 20 CRI grid points as described in the last section (true PSD: solid black line).

Figure 3 shows eight examples of retrieved PSDs for the case of noiseless data. In that case it is possible to retrieve the monomodal lognormal shape of the input PSD very well with regard to accuracy and precision in almost all examples, in particular peak height and location match accurately (Fig. 3). We note (not shown here) that a second peak very often occurs for coarse-mode particles (gsd $\sigma = 2.3$), although this second peak is very small in most cases. The exceptions are weakly absorbing (0.005i, 0.01i) particles with real parts around 1.4. In that case the PSD is monomodal. For non-absorbing particles or strongly light-absorbing particles (0.05i) with real part 1.4, the second peak also occurs for gsd $\sigma = 2.1$ (Fig. 3d) and gsd $\sigma = 1.9$ (not shown here).

In summary, the retrieval of the PSD in the case of noiseless data is excellent with regard to retrieval accuracy and retrieval precision. Only for almost all real and imaginary parts in the coarse-mode case and for fine-coarse-mode cases with real parts around 1.4 and imaginary parts of 0i and 0.05i respectively, the retrieved PSDs show a small second peak. But this second peak probably is only a mathematical feature (oscillation) which cannot be smoothed out during the regularization process. The reason for this second peak is probably the larger ill-posedness of the underlying mathematical problem.

Figure 4 shows 16 examples of retrieved PSDs for the case of noisy data (random Gaussian noise of 15 %, 10 runs per example). Most of the results are good, in particular for gsd 1.7 and 1.9 where the accuracy of the results is very good (Fig. 4, second and third column). In the case of noisy input data, oscillations of the PSDs that describe coarse-mode particles (gsd $\sigma = 2.3$) occur. These PSDs show a second or even a third peak for all CRIs considered in this study. However, the retrieval accuracy is good (except for strongly light-absorbing particles with 0.05i (Fig. 4p). In contrast, retrieval precision is not that good. Additionally, all PSDs coupled with real part 1.4 and gsd 1.7–2.1 show more or less a second peak for all imaginary parts (not shown here).

The retrieved PSDs of fine-mode particles with gsd $\sigma = 1.5$ are always monomodal. However, the peak height is underestimated and the peak is shifted to larger radii (see Fig. 4a, e, i, m); the precision is still good but the accuracy is lower in comparison with gsd 1.7 and 1.9 (Fig. 4, second and third column). The impact (quantity) of both effects (location and height of the peak) decreases with increasing real part,
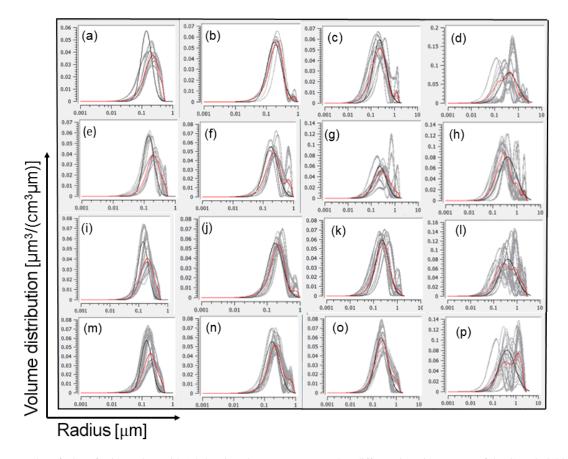
Figure 4. Examples of PSDs for input data with 15 % noise: the rows correspond to different imaginary parts of the CRI: 0, 0.005, 0.01, and 0.05. The columns correspond to gsd: 1.5, 1.7, 1.9, 2.3. The real part of CRI is 1.5. The mean PSD, i.e. the solution, is shown as a solid red line.

independent of the value of the imaginary part (not shown here).

For the combination of real parts 1.5 and 1.6 and $\sigma = 1.7, 1.9$ the retrieved PSDs (red solid lines) are monomodal. The peak height and its location show excellent overlap to the peak height and location of the initial distribution (black solid lines) for all imaginary parts. Examples are shown in Fig. 4, second and third column, for real part 1.5. The accuracy and precision is very high even in the case of noisy data. In contrast, for $\sigma = 2.1$ (not shown here) the peak height is often a little bit overestimated, but the peak location is only slightly shifted to the left or right.

In summary, if 15 % noise is added to the input data, the PSD can still be retrieved quite well for fine-coarse-mode particles in the case of real parts 1.5 and 1.6 in combination with all imaginary parts.

3 Simulation results and discussion

3.1 Generation of optical data for retrieving microphysical parameters

Table 1 shows the parameters of the particle size distributions (effective radius and geometric standard deviation) and the CRIs that were used for the computations of the optical input data. We used five different effective radii.

In our study we consider particles of radius below 500 nm as fine-mode particles and particles above 500 nm as coarse-mode particles. We did not investigate explicitly bimodal particle size distributions. All tests were done with monomodal particle size distributions. Nevertheless we think that the choice of our particle size distributions still allows us to infer conclusions on the performance of our algorithms with respect to what is usually denoted fine-mode fraction and coarse-mode fraction of bimodal particle size distributions, even if the investigated size distributions are only monomodal. The sensitivity of the algorithm toward particle size depends on the underlying Mie-scattering efficiencies for single particles, and this dependence does not change in dependence of the modality of the particle size distribution.

Effective radii of 0.15 and 0.2 µm are in the size range of the fine-mode fraction of particle size distributions. Effective radii of 0.28 and 0.4 µm describe particle size distributions that have a significant share of particles in the coarse-mode fraction and the fine-mode fraction. An effective radius of 0.57 µm describes a size distribution for which most of the particles are in the coarse-mode fraction of the particle size distribution.

With regard to the complex refractive indices we tested three real parts, i.e. 1.4, 1.5, and 1.6. In our opinion these values cover a realistic range of real parts that can be expected for atmospheric particles. Values of around 1.4 describe highly refractory particles. Sea salt belongs to this class of particles. The value of 1.5 can be used to describe industrial pollutants, for example sulfuric acid. Soot has a high real part of 1.8, but it is usually not found in pure form. Thus, if soot mixes with other aerosol components the real part reduces. We estimate that 1.6 is a representative value for pollution particles, for example biomass burning particles that contain some amount of black carbon.

The imaginary part varies over several orders of magnitude. Our goal in our software development is that we can find the correct value at least to within $\pm 50\%$.

An accuracy of approximately ± 0.03 for single-scattering albedo is often assumed as a basic requirement in order to test the sensitivity of light-absorbing aerosols in climate change studies. This accuracy, however, cannot be expressed in terms of a single number of the accuracy of the imaginary part. The reason for it is that the particle size distribution also influences the value of the SSA. A small change of the imaginary part may have a significant impact on the SSA if the particles are in a specific radius range. A small change of the imaginary part may not have a significant impact on the SSA if the particles are in another part of the radius range of atmospheric particles. We currently cannot quantify by how much the SSA changes if particle size and/or complex refractive index change by a certain amount. Detailed simulation studies (e.g. forward Mie-scattering computations) for a wide range of scenarios of PSDs and CRIs would be required. This work goes beyond the scope of the current study.

In the first step we used the UP algorithm to test if we are able to derive the imaginary part to within ± 0.005 in absolute values. We used the TROPOS/UH algorithm to find out the accuracy of all other retrieved microphysical parameters, assuming that the real part can be derived to ± 0.05 or ± 0.1 accuracy. We investigated if this constraint on the real part allows for retrieving the imaginary part to ± 0.005. This accuracy of ± 0.005 may in fact not be achievable, at least not for arbitrary particle scenarios, but it would significantly increase our chances to retrieve highly accurate values of single-scattering albedo.

We had to restrict our simulations to a few imaginary parts because the inversion algorithms are manually operated and the data analysis is time consuming. We selected a few imaginary parts that were meant to give us a reasonable overview on the retrieval performance if particles are non-absorbing (imaginary part $= 0i$) and if particles are highly light-absorbing (imaginary part $= 0.05i$).

One point that must be considered in these simulations is the fact that, with regard to experimental conditions, we likely will not find all possible combinations of the particle size parameters (effective radius, real, and imaginary part of the complex refractive index) listed in Table 1. The reason why we believe that this will not happen can be seen from the following Table 2.

Table 2 shows the minimum and maximum values of the extinction-related and backscatter-related Ångström exponents that follow from the combinations of the five effective radii and the CRIs listed in Table 1. The other three values (five in total) for each of the Ångström exponents fall within these minimum and maximum values. We also show the individual values of the lidar ratios, and the ratio of the two lidar ratios (at 355 and 532 nm) that we obtain from all the combinations of the real and imaginary parts for each effective radius tested in this sensitivity study.

The extinction-related Ångström exponents largely cover the range of values we found from measurements of extinction coefficients at 355 and 532 nm. We regularly measure extinction-related Ångström exponents of 1–2 in regions that are affected by anthropogenic pollution. Maximum values that have been measured are as high as 2.5, but we did not test this scenario in our study.

Values below 1 describe large particles in the coarse-mode fraction of particle size distributions. The most likely candidate of an aerosol type with extinction-related Ångström exponents below 1 is mineral dust. However, we cannot simulate with reasonable confidence optical data that describe mineral dust particles. Until now we could not identify a light-scattering model that would allow us to model trustworthy values of particle backscatter coefficients, i.e. scattering at 180°. However, in the past we measured extinction-related Ångström exponents of 0.5–1. Such values are related to aged biomass burning aerosols (Müller et al., 2005). In a few instances, extinction-related Ångström exponents of aged biomass burning aerosols were less than 0.5.

We also considered the case of extinction-related Ångström exponents around 0. Again, this scenario most likely occurs if large mineral dust particles are present. However, large sea-salt particles may also show extinction-related Ångström exponents of around 0 and for that reason we simulated such cases as well. We point out that it is not unlikely to find extinction-related Ångström exponents slightly less than 0, as shown in Table 2. It is unclear if negative values can be resolved by Raman lidar measurements in view of realistic measurement errors of 10–20%.

The critical point in this table is the lidar ratios. Several values are quite clearly very unusual. Some lidar ratios are considerably higher than 100 sr, and some values are considerably lower than 20 sr.

Table 2. Input parameters of the size distributions used in the simulation studies. We used monomodal PSDs with mean radius 100 nm and 5 different geometric standard deviations. We used PSDs normalized to one particle per cm^3. The parameter ext-A(355/532) denotes the extinction-related Ångström exponent for the wavelength pair 355/532 nm, while bsc-A(355/532) and bsc-A(532/1064) denote the backscatter-related Ångström exponents for the wavelength pairs 355/532 nm and 532/1064 nm respectively. LR(355) and LR(532) denote the lidar ratios at 355 and 532 nm respectively. LR(532) / LR(355) denotes the colour ratio of the lidar ratios at 355 to 532 nm. We use five different imaginary parts for each pair of effective radius and real part. For that reason we obtain five different lidar ratios for each of these pairs. In the case of the Ångström exponents, we show the range from minimum to maximum value instead of the five individual values. That range of values arise from the choice of different input parameters used for the particle size distributions.

er	Optical	Real part		
(μm)	Parameter	1.4	1.5	1.6
0.15	ext-A(355/532)	1.45–1.78	1.29–1.49	1.09–1.23
	bsc-A(355/532)	0.36–1.38	0.53–1.79	1.06–2.37
	bsc-A(532/1064)	0.59–1.11	0.6–1.23	0.63–1.4
	LR(355)	81, 89, 99, 142, 191	55, 63, 11, 107, 153	34, 39, 44, 70, 106
	lr(532)	69, 73, 79, 101, 122	62, 67, 8, 92, 113	54, 59, 64, 84, 105
	LR(532) / LR(355)	0.64–0.85	0.74–1.13	0.99–1.59
0.2	ext-A(355/532)	0.94–1.11	0.73–0.85	0.54–0.59
	bsc-A(355/532)	0.19–1.19	0.61–1.76	1.16–2.12
	bsc-A(532/1064)	0.46–1.23	0.59–1.59	0.89–2.03
	LR(355)	72, 84, 96, 155, 229	34, 40, 47, 85, 138	16, 19, 23, 42, 72
	LR(532)	75, 82, 91, 129, 168	49, 55, 62, 95, 132	30, 34, 39, 62, 93
	LR(532) / LR(355)	0.73–1.03	0.96–1.45	1.29–1.86
0.28	ext-A(355/532)	0.55–0.62	0.37–0.4	0.23–0.24
	bsc-A(355/532)	0.03–1.08	0.48–1.31	0.75–1.47
	bsc-A(532/1064)	0.31–1.19	0.6–1.62	0.99–1.95
	LR(355)	56, 72, 87, 161, 257	23, 28, 34, 69, 124	10, 12, 15, 32, 59
	LR(532)	68, 79, 92, 146, 208	32, 39, 46, 81, 130	16, 19, 23, 43, 73
	LR(532) / LR(355)	0.81–1.21	1.05–1.44	1.24–1.64
0.4	ext-A(355/532)	0.3–0.32	0.16–0.17	0.07–0.08
	bsc-A(355/532)	−0.09–0.99	0.23–0.85	0.32–0.94
	bsc-A(532/1064)	0.17–1.1	0.5–1.31	0.78–1.48
	LR(355)	41, 60, 77, 163, 280	18, 24, 30, 64, 121	7, 10, 13, 29, 57
	LR(532)	54, 70, 85, 155, 239	23, 30, 36, 71, 124	10, 13, 17, 34, 63
	LR(532) / LR(355)	0.86–1.31	1.03–1.31	1.1–1.42
0.57	ext-A(355/532)	0.14	0.04–0.05	−0.01 to −0.02
	bsc-A(355/532)	−0.2–0.84	−0.03–0.51	−0.02–0.56
	bsc-A(532/1064)	0.04–1	0.31–0.94	0.46–1.04
	LR(355)	31, 52, 72, 169, 305	16, 22, 29, 65, 128	6, 9, 12, 30, 60
	LR(532)	41, 60, 78, 160, 265	19, 25, 32, 67, 123	8, 11, 14, 31, 60
	LR(532) / LR(355)	0.87–1.33	0.97–1.21	1–1.27

There remains the question of whether it is justified to simulate scenarios in which lidar ratios considerably exceed 100 sr or drop below 20 sr. We believe that we should consider such extreme outliers in at least a few studies for two reasons. First, we can test the robustness of our algorithms for such extreme cases. The second point is that the underlying microphysical properties do not seem to be completely out of range of numbers we can expect for atmospheric particles. It is simply the combination of specific values of particle size distribution and CRI that creates these outliers of lidar ratios.

3.2 Simulation results of TROPOS/UH algorithm

We carried out the inversions as described in the methodology section. We tested how well we can retrieve the parameters of interest under favourable circumstances. In this study we define favourable circumstances as the situation in which we have approximate knowledge of the real part of the complex refractive index. This latter assumption is based on a recent study by Chemyakin et al. (2014), who show that the arrange and average algorithm may be a method that could allow us to constrain the real part of the refractive index to 0.1 or even 0.05 uncertainty (Table 3 in Chemyakin et al.

Table 3. Percentage of simulation cases that result in given retrieval accuracy.

	Results for case of accuracy of real part retrieval of 0.1				
	Retrieval accuracy	TROPOS/UH		UP	
		correct data	15 % error	correct data	15 % error
Effective radius	20 %	79	31	97	43
Surface area conc.	20 %	95	76	97	99
Volume conc.	20 %	77	48	92	47
Real part	0.1	–	–	99	100
Imag. part	0.005	57	36	70	82
SSA (355 nm)	0.05	81	47	91	88
SSA (532 nm)	0.05	85	49	92	88
	Results for case of accuracy of real part retrieval of 0.05				
	Retrieval accuracy	TROPOS/UH		UP	
		correct data	15 % error	correct data	15 % error
Effective radius	20 %	87	27	–	–
Surface area conc.	20 %	95	72	–	–
Volume conc.	20 %	89	57	–	–
Real part	0.05	–	–	83	64
Imag. part	0.005	68	48	–	–
SSA (355 nm)	0.05	85	52	–	–
SSA (532 nm)	0.05	87	67	–	–

(2014)). We made use of this possibility in this study. We explicitly did not attempt to further optimize our inversion results by selecting a subset of best possible solutions in the sense of manually selecting solutions from the solution space that follows from constraining the real part to either 0.05 or 0.1. We obtain a family of individual solutions for which the real part is within either 0.1 or 0.05 deviation from the true value. We average this family of solutions and thus obtain mean value and uncertainty, which in this study will be expressed in terms of accuracy (systematic error or bias) and precision (statistical error or noise).

We also tested our algorithm under the assumption of comparably unfavourable measurement error scenarios. We distorted each optical data point by its maximum value of either 5, 10, or 15 %. In that regard, errors of 15 % represent the worst case scenario in this study. We did this distortion without considering the possibility that data points may be correlated to each other and thus error bars may also not be independent of each other. We assume that the inversion of such "extremely" distorted backscatter and extinction spectra would result in microphysical parameters that also deviate to a maximum value from the correct values. This assumption of course has the flaw that the inversion is a non-linear problem. That means an extreme distortion of optical input data may not necessarily need to lead to a maximum deviation of the retrieved microphysical parameters from their true values. However, we believe that we will learn more about this

type of error analysis in this first attempt and that we can refine it in future studies.

3.2.1 Results for error cases 0 and 15 %

Figure 5 shows a summary of the retrieval performance of the TROPOS/UH algorithm with regard to effective radius, number, surface area, and volume concentration, and the real and imaginary part. The left panel shows the results of simulations carried out for error-free data. The right panel shows the extreme error estimation of the retrieval results.

According to Chemyakin et al. (2014), the real part of the CRI of the optical data we used in this study can be retrieved to an accuracy of approximately 0.1 with 2σ confidence for the case of noiseless optical data (Table 3 in Chemyakin et al. (2014)). We made use of this result in our simulations as we are mainly interested in finding out about the performance of the algorithm under favourable conditions of the input parameters. For example, we are looking for ways that allow us to constrain the search space of the refractive index grid, as the retrieval of the refractive index poses the greatest challenge in our work.

Under this assumption of known refractive index, the other parameters can be derived accordingly. The results are shown as squares in Fig. 5. We see that effective radius can be derived to high accuracy if the particles are in the fine-mode fraction of the particle size distribution. With regard to particles in the intermediate state (transition from fine mode to coarse mode, i.e. effective radii of 0.28 and 0.4 μm) and the
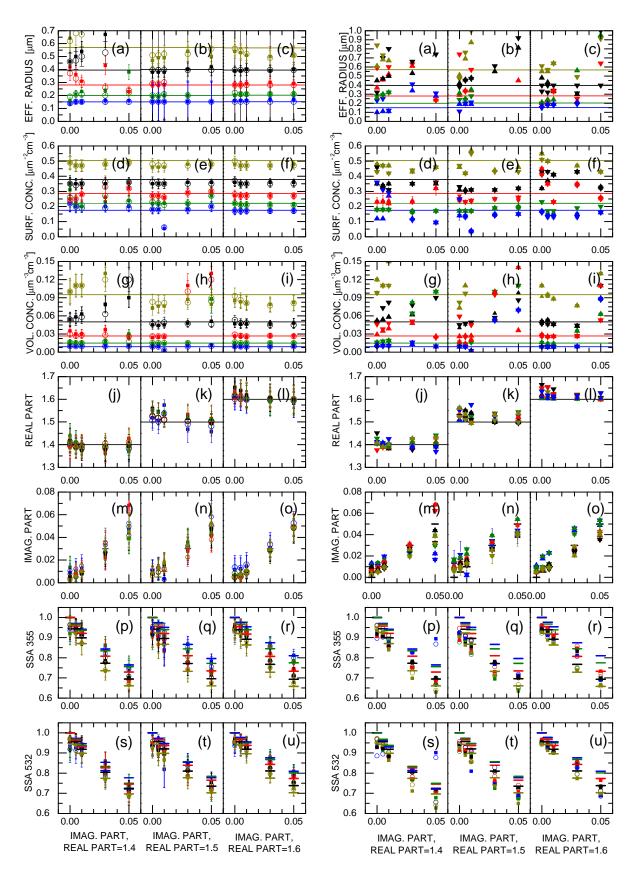
Figure 5.

Figure 5. Left panels: TROPOS/UH algorithm retrieval results for **(a–c)** effective radius, **(d–f)** surface area concentration, **(g–i)** volume concentration, **(j–l)** real part, and **(m–o)** imaginary part of the complex refractive index, and single-scattering albedo at **(p–r)** 355 nm and **(s–u)** 532 nm. The left segment shows the results for real part 1.4, the centre segment shows the results for real part 1.5, and the right segment shows the results for real part 1.6. Our inversion results are obtained with an upper threshold of the real part of 0.1 (boxes) and a threshold of 0.05 accuracy (circles). The error bars denote 1 standard deviation. Right panels: results of our extreme error analysis for the case of 15 % measurement error. Upward-pointing triangles refer to results for which the real part was constrained to 0.05. Downward-pointing triangles refer to inversion results for which the real part was constrained to 0.1. Meaning of the other symbols is the same as in the left panel.

coarse mode, the performance on average still is acceptable (in view of the uncertainty bars) although we find outliers for particles with real parts of 1.4.

Surface area concentration on average shows exceptionally high accuracy. The precision (the uncertainty bars reflect the statistical noise in terms of 1 standard deviation) is high compared to the error bars we obtain for effective radius and volume concentration. We find one outlier (see Fig. 5e) which cannot be explained at the moment.

Volume concentration in general shows the same features as effective radius, i.e. the retrieval accuracy is generally good for particles in the fine-mode fraction. We find some outliers for the intermediate cases and the coarse-mode case, mainly for real parts of 1.4.

The real parts follow from the application of the methodology suggested by Chemyakin et al. (2014) (see Table 3 in Chemyakin et al. (2014)). If the real part is known, we can derive the imaginary part too, as the solutions for the real and imaginary part are correlated. An example of what this correlation looks like is shown for the UP algorithm in Fig. 1. Similar behaviour has also been found for the TROPOS/UH algorithm (see for example Fig. 3 in Müller et al. (2001b)).

The imaginary part can be found within ±50 % uncertainty if imaginary parts are 0.01 or larger. This result has already been found in previous studies, e.g. Müller et al. (1999b). We know that we cannot derive the exact value if the imaginary part is $0i$. The inversion results for each data product are always the mean of several results. Thus, even if we assume that the correct optical data are used in the inversion, i.e. no error bars are considered, we end up with an overestimation of the imaginary part. The reason for this situation is that we carry out the inversion for a grid of complex refractive indices. In general we will find acceptable inversion results even for cases in which the imaginary part is not $0i$ (Ansmann and Müller, 2005; Müller et al., 1999b). If we average these individual values, we naturally obtain a bias toward mean values larger than $0i$.

The noteworthy point and a new result compared over previous studies is that our simulations give us an impression of the likely value of the overestimation of the imaginary part. If the true imaginary part is less than 0.01 we may overestimate the imaginary part by $0.005i$ on average. We hope that we can reduce this overestimation in future. One option could be that we refine our search grid of the imaginary part which is currently set to $0.003i$. That means the next nearest value to $0i$ that can be found with our algorithm is $0.003i$. We assume

that this stepsize is likely to be one reason for the overestimation of the mean imaginary part (in the retrieval), which in turn leads to an underestimation of the single-scattering albedo for low light-absorbing aerosols particles.

Figure 5 also shows results if we assume that we could retrieve the real part to an accuracy of 0.05 (open circles). In that case we again make use of the results published by Chemyakin et al. (2014), who show that an uncertainty of 0.05 is possible with 1-σ confidence if the arrange and average algorithm is applied.

The microphysical properties do not differ significantly from the results if the real part is retrieved to 0.1 accuracy and if we take account of the overall uncertainties of our inversion results. That is, any attempt to further improve the real part retrieval may not necessarily result in significant improvement of other data products and we consider this an important outcome of our study.

With regard to single-scattering albedo, Fig. 5 shows the results for each case separately, i.e. for each value of the true imaginary part we show five different results corresponding to the five different particle size distributions that we tested in the simulations. Thus, we show in total 25 different results for each real part. The true value of single-scattering albedo in each column (corresponding to a specific value of the imaginary part and one of five possible effective radii) is shown by a thick horizontal coloured line. Each of these coloured horizontal lines can be linked to one symbol of the same colour. The symbol with that same colour represents the inversion result (mean value) of the underlying optical data set.

The retrieved single-scattering albedo on average can be derived to within ±0.05 regardless of whether the real part is known to within 0.05 or 0.1. If single-scattering albedo is larger than 0.9, the inversion results become worse and tend to underestimate the true single-scattering albedo. If the true single-scattering albedo is 1, the retrieved value is around 0.95.

With regard to the impact of measurement errors of the optical data on our inversion results, we carried out numerous simulation studies in the past (Müller et al., 1999b; Veselovskii et al., 2002, 2004). We also learnt about the impact of measurement errors from the analysis of experimental data. We know that on average measurement uncertainties should be less than 20 % in order to obtain microphysical parameters with low uncertainty. One point we are struggling with is that uncertainty needs to be expressed in terms of

accuracy and precision. All previous studies were only designed to estimate the uncertainty and we kept our studies fairly simple with respect to error analysis of the inversion results. We did not separately study systematic and statistical errors.

In this study we made the first step toward refining our error analysis. We used 5, 10, and 15 % noise for each optical data point. We wanted to test by how much the microphysical parameters could shift for a given measurement uncertainty. In this study we are mainly interested in the likely maximum shift of the mean value of each data product. We try to provide a first answer to that question by what we call extreme error estimation. We distorted the backscatter and extinction spectra to their extremes and took the average of the eight extreme error runs as the final inversion result.

Another idea behind this approach of using maximum distortion is to find out what could possibly go wrong in data inversion if we do not know the error model that describes the uncertainty bars of the optical input data. For example we need to know how systematic and statistical error are computed from the Raman lidar data. Knowledge of the systematic and the statistical uncertainty would allow us to compute the uncertainty bars of the microphysical parameters in a more refined way. We need to know if a Gauss-like error distribution function sufficiently well describes the error bars of the optical input data, or if a different error distribution function is more appropriate. We emphasize that error distribution functions contain information that can be used in data inversion as additional pieces of constraint that might help us to improve the inversion results.

Our procedure needs to be taken with caution. It is just the first step toward dealing with the concept of error analysis in a more concise way. For example, we applied the noise for each data point individually, i.e. we did not consider that errors of the individual data points of a given $3\beta + 2\alpha$ data set may be correlated to each other. Such an effect would likely have a significant impact on the inversion results. We also did not consider that certain situations of distorted data (i.e. values of Ångström exponents) may be less likely than other situations. We would like to use these results as motivation regarding the need for clearly defined data analysis protocols, because the quality of the optical data significantly impacts the quality of the microphysical parameters, or in other words: if the input optical data do not meet certain quality standards, the inversion algorithms cannot fix that problem.

In the following we only discuss the results for 15 % error (right panel of Fig. 5). 15 % error mainly disrupts the retrieval results if lidar ratios are above 100 sr, regardless of particle size and the real part. We notice that the effect is particularly pronounced for imaginary parts of 0.03 and 0.05. A systematic shift could occur (loss of accuracy), but it would not necessarily lead to a change of the statistical uncertainty (loss of precision).

This effect of the loss of accuracy is obvious for effective radius. The uncertainties could become unacceptably large

for some of the retrieved effective radii if we do not use other types of information to constrain the results or if we do not know the mathematical function that describes the error bars.

Surface area concentration remains within a reasonable range of uncertainty if we use 50 % deviation from the true values as benchmark. The retrieval error remains rather well behaved for the two PSDs that describe fine-mode particles, but also the other types of PSDs, i.e. the transition type and the coarse-mode type still deliver useful results.

Volume concentration shows reasonable results (we again use 50 % retrieval error as benchmark when we talk about useful results) for all imaginary parts except the value $0.05i$. The results worsen for large particles, i.e. for effective radii of 0.57 µm, but we can still derive meaningful results even if coarse-mode particles dominate the investigated particle size distribution. We measure the extinction-related Ångström exponent, and large particles are linked to low Ångström exponents. Thus we can use the values from extinction measurements at two wavelengths as quality flag parameter of the inversion results.

The imaginary part is our main target of future studies as it allows us to derive light absorption coefficients of particles. We find that the uncertainty (variation) of the imaginary parts that we obtain from the inversion of erroneous data to large part does not differ from the results we obtain from the inversion of error-free data. The exception is high imaginary parts ($0.05i$). In that case there is a high probability that we would systematically underestimate the imaginary part.

We find on average that single-scattering albedo is underestimated, regardless of the true value of the single-scattering albedo. We do not see a significant difference of the quality of the results (accuracy and precision) of single-scattering albedo at 355 nm and single-scattering albedo at 532 nm.

3.2.2 Results in terms of correlation plots

Figure 6 presents the inversion results in terms of correlation plots. This representation allows for easier identification of the general performance of the inversion algorithm. The left panel of Fig. 6 shows the results for error-free optical data, the right panel shows the results for optical data errors of 15 %. The left column of each panel shows the results if the real part is known to 0.05. The right panel shows the results if the real part is known to 0.1.

The main results were already discussed in the context of Fig. 5. We find that on average the results slightly worsen if the real part is known to 0.1 instead of 0.05 accuracy. Figure 5i, j clearly shows the overestimation of the retrieved imaginary part if the true value drops below $0.01i$. Single-scattering albedo is accordingly underestimated, and this underestimation becomes quite apparent if the true single-scattering albedo rises above 0.95.

If we introduce optical data errors, the quality of the inversion results worsens. Our approach of creating an extreme distortion of the optical spectra gives us some insight into
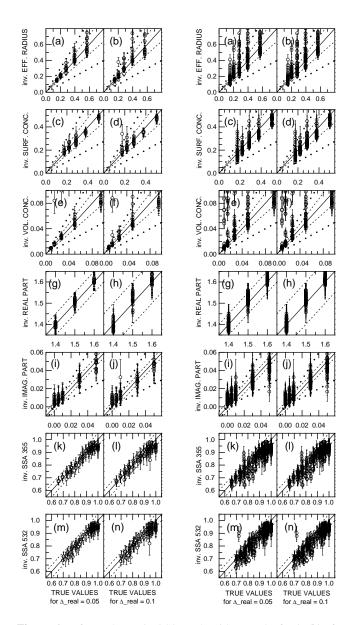
Figure 6. Left panels: TROPOS/UH algorithm results for **(a, b)** effective radius, **(c, d)** surface area concentration, **(e, f)** volume concentration, **(g, h)** real part and **(i, j)** imaginary part of the complex refractive index, and single-scattering albedo at **(k, l)** 355 nm and **(m, n)** 532 nm. Right panels: retrieval results for the case of 15 % data uncertainty. The left (right) column in each of the two panels shows the results if the real part is known to an accuracy of 0.05 (0.1). The error bars denote 1 standard deviation. The 1 − 1 lines are shown. The lines for ±20 % deviation from the true results (dashed lines) and for ±50 % deviation (dotted lines) are shown for effective radius, and surface area and volume concentration. In the case of the real part, the dashed lines indicate ±0.05 deviation (left column in both panels) and ±0.1 deviation (right column in both two panels). In the case of the imaginary part, the dashed (dotted) lines indicate a deviation of ±0.005 (±50 %). In the case of single-scattering albedo the dashed lines show a deviation of ±0.05.

the maximum uncertainties that may occur. We find that effective radius and volume concentration may exceed 50 % uncertainty. Surface area concentration in most cases stays within ±50 % uncertainty. The imaginary part stays within ±50 % uncertainty if the true value is above $0.01i$. If the true value is below $0.01i$ the retrieved imaginary part may show a positive bias of 0.01 to 0.02. Regardless of the increased uncertainty of the microphysical properties single-scattering albedo still can be derived to ±0.05 in many cases. However, as noted before (see Fig. 5) single-scattering albedo is on average underestimated.

3.3 Simulation results of UP algorithm

The simulation results of the UP algorithm are shown in Figures 7, 8, 9, and 10. We split the discussion of our results into two parts: the first part deals with the results obtained with noiseless data. The second part deals with the case of noisy data. Moreover, we note that the complete evaluation is done without any restriction or constraints with respect to the real part of the CRI as it was used previously for the evaluation of the TROPOS/UH algorithm. We note that the UP error analysis was made with Gaussian noise.

3.3.1 The noiseless data case

First, we evaluate the effective radius retrieval and the retrievals of total surface area and volume concentration. In the case of noiseless input data, for almost all fine-mode particle examples ($\sigma = 1.5, 1.7$) the effective radius can be retrieved with a relative error less than 7.5 % except for CRI $1.6 + 0i$, $1.6 + 0.05i$ with $\sigma = 1.5$ (not shown here) and for CRI $1.5 + 0.03i$, $1.5 + 0.05i$, $1.6 + 0.03i$, $1.6 + 0.05i$, i.e. for strong light-absorbing particles, with $\sigma = 1.7$ (Fig. 7c). In those cases the relative error is less than 10 %.

For almost all fine-coarse-mode and coarse-mode particles ($\sigma = 2.1, 2.3$), the effective radius is underestimated (Fig. 7a). The underestimation is between 2 and 13 %. We find outliers for coarse-mode particles for which $\sigma = 2.3$ and CRI $1.4 + 0i$ and $1.6 + 0i$, i.e. non-absorbing particles. The relative retrieval error of the effective radius is only about 25 % in that case. In 90 % of all simulation examples the retrieval errors are below 14 % (Fig. 11h).

The total surface area and volume concentrations are two very stable microphysical parameters. In almost all cases these parameters can be retrieved to better than 15 %. With regard to surface area concentration there are only two outliers, i.e. for $\sigma = 1.5$ with CRI $1.4 + 0.01i$ and $1.4 + 0.03i$. In case of volume concentration we find six outliers. Five of the outliers have a real part of 1.4. In almost all cases, the surface area concentration is overestimated for $\sigma = 1.5, 1.9, 2.1$ (Fig. 7d). The volume concentration shows a very stable behaviour (Fig. 7g). In summary, 95 % of all examples of total surface area concentration and 90 % of all examples of to-
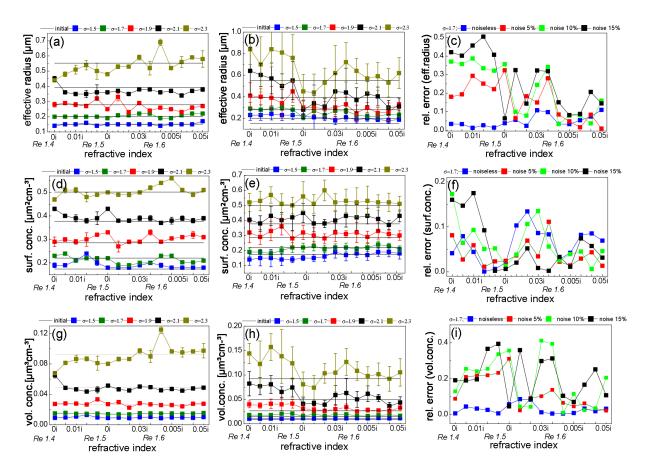
Figure 7. Retrieval results for effective radius (first row), surface area concentration (second row) and volume concentration (third row). The first column corresponds to noiseless data. The second column corresponds to data with 15 % noise. The last column shows the relative retrieval error for increasing noise level of the input data: 0, 5, 10, 15 % for gsd 1.7 as an example. Each subfigure **(a)**–**(i)** is compartmentalized into three partition with respect to the horizontal axis (refractive index). The partitions correspond to different real parts of CRI, namely 1.4, 1.5, and 1.6. Each partition shows the result for different imaginary parts of the CRI: 0, 0.005, 0.01, 0.03, 0.05i. The error bars denote the standard deviation from the mean value of all selected best solutions.

tal volume concentration result in retrieval errors well below 15 % (Fig. 11i, j).

Second, we evaluate the retrieval of the CRI. The retrieval of the real part is very stable. In particular for 1.5 (Fig. 9a, c) the relative error is below 3 % except for two outliers (not shown here). In summary, the relative error is below 5 % for all real parts and all simulation examples (two outliers) and the absolute deviation (90 % of all examples) is below 0.06 (see Figs. 8b and 11k respectively).

The imaginary part has a larger variation if the particles are more absorbing (Fig. 9d). This result holds true for all real parts and is more pronounced for the real part 1.4. The error bars are larger (standard deviation of all selected CRI grid points from the mean). Moreover, for the real part 1.4 (except 1 outlier) the imaginary part is always overestimated for strong light-absorbing particles.

The relative error is less than 55 % except for eight outliers which for most part have an imaginary part of 0.01i. At the moment the reason for this behaviour is unknown.

In summary, with respect to the imaginary part, 83 % of all examples show retrieval errors below 50 %, and 94 % of all investigated imaginary parts show retrieval errors below 70 % (Fig. 8b). We note that we do not include non-absorbing particles with 0i for our evaluation of the relative error since division by zero fails. Figure 9d shows that both the absolute deviation from zero and the error bars are very small. With regard to the absolute error we find that 70 % of all examples are below 0.005 absolute deviation and 85 % are below 0.01 absolute deviation (Fig. 11l).

3.3.2   The noisy data case

In the case of noisy input data (15 %), the relative error of the effective radius is monotonically decreasing with increasing real part of the CRI, i.e. for real part 1.4 the relative error is less than 60 %, for 1.5 it is less than 40 %, and for 1.6 it is less than 30 %. For real part 1.4, the effective radius is overestimated in all examples of investigated PSDs. For the real parts 1.5 and 1.6, this overestimation still occurs in the case
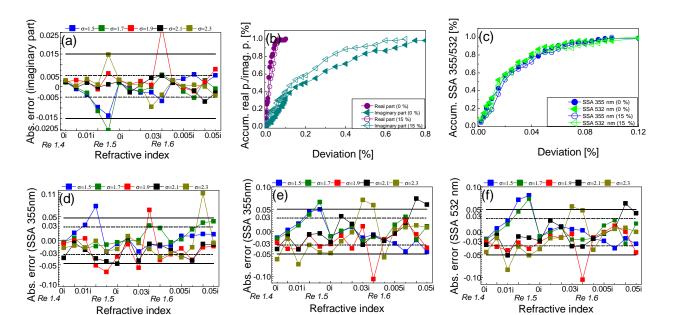
Figure 8. Accumulative distribution of inversion results for UP algorithm **(b, c)**. In the case of **(b)** complex refractive index and **(c)** single scattering albedo we show the accumulative sum (1.0 means 100 %) in dependence of the deviation in percent from the true result. Cases **(a)** and **(d)**–**(f)** show absolute errors of the retrieval results: **(a)** imaginary part of CRI for 15 % noisy input data, **(d)** SSA 355 nm (noiseless), **(e)** SSA 355 nm (15 % noisy input data) and **(f)** SSA 532 nm (15 % noisy input data). Each subfigure **(a)**, **(d)**–**(f)** is compartmentalized into three partition with respect to the horizontal axis (refractive index). The partitions correspond to different real parts of CRI namely 1.4, 1.5, and 1.6. Each partition shows the result for different imaginary parts of the CRI: 0, 0.005, 0.01, 0.03, 0.05i.
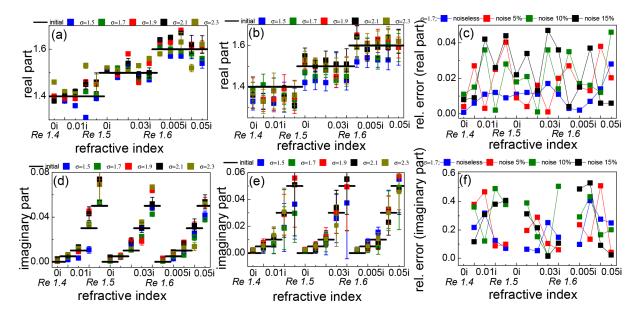

Figure 9. Retrieval results for real part of CRI (first row) and imaginary part of CRI (second row). First column corresponds to noiseless and second to 15 % noisy input data. The last column shows the relative retrieval error with increasing noise level of the input data: 0, 5, 10, 15 % for gsd 1.7 as an example. Each subfigure **(a)**–**(f)** is compartmentalized into three partition with respect to the horizontal axis (refractive index). The partitions correspond to different real parts of CRI namely 1.4, 1.5, and 1.6. Each partition shows the result for different imaginary parts of the CRI: 0, 0.005, 0.01, 0.03, 0.05i. The error bars denote the standard deviation from the mean value of all selected best solutions.
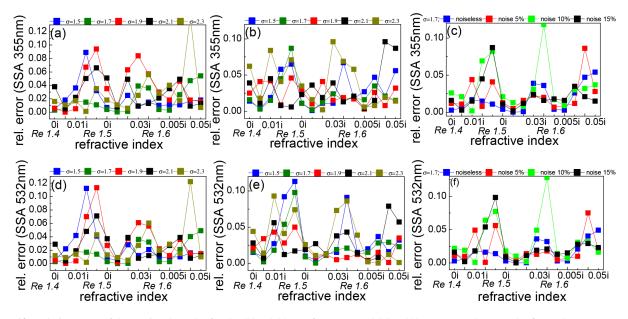
Figure 10. Relative error of the retrieval results for the SSA 355 nm (first row) and SSA 532 nm (second row). The first column corresponds to noiseless and second to 15 % noisy input data. The last column shows the relative retrieval error with increasing noise level of the input data: 0, 5, 10, 15 % for gsd 1.7 as an example. Each subfigure **(a)**–**(f)** is compartmentalized into three partitions with respect to the horizontal axis (refractive index). The partitions correspond to different real parts of CRI, namely 1.4, 1.5, and 1.6. Each partition shows the result for different imaginary parts of the CRI: 0, 0.005, 0.01, 0.03, 0.05i.

of the fine-mode particles (Fig. 7b). It is worth mentioning that the relative errors for fine-mode particles are for the most part larger than for the intermediate and coarse-mode particles. In summary, 85 % of all investigated examples show retrieval errors below 50 % (Fig. 11h). However, if we include only the real parts 1.5 and 1.6, 98 % of all examples show retrieval errors below 40 %.

The surface area concentration is the most stable parameter in the case of inversion of noisy data. The retrieval error is less than 20 % (1 outlier) with respect to the real part 1.4. In the case of real parts of 1.5 and 1.6, the retrieval error is less than 15 % (two outliers). The parameter is slightly overestimated for fine-coarse- and coarse-mode particles (Fig. 7e) but stays within the error bars. The retrieval for fine-mode particles is excellent for the real parts 1.5 and 1.6 and all imaginary parts. In summary, 99 % of all investigated examples show retrieval errors below 20, and 85 % of all investigated examples show retrieval errors below 15 % (Fig. 11i). We note that the retrieval error for most of the examples does not exceed the level of the input data error.

The volume concentration is also slightly overestimated, but in contrast to surface area concentration this overestimation is for the most part only happening for fine-mode particles (Fig. 7h). The overestimation stays mostly within the error bars for the real parts 1.5 and 1.6. Moreover, volume concentration is overestimated for all particles with real part 1.4. The retrieval error is less than 35 % for real parts 1.5 and 1.6. This means that on average it stays within the input data error level. For the real part 1.4, the volume concentration has a

retrieval error less than 55 % except for $\sigma = 2.1$ and $\sigma = 2.3$ where it is 70 %. For fine-mode particles the volume concentration has a retrieval error less than 40 % for all CRIs. In summary, 90 % of all investigated examples of volume concentration have a retrieval error below 50 % (Fig. 11j). If we consider only the real parts 1.5 and 1.6, the retrieval error is below 40 % in 100 % of these examples.

In contrast to the noiseless data case, the real part of CRI 1.4 is always underestimated (1 outlier) in the case of noisy data (Fig. 9b). Moreover, the CRI (all real parts) is nearly always underestimated for the case of fine-mode particles. The relative error is larger than the relative error of the CRI of intermediate and coarse-mode particles. Still, the retrieval error is very small, i.e. 89 % of all examples have an error below 5, and 100 % of all examples have a retrieval error below 6 %. With respect to the absolute deviation, 83 % (noiseless case) and 64 % (noisy case) are below ±0.05, and nearly 100 % (noiseless and noisy case) are well below ±0.1 (two outliers) (Fig. 11k).

With regard to the imaginary part of the CRI, we find that the error bars are very large for strong light-absorbing particles (precision is low), i.e. imaginary parts are 0.03i and 0.05i (Fig. 9e). The accuracy and precision is good for non- and weak light-absorbing particles. In summary, in 90 % of all investigated examples the retrieval error is below 40 %, in 93 % of all cases it is below 50 %, and in 100 % (only two outliers) of all examples the retrieval error is below 60 % (Fig. 8b). We note again that non-absorbing particles with 0i are not included in the evaluation of the relative error. How-
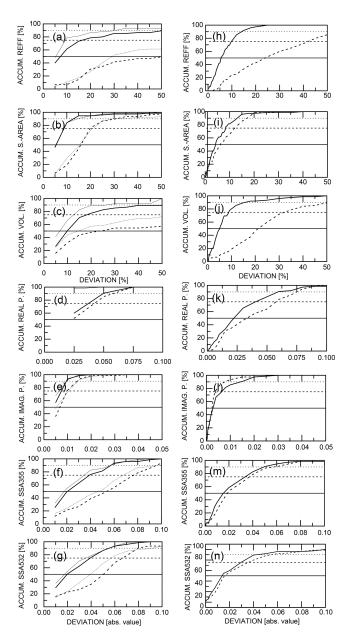
Figure 11. Accumulative distribution of inversion results for TROPOS/UH algorithm (left panel) and UP algorithm (right panel). TROPOS/UH algorithm: results are shown for the constraint that the real part can be extracted to an accuracy of 0.05 (solid thin line) and 0.1 (solid thick line), according to Chemyakin et al. (2014) and the data error is 0 %. Results for the data error of 15 % are shown too (0.05 accuracy as thin dashed line and 0.1 accuracy as thick dashed line). UP algorithm: results are shown for the noiseless case (solid thick line) and for the data error of 15 % (thick dashed line) without any constraint for the real part in both cases. In the case of (**a, h**) effective radius, (**b, i**) surface area and (**c, j**) volume concentration we show the accumulative sum in dependence of the deviation in percent from the true results. In the case of the (**d, k**) real part, (**e, l**) imaginary part, and single-scattering albedo at (**f, m**) 355 nm and (**g, n**) 532 nm we show absolute deviations. The solid horizontal line shows the 50 % level (accumulative sum). The dashed horizontal line shows the 75 % level, and the dotted horizontal line shows the 90 % level.

ever, Fig. 9e shows that the absolute deviation as well as the error bars are very small for non-absorbing particles.

As already mentioned before, if we want to obtain an absolute accuracy of approximately ±0.03 for the SSA, which is a basic requirement for climate change studies, one needs to test if it is possible to retrieve the imaginary part of the CRI to an absolute error of less than ±0.005. We found that for non- ($0i$) and weak light-absorbing ($0.005i$, $0.01i$) particles this is possible, i.e. the accuracy of ±0.005 in absolute values is achievable; for the real part 1.6 it seems to be even possible for strong light-absorbing particles ($0.03i$, $0.05i$) (Fig. 8a). For the rest of strong light-absorbing cases the absolute error stays within ±0.015 except for two outliers. In summary, 82 % of all examples stay below ±0.005 and 92 % below ±0.01 absolute retrieval error (Fig. 11l). A closer look at Figs. 11l, 8a shows that for the imaginary part, the noisy data case delivers better results than the noiseless one. This fact needs further investigation in future.

3.3.3 Evaluation results for SSA

We evaluate the accuracy of single-scattering albedo for the case of noiseless and noisy optical data. As expected from the absolute retrieval error of the imaginary part of the CRI – which is larger for strong light-absorbing particles – the single-scattering albedo at 355 and 532 nm shows more or less the same behaviour in the noiseless and noisy cases respectively (Fig. 8d–f and Fig. 10a, b, d, e). For noiseless data, 91 % of all examples (355 and 532 nm) are below 6 % retrieval error. 99 % of all examples at 355 nm, and 95 % of all examples at 532 nm are below 10 % retrieval error (Fig. 8c).

With respect to 15 % error of the optical data, we found that in 100 % of all examples the retrieval errors are below 12 % for the single-scattering albedo at 355 and 532 nm. The retrieval errors are well below 6 % for 87 % of all examples at the wavelength 355 nm and for 88 % of all examples at the wavelength 532 nm (Fig. 8c).

Regarding the absolute deviation and the cases of noiseless input data, we found an underestimation of the SSA at 355 nm (and similarly for SSA at 532 nm) for non- and weak light-absorbing particles in the intermediate and coarse mode. For non- and weak light-absorbing particles in the fine-mode, the accuracy limit of ±0.03 (in absolute values) is achieved (Fig. 8d). In the case of noisy data, this underestimation (in the intermediate and coarse mode) of the SSA at 355 and 532 nm occurs for all imaginary parts at real part 1.4 (Fig. 8e, f). For nearly all non- and weak light-absorbing particles with real parts 1.5 and 1.6 the accuracy limit ±0.03 is achieved for all five modes at 532 nm. The accuracy limit is ±0.05 (one outlier) at 355 nm in these cases (Fig. 8e, f). In summary, 70 % (65 %) of all examples of the SSA at 532 nm (355 nm) have an uncertainty of less than ±0.03. In 88 % of all examples (both wavelengths) the uncertainty is below ±0.05 (Fig. 11m, n).

3.3.4 Results in dependence of increasing noise level

We complete our evaluation by showing how the retrieval errors increase with respect to increasing noise level. For that purpose we use one example, i.e. fine-mode particles with gsd $\sigma = 1.7$. We investigated 0, 5, 10, and 15 % input noise. Figure 7c shows the retrieval error of the effective radius for increasing noise level. For the real part 1.4 we clearly see the expected behaviour, i.e. the retrieval error increases with increasing noise level. The same is true for real part 1.5 and strong light-absorbing particles. For all other CRIs, the noise level seems to have less influence on the retrieval results. Additionally, Fig. 7c, i shows that we obtain very good retrieval results for effective radius and for volume concentration, more or less independently of the CRI, in the case of noiseless data. However, Fig. 7i shows that for the volume concentration the noise level itself seems to nearly always have only a minor impact on the results.

Although the surface area concentration is a very stable parameter in the retrieval process. Surprisingly, Fig. 7f shows reverse results with respect to increasing noise level, in particular for real parts 1.5 and 1.6.

The retrieval of the real part of the CRI (Fig. 9c) shows more or less what one would expect for noiseless data. In contrast the noise level seems to be again less important with regard to the relative retrieval error. In contrast the imaginary part of the CRI (Fig. 9f) and the single-scattering albedo (Fig. 10c, f) show no regular behaviour even for noiseless input data, and particularly for strong light-absorbing particles in the case of real parts of 1.5 and 1.6. This behaviour needs further investigation.

4 Summary

We summarize the status of two manually operated data inversion algorithms (TROPOS/UH algorithm and UP algorithm) that are used to derive microphysical parameters of atmospheric particles. The optical input data for these algorithms are collected with EARLINET multiwavelength $(3\beta + 2\alpha)$ Raman lidar. The algorithms provide particle effective radius, surface area and volume concentration, and the real and imaginary part of the complex refractive index. Single-scattering albedo can be computed from the derived particle size distributions and complex refractive indices with Mie-scattering algorithms.

4.1 TROPOS/UH algorithm

We tested the algorithms ability to derive the investigated particle parameters as accurately as possible, if the optical input data are error free and the real part of the complex refractive index is known to 0.1 or 0.05 uncertainty. This latter assumption was introduced in our simulation study in view of results published by Chemyakin et al. (2014), who show that it may be possible to constrain the result for the real part

by the arrange and average algorithm. Investigations with regard to how much the real part can be constrained are under way. Thus, our own study can merely be considered as a first exploratory study. We also investigated the situation of noisy optical input data. We were interested in the question of how much the derived microphysical parameters could deviate from the true results if a certain uncertainty level (5, 10, and 15 %) is reached. For that purpose we distorted the optical spectra (backscatter at 3 wavelengths and extinction at 2 wavelengths) to a maximum for each of these three uncertainty levels, which likely will not occur under real experimental situations. We made the simplifying assumption that an extreme distortion of the optical input spectra leads to an extreme (maximum) deviation of the microphysical parameters from the true values. We acknowledge that this assumption requires more studies to corroborate it.

We find that the effective radius of the PSDs in the fine-mode, the intermediate case (fine- and coarse-mode particles), and the coarse-mode can be retrieved well. Accuracy usually is better than 25 % in all cases, though we notice outliers. Surface area concentration can be retrieved with an accuracy of approximately 10 %. Statistical errors (precision) is just about large enough to include the true values too. In most cases we slightly underestimate the true values. Volume concentration can be retrieved well for the fine-mode, intermediate case, and coarse mode. Accuracy is better than 20 % except for six outliers out of 75 investigated cases.

The real part was retrieved to either 0.1 or 0.05 uncertainty according to the methodology by Chemyakin et al. (2014), and thus is not part of the performance study of our inversion method.

It seems that the accuracy of the retrieved imaginary part in general is better than 30 % for all investigated cases if the true imaginary part is ≥ 0.01. Taking into account precision (statistical error), we find that we still can derive the imaginary part to within ± 50 %.

If the imaginary part is less than 0.01 we can merely decide if the mean value is ≤ 0.015. If we include statistical noise we lose precision as well. This results confirm previous, less systematic studies in which we found that we can merely determine the order of magnitude if the imaginary part is below 0.01.

Single-scattering albedo can be derived to approximately 0.05 in many of the investigated cases. Accuracy is in part better than 0.05, but if we include statistical errors we find deviations up to 0.1. These results however depend on the constraint of the real part. We find that if we can constrain the real part to 0.05, single-scattering albedo can be derived to better than 0.05 as long as its true value is ≤ 0.9. We find a systematic underestimation of single-scattering if the true value is above 0.9. This effect is caused by the fact that our TROPOS/UH inversion method cannot accurately derive imaginary parts if they are less than 0.01. Particularly the fact that we cannot derive an imaginary part of 0 currently poses the greatest challenge.

We finally notice that it may not be so important for the retrieval quality of effective radius, and surface area and volume concentration if the real part cannot be constrained to better than 0.1. Effects are more pronounced with regard to the imaginary part, though accuracy nearly always stays within ±50 % as long as the true imaginary part is ≥ 0.01. Below 0.1 (true) value the imaginary part is overestimated by on average 0.01. The quality of single-scattering albedo also depends on how accurately the real part can be retrieved. If it can be retrieved to 0.05 we find that the retrieved single-scattering (at 355 and 532 nm) in nearly all investigated cases remains within 0.05 accuracy. Significant underestimation occurs if single-scattering albedo is above 0.9. In that case, overestimation of the imaginary part plays a role. The quality of single-scattering albedo becomes notably worse if the real part is retrieved to only 0.1. In that case the accuracy may be only 0.1.

With regard to our extreme error computations we find that results for the most part show an accuracy better than 50 %. In many cases the accuracy is significantly better than 50 %. We find outliers. We do not have sufficient information from this limited set of simulations that allows us to identify a pattern that could explain when the outliers occur. We believe that the inversion method is robust enough to provide microphysical size parameters with ≤ 50 % error (accuracy plus precision), even in cases of 15 % measurement error if the exact error-model of the optical uncertainty bar is unknown. We emphasize that one goal of optical data analysis must be the characterization of the underlying error model of the optical data as it would likely improve significantly the microphysical inversion products.

With regard to the imaginary part we notice that on average the accuracy worsens with increasing noise level. We find more cases of overestimation and underestimation of the true imaginary part compared to the case of correct optical data. Results slightly differ if the real part is constrained to 0.05 or 0.1. However, if we take only results for which the true imaginary part is larger than 0.01, we still find that the retrieved imaginary parts are within ±50 % from the true values.

Single-scattering albedo seems to be significantly underestimated compared to the results for single-scattering albedo in the case of noiseless optical data. In particular we notice a wider scatter of values of single-scattering albedo around the true values. In contrast, in the case of noiseless optical input data, the results for single-scattering albedo remain rather well confined within a narrow band of ±0.05 deviation around the true values, except for the cases of single-scattering albedo ≥ 0.9 (see comments in Sect. 3.1.1).

4.2 UP algorithm

We summarize here our results of the UP algorithm for 15 % noisy input data.

The total surface area concentration is the most stable parameter within the microphysical retrieval procedure, i.e.

99 % of all examples are below 20, and 85 % are below 15 % retrieval error (Table 3). The retrieval error level stays mostly within the input data error/noise level. With respect to the effective radius, 85 % of all examples are below 50 % retrieval error. Additionally, by including only the real parts 1.5 and 1.6, 98 % of these examples are below 40 % retrieval error. For the total volume concentration we found 90 % of all examples are below 50 % retrieval error and, moreover, by including only real parts 1.5 and 1.6 the retrieval error stays below 40 % in 100 % of the examples.

We found for the imaginary part of the CRI that for non- (0i) and weak-absorbing (0.005i, 0.01i) particles in all modes (fine, intermediate, coarse), the accuracy ±0.005 in absolute values is achievable. For real part 1.6, we achieve this accuracy even for strong light-absorbing particles (0.03i, 0.05i). In summary, 82 % of all examples stay below ±0.005. Concerning the real part of CRI, the relative retrieval error is very small, i.e. 100 % of all examples are below 6 % retrieval error. With respect to the absolute deviation, 64 % of all examples are below ±0.05, and 100 % are well below ±0.1.

The evaluation statistics show that the relative retrieval errors of effective radius and total volume concentration are prominently larger for the real part 1.4. The retrieval of the imaginary part of the CRI has significantly larger error bars for strongly light-absorbing particles of 0.03i and 0.05i. The error bar behaviour is monotonically increasing with an increasing light absorption level. These observations are in agreement with mathematical theoretical studies which show that in general the degree of ill-posedness of the inverse problem grows both with increasing imaginary part and decreasing real part.

For fine-mode particles with gsd $\sigma = 1.5$ the relative retrieval errors of the effective radius and real part of the CRI are largest for almost all CRIs. The same is true for the total surface area concentration but here only for real part 1.4, or for total volume concentration only for real part 1.6. For coarse-mode particles with $\sigma = 2.3$, the post-processing procedure is very complicated and time consuming since the CRI domain often has neither a diagonal structure nor any other well-defined domain. It is sometimes even speckled. Selecting the CRI associated with the PSD may be difficult.

As expected from the absolute retrieval error of the imaginary part of the CRI, which is larger for strong absorbing particles, the single scattering albedo shows more or less the same behaviour for 355 and 532 nm. For the single scattering albedo, the relative errors stay below 12 % in 100 % of all examples. In more detail, 87 % of all examples for 355 nm and 88 % of all examples for 532 nm are well below 6 %. With respect to the absolute error we achieve an accuracy limit of ±0.03 for nearly all non- and weak-absorbing particles in all three modes (fine, intermediate, coarse) and for which real parts are 1.5 and 1.6. We note that this accuracy limit is a basic requirement for climate change studies. In summary, 70 % (65 %) of all examples stay below ±0.03 for SSA 532 nm (355 nm).

4.3   Comparison of results

Figure 11 shows the accumulative distribution, i.e. the number of case studies for which each of the investigated parameters deviates less than a given value (shown on the *x* axis of each plot).

With regard to the UH/TROPOS algorithm we see that, if we have knowledge of the real part, approximately 50 % of all solutions for effective radius deviate by less than 5–8 % from the true results. We find that 75 % of all solutions have uncertainties of less than 15 %. Each solution deviates less than 50 % from the true results.

If uncertainties of the optical data are 15 %, only 50 % of all solutions have errors less than 50 %. The uncertainties of the other 50 % of the solutions may be considerably larger. We note that this is an extreme estimate, as we distorted the optical backscatter and extinction spectra to their maximum.

In the case of surface area concentration, 75 % of all simulated cases deviate less than 10 % from the true results, regardless of whether the real part is known to 0.05 or 0.1 accuracy. If we introduce uncertainties, we find that 75 % of all solutions deviate less than 20 % from the true results.

In the case of volume concentration we find that 50 % of all solutions deviate less than 10 % from the true values, 75 % of all cases deviate less than 20 % from the true values, and in 90 % of all simulated cases the retrieval error is less than 50 %. If we assume uncertainties of 15 % for the optical data, volume concentration may deviate less than 20 % from the true value in 50 % of all simulation cases.

With respect to the imaginary part the deviation is less than approximately 0.006 in 75 % of all cases. If measurement errors are included the deviation is ≤ 0.01 in 75 % of all cases.

Single-scattering albedo at 355 and 532 nm can be derived to better than 0.04 in 75 % of all cases if the real part is known to 0.05–0.1. In 90 % of all simulation cases we can retrieve single-scattering albedo to better than 0.06. If we introduce measurement errors the uncertainty of single-scattering albedo is ≤ 0.06 in 50 % of all cases and ≤ 0.08 in 75 % of all cases.

The next paragraphs describe the retrieval results of the UP algorithm. With regard to the noiseless data case, we find that approximately 50 % of all solutions for the effective radius deviate by less than 5, and 90 % of all solutions deviate by less than 14 % from the true values. Each solution deviates less than 25 % from the true values. In case of uncertainties of 15 % of the optical data, only 50 % of all solutions have errors less than 25, and 85 % of all solutions stay below 50 % deviation.

The total surface area concentration proves to be the most stable parameter as already mentioned. That is, 75 % of all simulated cases have uncertainties less than 8, and 95 % of all cases deviate less than 15 % from the true values. Even in the data case of 15 % noise, 85 % of all investigated cases deviate by less than 15 % from the true values, i.e. the data error is not amplified during the retrieval process.

Concerning the noiseless data case of the total volume concentration, we find that 50 % of all examined cases deviate less than 5 % from the true results, and 90 % of all cases have less than 15 % uncertainty. If we assume 15 % optical data error, the qualitative behaviour of the uncertainty behaviour of volume concentration is similar to the uncertainty behaviour of effective radius. Only 47 % of all retrieved solutions show errors less than 20 %, and 90 % of all solutions deviate less than 50 % from the true values.

With respect to the real part of the CRI we find that the differences between the retrieved and the true values in the noiseless data case are ≤ 0.04 in 75 % (≤ 0.06 in 90 %) of all investigated cases. In the noisy data case the deviation is ≤ 0.055 in 75 % of all cases. In both cases all solutions deviate less than 0.1 from the true values.

With regard to the imaginary part we observe that there is no big difference between the noiseless and noisy data case. On average the deviation between retrieved values and true values is ≤ 0.005 in 75 % of all simulated cases.

The single-scattering albedo at 355 and 532 nm shows a similar qualitative behaviour as the imaginary part of the CRI, namely, there is only a quite small difference between the noiseless and noisy data case. On average the deviation is ≤ 0.03 for SSA at 532 nm (≤ 0.04 for SSA at 355 nm) in 72 % of all investigated cases. The deviation is ≤ 0.05 for SSA at 355 and 532 nm in 90 % of all cases.

Table 3 presents a few numbers regarding the performance of the two algorithms. The real part was assumed known to within ±0.05 or ±0.1, in the case of the TROPOS/UH algorithm. No such assumptions were made for the simulations with the UP algorithm.

We think that we can achieve a retrieval accuracy of the size parameters of 20 %. Any better value seems unrealistic in view of measurement errors and errors caused by the inversion method. A retrieval accuracy of 0.005 of the imaginary part may be possible. We furthermore target a retrieval accuracy of 0.05 for single-scattering albedo.

We note that both algorithms have advantages and drawbacks. The simulation studies use the same size distributions and complex refractive indices but follow different aims in testing each algorithm. Therefore, a truthful comparison between the algorithms is not possible. Nevertheless, it is remarkable that the surface area concentration is found to be the most stable microphysical parameter in the retrieval process. For example, in the noiseless optical data case and the case of surface area concentration, a deviation of less than 20 % compared to the true result is achieved in 95 % (for TROPOS/UH) and 97 % (for UP) of all investigated cases. Moreover, for the case of 15 % noise of the optical data, volume concentration deviates less than 20 % from the true values in 48 % (for TROPOS/UH) and 47 % (for UP) of all simulated cases (Table 3).

5 Outlook

Future development work of the TROPOS/UH algorithm will focus on deriving fine mode and coarse mode particle parameters separately. We want to investigate to what extent a wavelength-dependent complex refractive index of aerosol particles influences the quality of the retrieval results. The reason for this study is that we assume a wavelength-independent refractive index in our retrievals. We want to investigate whether we can derive the single-scattering albedo with less uncertainty if the true value of SSA is above 0.9.

In future the Potsdam group will improve the presented software using the two regularization methods, namely TSVD and iterative Páde method in parallel to utilize advantages of both methods simultaneously.

Currently the Potsdam group is investigating a microphysical retrieval procedure via regularization for non-spherical particles, in particular, spheroidal particles, i.e. oblate and prolate particles. In future a user-friendly software tool for non-spherical particles could be developed.

References

Alados-Arboledas, L., Müller, D., Guerrero-Rascado, J. L., Navas Guzmán, F., , Pérez-Ramírez, D., and Olmo, F. J.: Optical and microphysical properties of fresh biomass burning aerosol retrieved by Raman lidar, and star- and sun-photometry, Geophys. Res. Letts., 38, L01807, doi:10.1029/2010GL045999, 2011.

Althausen, D., Müller, D., Ansmann, A., Wandinger, U., Hube, H., Clauder, E., and Zörner, S.: Scanning 6-wavelength 11-channel aerosol lidar, J. Atmos. Ocean. Tech., 17, 1469–1482, 2000.

Amato, U., Carfora, M. F., Cuomo, V., and Serio, C.: Objective algorithms for the aerosol problem, Appl. Opt., 34, 5442–5452, 1995.

Ansmann, A. and Müller, D.: Lidar and atmospheric aerosol particles, in: Lidar. Range-Resolved Optical Remote Sensing of the Atmosphere, edited by: Weitkamp, C., 105–141, Springer, New York, 2005.

Böckmann, C.: Hybrid regularization method for the ill-posed inversion of multiwavelength lidar data to determine aerosol size distributions, Appl. Optics, 40, 1329–1342, 2001.

Böckmann, C. and Wauer, J.: The influence of spheroids on the inversion in the retrieval of microphysical particle parameters from lidar data, Proc. SPIE, 4150, 282–289, 2001.

Böckmann, C. and Wauer, J.: Algorithms for the inversion of light scattering data from uniform and non–uniform particles, Journal of Aerosol Sci., 32, 49–61, 2001.

Böckmann, C., Mironova, I., Müller, D., Schneidenbach, L., and Nessler, R.: Microphysical aerosol parameters from multiwavelength lidar, Journal of the Optical Society of America A, Optics, image science, and vision, 22, 518–528, 2005.

Böckmann, C. and Kirsche, A.: Iterative regularization method for lidar remote sensing, Comput. Phys. Commun., 174, 607–615, 2006.

Böckmann, C. and Pornsawad, P.: Iterative Runge–Kutta–type methods for nonlinear ill–posed problems, Inverse Problems, 24, 025002 , doi:10.1088/0266-5611/24/2/025002, 2008.

Böckmann, C., Osterloh, L., Rother, T., Serikov, I., Linne, H., and Kinne, S.: Optical and Microphysical Properties from Raman Lidar and Depolarization Data, p. 4, presented on 26th International Lidar Radar Conference, 25–29 June 2012, Porto Heli, Greece, 2012.

Böckmann, C. and Osterloh, L.: Runge–Kutta type regularization method for inversion of speroidal particle distribution from limited optical data, Inverse Problems in Science and Engineering, 22, 150–165, 2014.

Bohren, C. F. and Huffman, D. R. (Eds.): Absorption and scattering of light by small particles, John Wiley, Hoboken, NJ, ISBN-13: 978-0-471-29340-8, 1983.

Chemyakin, E., Müller, D., Burton, S., Kolgotin, A., Hostetler, C., and Ferrare, R.: Arrange and average algorithm for the retrieval of aerosol parameters from multiwavelength highspectral-resolution lidar/Raman lidar data, Appl. Opt., 53, 7252–7266, 2005.

Doicu, A., Trautmann, T., and Schreier, F.: Numerical Regularization for Atmospheric Inverse Problems, Springer Berlin Heidelberg, ISBN-13: 978-3-642-05438-9, 2010.

Engl, H. W., Hanke, M., and Neubauer, A.: Regularization of Inverse Problems, Springer Netherlands, ISBN-13: 978-0-7923-4157-4, 1996.

Eixmann, R., Böckmann, C., Fay, B., Matthias, V., Mattis, I., Müller, D., Kreipl, S., Schneider, J., and Stohl, A.: Tropospheric aerosol layers after a cold front passage in January 2000 as observed at several stations of German Lidar Network, Atmos. Res., 63, 39–58, 2002.

Hansen, P. C.: Discrete Inverse Problems, Insight and Algorithms, SIAM, 2010.

Heintzenberg, J., Müller, H., Quenzel, H., and Thomalla, E.: 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.

Hoffmann, A., Osterloh, L., Stone, R., Lampert, A., Ritter, C., Stock, M., Tunved, P., Hennig, T., Böckmann, C., Li, S.-M.,

Eleftheriadis, K., Maturilli, M., Orgis, T., Herber, A., Neuber, R., and Dethloff, K.: Remote sensing and in-situ measurements of tropospheric aerosol, a PAMARCMiP case study, Atmos. Environ., 52, 56–66, 2012.

Kirsche, A. and Böckmann, C.: Páde iteration method for regularization, Appl. Math. Comput,, 180, 648–663, 2006.

Mishchenko, M. I., Cairns, B., Hansen, J. E., Travis, L. D., Burg, R., Kaufman, Y. J., Martins, J. V., and Shettle, E. P.: Monitoring of aerosol forcing of climate from space: Analysis of measurement requirements, J. Quant. Spectrosc. Ra., 88, 149–161, doi:10.1016/j.jqsrt.2004.03.030, 2004.

Müller, D., Wandinger, U., and Ansmann, A.: Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: Theory, Appl. Opt., 38, 2346–2357, 1999a.

Müller, D., Wandinger, U., and Ansmann, A.: Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: Simulation, Appl. Opt., 38, 2358–2368, 1999b.

Müller, D., Franke, K., Wagner, F., Althausen, D., Ansmann, A., Heintzenberg, J., and Verver, G.: Vertical profiling of optical and physical particle properties over the tropical Indian Ocean with six-wavelength lidar, 2. Case studies, J. Geophys. Res., 106, 28577–28595, 2001a.

Müller, D., Wandinger, U., Althausen, D., and Fiebig, M.: Comprehensive particle characterization from three-wavelength Raman-lidar observations, Appl. Opt., 40, 4863–4869, 2001b.

Müller, D., Mattis, I., Wandinger, U., Ansmann, A., Althausen, D., and Stohl, A.: Raman lidar observations of aged Siberian and Canadian forest-fire smoke in the free troposphere over Germany in 2003: Microphysical particle characterization, J. Geophys. Res., 110, D17201, doi:10.1029/2004JD005756, 2005.

Müller, D., Veselovskii, I., Kolgotin, A., Tesche, M., Ansmann, A., and Dubovik, O.: Vertical profiles of pure dust and mixed smoke-dust plumes inferred from inversion of multiwavelength Raman/polarization lidar data and comparison to AERONET retrievals and in situ observations, Appl. Opt., 52, 3178–3202, 2013.

Noh, Y. M., Müller, D., Mattis, I., Lee, H., and Kim, Y. J.: Vertically resolved light-absorption characteristics and the influence of relative humidity on particle properties: Multiwavelength Raman lidar observations of East Asian aerosol types over Korea, J. Geophys. Res., 116, D06206, doi:10.1029/2010JD014873, 2011.

Osterloh, L., Pérez, C., Böhme, D., Baldasano, J. M., Böckmann, C., Schneidenbach, L., and Vicente, D.: Parallel software for retrieval of aerosol distribution from LIDAR data in the framework of EARLINET-ASOS, Comput. Phys. Commun., 180, 2095–2102, 2009.

Osterloh, L., Böckmann, C., Mamouri, R.-E., and Papayannis, A.: An Adaptive Base Point Algorithm for the Retrieval of Aerosol Microphysical Properties, Open Atmospheric Science Journal, 5, 61–73, 2011.

Osterloh, L., Böckmann, C., Nicolae, D., and Nemuc, A.: Regularized Inversion of Microphysical Atmospheric Particle Parameters: Theory and Application, J. Comput. Phys., 237, 79–94, 2013.

Papayannis, A., Mamouri, R. E., Nenes, A., Tsaknakis, G., Georgoussis, G., Avdikos, G., Böckmann, C., Osterloh, L., Eleftheriadis, K., Böhme, D., and Stohl, A.: Optical, microphysical

and chemical properties of tropospheric aerosols retrieved by a 6-wavelength RAMAN lidar system during a biomass burning event over Athens, Greece", Proceeding of the 24th International laser radar conference, 381–384, Boulder, Colorado, USA, 23–27 June 2008.

Phillips, D. L.: A technique for the numerical solution of certain integral equations of the first kind, J. Assoc. Comput. Mach., 9, 84–97, 1962.

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

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific, ISBN-13: 978-981-02-2740-1, 2000.

Samaras, S., Nicolae, D., Böckmann, C., Vasilescu, J., Binietoglou, J., Labzovskii, L., Toanca, F., and Papayannis, A.: Using Raman-lidar-based regularized microphysical retrievals and Aerosol Mass Spectrometer measurements for the characterization of biomass burning aerosols, Journal of Computational Physics, 299, 156–174, 2015.

Tesche, M., Müller, D., Groß, S., Ansmann, A., Althausen, D., Freudenthaler, V., Weinzierl, B., Veira, A., and Petzold, A.: Optical and microphysical properties of smoke over Cape Verde inferred from multiwavelength lidar measurements, Tellus, Ser. B, 63, 677–694, 2011.

The EARLINET publishing group 2000–2010: Adam, M., Alados-Arbolas, L., Althausen, D., Amiridis, V., Amodeo, A., Ansmann, A., Apituley, A., Arshinov, Y., Balis, D., Belegante, L., Bobrovnikov, S., Boselli, A., Bravo-Aranda, J. A., Bosenberg, J., Carstea, E., Chaikovsky, A., Comeron, A., D'Amico, G., Daou, D., Dreischuh, T., Engelmann, R., Finger, F., Freudenthaler, V., Garcia-Vizcaino, D., Garcia, A. J. F., Geiss, A., Giannakaki, E., Giehl, H., Giunta, A., de Graaf, M., Granados-Munoz, M. J., Grein, M., Grigorov, I., Gross, S., Gruening, C., Guerrero-Rascado, J. L., Haeffelin, M., Hayek, T., Iarlori, M., Kanitz, T., Kokkalis, P., Linné, H., Madonna, F., Mamouri, R.-E., Matthias, V., Mattis, I., Menendez, F. M., Mitev, V., Mona, L., Morille, Y., Muñoz, C., Müller, A., Müller, D., Navas-Guzman, F., Nemuc, A., Nicolae, D., Pandolfi, M., Papayannis, A., Pappalardo, G., Pelon, J., Perrone, M. R., Pietruczuk, A., Pisani, G., Potma, C., Preissler, J., Pujadas, M., Putaud, J., Radu, C., Ravetta, F., Reigert, A., Rizi, V., Rocadenbosch, F., Rodriguez, A., Sauvage, L., Schmidt, J., Schnell, F., Schwarz, A., Seifert, P., Serikov, I., Sicard, M., Silva, A. M., Simeonov, V., Siomos, N., Sirch, T., Spinelli, N., Stoyanov, D., Talianu, C., Tesche, M., De Tomasi, F., Trickl, T., Vaughan, G., Volten, H., Wagner, F., Wandinger, U., Wang, X., Wiegner, M., and Wilson, K. M.:EARLINET all observations (2000–2010), World Data Center for Climate (WDCC), doi:10.1594/WDCC/EN_all_measurements_2000-2010, 2014.

Tikhonov, A. N. and Arsenin, V. Y. (Eds.): Solutions of Ill-posed Problems, John Wiley, New York, ISBN-10: 0470991240, 1977.

Twomey, S.: The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements, J. Franklin Inst., 279, 95–109, 1965.

Twomey, S. (Ed.): Introduction to the mathematics of inversion in remote sensing and indirect measurements, Elsevier, Amsterdam, ISBN-13: 978-0444415479, 1977.

Veselovskii, I., Kolgotin, A., Griaznov, V., Müller, D., Wandinger, U., and Whiteman, D. N.: Inversion with regularization for the retrieval of tropospheric aerosol parameters from multiwavelength lidar sounding, Appl. Opt., 41, 3685–3699, 2002.

Veselovskii, I., Kolgotin, A., Griaznov, V., Müller, D., Franke, K., and Whiteman, D. N.: Inversion of multiwavelength Raman lidar data for retrieval of bimodal aerosol size distribution, Appl. Opt., 43, 1180–1195, 2004.

Veselovskii, I., Dubovik, O., Kolgotin, A., Lapyonok, T., Girolamo, P. D., Summa, D., Whiteman, D. N., Mishchenko, M. I., and Tanré, D.: Application of randomly oriented spheroids for retrieval of dust particle parameters from multiwavelength lidar measurements, J. Geophys. Res., 115, D21203, doi:10.1029/2010JD014139, 2010.

Vogel, C. R.: Computational Methods for Inverse Problems, SIAM Philadelphia, ISBN-10: 0898715504, 2002.

Wandinger, U. and Ansmann, A.: Experimental determination of the lidar overlap profile with Raman lidar, Appl. Opt., 41, 511–514, 2002.

Wandinger, U., Müller, D., Böckmann, C., Althausen, D., Matthias, V., Bösenberg, J., Weiß, V., Fiebig, M., Wendisch, M., Stohl, A., and Ansmann, A.: Optical and microphysical characterization of biomass–burning and industrial–pollution aerosols from–multiwavelength lidar and aircraft measurements, J. Geophys. Res.-Atmos., 107, LAC 7-1–LAC 7-20, doi:10.1029/2000JD000202, 2002.

Zuev, V. E. and Naats, I. E. (Eds.): Inverse problems of lidar sensing of the atmosphere, Springer, Berlin, ISBN-13: 978-3-540-38802-9, 1983.