

# Statistical estimation of stratospheric particle size distribution by combining optical modelling and lidar scattering measurements

J. Jumelet, S. Bekki, C. David, and P. Keckhut

UPMC Paris 6, Service d'Aéronomie du CNRS/IPSL, Paris, France

Received: 4 March 2008 – Published in Atmos. Chem. Phys. Discuss.: 20 May 2008

Revised: 21 July 2008 – Accepted: 14 August 2008 – Published: 10 September 2008

**Abstract.** A method for estimating the stratospheric particle size distribution from multiwavelength lidar measurements is presented. It is based on matching measured and model-simulated backscatter coefficients. The lidar backscatter coefficients measured at the three commonly used wavelengths 355, 532 and 1064 nm are compared to a precomputed look-up table of model-calculated values. The optical model assumes that particles are spherical and that their size distribution is unimodal. This inverse problem is not trivial because the optical model is highly non-linear with a strong sensitivity to the size distribution parameters in some cases. The errors in the lidar backscatter coefficients are explicitly taken into account in the estimation. The method takes advantage of the statistical properties of the possible solution cluster to identify the most probable size distribution parameters. In order to discard model-simulated outliers resulting from the strong non-linearity of the model, solutions farther than one standard deviation of the median values of the solution cluster are filtered out, because the most probable solution is expected to be in the densest part of the cluster. Within the filtered solution cluster, the estimation algorithm minimizes a cost function of the misfit between measurements and model simulations.

Two validation cases are presented on Polar Stratospheric Cloud (PSC) events detected above the ALOMAR observatory (69° N – Norway). A first validation is performed against optical particle counter measurements carried out in January 1996. In non-depolarizing regions of the cloud (i.e. spherical particles), the parameters of an unimodal size distribution and those of the optically dominant mode of a bimodal size distribution are quite successfully retrieved, especially for the median radius and the geometrical standard deviation. As expected, the algorithm performs poorly when

solid particles drive the backscatter coefficient. A small bias is identified in modelling the refractive index when compared to previous works that inferred PSC type Ib refractive indices. The accuracy of the size distribution retrieval is improved when the refractive index is set to the value inferred in the reference paper.

Our results are then compared to values retrieved with another similar method that does not account for the effect of the measurements errors and the non-linearity of the optical model on the likelihood of the solution. The case considered is a liquid PSC observed over northern Scandinavia on January 2005. An excellent agreement is found between the two methods when our algorithm is applied without any statistical filtering of the solution cluster. However, the solution for the geometrical standard deviation appears to be rather unlikely with a value close to unity ( $\sigma \approx 1.04$ ). When our algorithm is applied with solution filtering, a more realistic value of the standard deviation ( $\sigma \approx 1.27$ ) is found. This highlights the importance of taking into account the non linearity of the model together with the lidar errors, when estimating particle size distribution parameters from lidar measurements.

## 1 Introduction

Stratospheric particles play an important role in atmospheric chemistry (WMO, 2007) and, in the case of large volcanic eruption, in the earth radiative budget (Robock, 2005). For example, Polar Stratospheric Clouds (PSC) are a key element in polar ozone depletion by providing the surfaces for chlorine activation through heterogeneous chemistry (Peter, 1997; Charlson and Heintzenberg, 1995). Characterizing and understanding stratospheric particles remains a major scientific issue. The most important characteristics of stratospheric particles are their size distribution, shape and composition. This information can be used to calculate radiative properties of particles, rates of heterogeneous chemistry or



Correspondence to: J. Jumelet  
(julien.jumelet@aero.jussieu.fr)

gravitational sedimentation (ASAP, 2006). But, very few measurements give access to the full size distribution of stratospheric particles. They are mostly balloonborne in-situ measurements (Hofmann and Rosen, 1980; Hofmann, 1990; Deshler et al., 2003) and remain very rare. Retrievals of stratospheric particles size distribution have been performed using aerosol extinctions from satellite measurements at multiple wavelengths, for instance using SAGE (Stratospheric Aerosol and Gas Experiment) and SAGE II data (McCormick et al., 1979). Here, we focus our attention on aerosol backscatter coefficients from lidar data. Lidar measurements have been used to detect stratospheric particles since the early 1960s (Junge et al., 1961; Fiocco and Smullins, 1963). They have been used as a proxy for the stratospheric aerosol loading (David et al., 1997; ASAP, 2006).

Direct determination of the particle size distribution from uniwavelength lidar measurements is theoretically precluded because of a lack of constraints (Müller and Quenzel, 1985). However, there have been attempts to characterize particle size distribution from multiwavelength Raman lidar data (Müller et al., 1998; Veselovskii et al., 2002, 2005). This typical inverse problem can be presented in the following way: given a set of lidar observations and optical modelling of the observations, how can the most probable aerosol size distribution be estimated? This problem is often addressed by assuming model linearity and normally distributed errors. Most retrieval methods are only based on the least square criterion (i.e. minimal discrepancy between measured and model-simulated quantities, variance weighted by the errors). Note that, in some cases, the criterion is used without even considering observation errors. This least-square criterion should lead to the maximum likelihood estimate and to a minimum variance for the analysis error (error in the estimation) if all the errors are Gaussian, unbiased and the model quasi-linear (Lahoz et al., 2007).

In our case, the model is highly non linear (see Sect. 3.3). It is very sensitive to the size distributions parameters in some cases, and, as a consequence, the least square estimator may not always give the most probable solution depending on the magnitude of the observation errors. In addition, even if the errors in the observables were to be normally distributed, the non linearity of the model would result in a Probability Density Function (PDF) of the solution that is not necessarily Gaussian, meaning that the PDF cannot be described by its first and second moment. Therefore, the entire shape of the solution PDF has to be considered.

The best match estimator, identified as the minimal difference between measured and modelled quantities in the least square sense, is usually the most probable solution when errors in the observables are negligible and the model quasi-linear. In our case, errors in the lidar retrieved backscatter coefficient are not at all negligible, and so it is important to check whether the best match solution is the most probable solution.

It is worth pointing out that the need to find a physical solution has led to added constraints on the form or properties of the solution. An example of such a constraint is the smoothness of the solution for vertical profile retrievals in regularization techniques (Tikhonov and Arsenin, 1977; Müller et al., 1999, 2000; Tarantola, 2005). Regularization methods account for the lidar errors in the retrieval process when setting the regularization parameter choice rules (Müller et al., 1999). An accurate retrieval also often requires numerous optical quantities (backscatter and/or extinction coefficients) at different wavelengths. With enough constraints, regularization techniques could also enable the determination of the refractive index (Veselovskii et al., 2004).

In the stratosphere, the particle extinction coefficient is measured with a limited accuracy, because very few photons are detected by the lidar at stratospheric altitudes. In addition, most of the best-equipped lidar stations only monitor the stratosphere at three wavelengths, typically at 355, 532 and 1064 nm. Because of these two limitations, the regularization technique is not necessarily the most suited approach for retrieving stratospheric particle size distributions from lidar data. In this paper, we explore an alternative method based on the comparison between measured and model-simulated Backscatter Coefficients (*BC*), using a Monte Carlo approach. The formulation of the method is inspired by the works of Beyerle et al. (1994) and Mehrtens et al. (1999). The retrieval algorithm minimizes a cost function of the misfit between measurements and model simulations with the control variables being the parameters of the PSC size distribution. The errors in the measurements are explicitly taken into account in the search for a solution. A cluster-based filtering of the solution pool ensures both stability and reliable error estimation. The refractive index is determined from the particle composition calculated by the microphysical model, taking into account both sulphuric acid aerosols and liquid type Ib PSC particles.

The paper is organized as follow. As the method is applied to the multiwavelength lidar measurements performed at the Arctic Lidar Observatory for Middle Atmosphere Research (ALOMAR–69° N–Norway), the setup of the Rayleigh/Mie/Raman (RMR) lidar is briefly presented in Sect. 2. The size distribution retrieval methodology is described in Sect. 3, introducing both the microphysical model and the size distribution retrieval algorithm. The fourth and fifth sections are dedicated to the validation against size-resolved PSC measurements and a similar approach using lidar measurements performed at ALOMAR. The last section is devoted to the summary and conclusions.

## 2 Lidar data

Two validation cases are presented in the result section: both feature PSC events observed above ALOMAR (69° N/16° E) where lidar measurements are performed routinely. The

ALOMAR Rayleigh/Mie/Raman lidar emits laser pulses at 355, 532 and 1064 nm. The beam is produced by a 30 Hz repetition rate pulsed Nd-YAG twin laser. The backscattered light is received by 180 cm diameter Cassegrain telescopes with a field-of-view of  $180 \mu\text{rad}$  and detected with photomultiplier tubes in photon-counting mode (von Zahn et al., 2000). The vibrational Raman at 387 and 607 nm (associated with the 355 and 532 nm wavelengths respectively) and the rotational Raman measurements at 529 and 530 nm are also performed simultaneously. The measurement integration time is typically 3 min with a vertical resolution of 150 m. Data are acquired with the lidar pointing to the zenith.

### 3 Methodology of the size distribution estimation

The particle size distribution is retrieved from comparisons between measured and model-simulated backscatter coefficients, taking the measurement errors into account. In the first phase, the model calculates the chemical composition according to the specific environmental conditions (pressure, temperature, total  $\text{HNO}_3$ , total  $\text{H}_2\text{O}$ ) for a range of particle size distributions. Then the refractive index is determined from the composition. In the second phase, we generate a *BC* look-up table as a function of the particle size distribution parameters. A Mie scattering module (code from Bohren and Huffman, 1983) is used to calculate the aerosol *BC* at different wavelengths (within the visible and near-infrared domain), for each model-simulated size distribution and chemical composition. In the third phase, the solution is searched by comparing the look-up table and the measurements, taking the errors in the lidar measurements into account. The three steps are summarized in Fig. 1.

We assume in our case that the stratospheric particle size distribution can be represented by a lognormal size distribution (Pinnick et al., 1976; WMO, 2007):

$$n(r) = \frac{N_0}{\sqrt{2\pi} \cdot r \cdot \ln(\sigma)} \times \exp\left(-\frac{\ln^2(r/r_m)}{2 \ln^2(\sigma)}\right) \quad (1)$$

where  $N_0$  is the total number of particles per unit volume and  $\sigma$  the geometrical standard deviation (hereafter called standard deviation) around the median radius  $r_m$ .

#### 3.1 Refractive index modelling

In the first phase, the model calculates the particle composition, ranging from a binary  $\text{H}_2\text{SO}_4/\text{H}_2\text{O}$  solution to a ternary  $\text{H}_2\text{SO}_4/\text{HNO}_3/\text{H}_2\text{O}$  (STS) solution (Larsen, 2000; Luo et al., 1996; Krieger et al., 2000) for every size distribution in the look-up table. The composition of the condensed phase is assumed to be in thermodynamic equilibrium with the gas phase. In order to determine the particle equilibrium composition (weight fractions of sulphuric acid, nitric acid and water), one needs to solve a set of 2 non-linear equations describing the equality between the partial pressures of  $\text{HNO}_3$

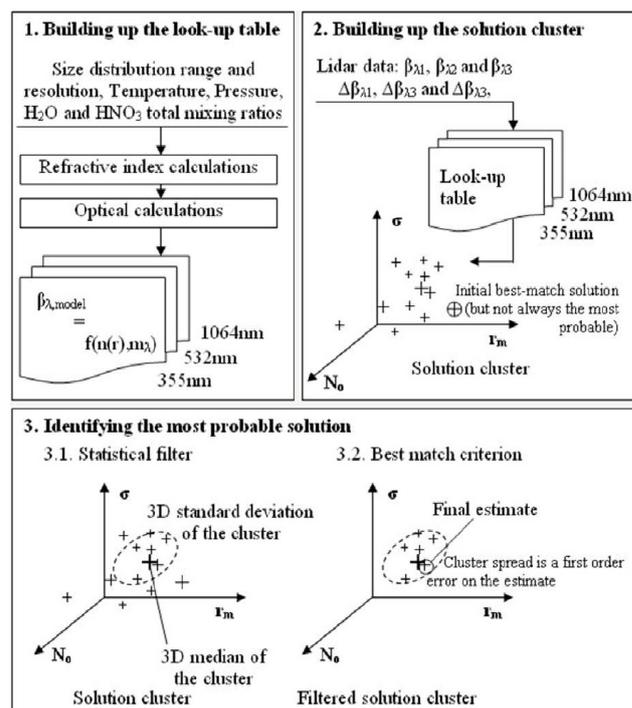
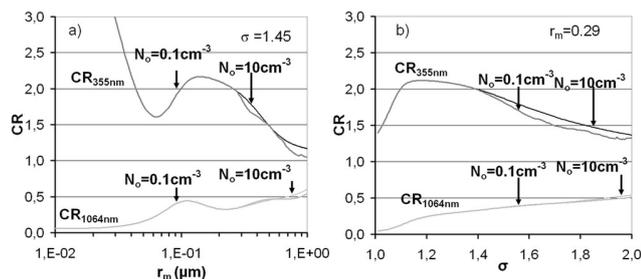


Fig. 1. The 3 steps of the retrieval methodology.

and  $\text{H}_2\text{O}$  just over the surface of the condensed phase and the partial pressures in the gas phase. The model is initialized with a value of the condensed  $\text{H}_2\text{SO}_4$  mixing ratio, the total (gaseous + condensed) amounts of  $\text{HNO}_3$  and  $\text{H}_2\text{O}$  and it then redistributes  $\text{HNO}_3$  and  $\text{H}_2\text{O}$  between the gas and the condensed phases according to the calculated particle composition. The iterative procedure of the equilibrium composition calculation ensures that the gas phase and condensed  $\text{HNO}_3$  and  $\text{H}_2\text{O}$  is equal to the initial total  $\text{HNO}_3$  and  $\text{H}_2\text{O}$ . The model then derives the condensed aerosol mass concentration (or aerosol volume concentration) which is compared to the value of the specified particle size distribution. The  $\text{H}_2\text{SO}_4$  mixing ratio is adjusted iteratively (and so is the composition) in order for the calculated aerosol volume density to match the one of the input size distribution. Finally, the refractive index is calculated from the equilibrium composition (Luo et al., 1996). The refractive indices at 355 nm and 1064 nm are assumed to be equal to the refractive indices at 360 nm and 1000 nm respectively because of the lack of available data beyond the 360–1000 nm range. Modelling the composition allows us to account for the rapid variations of the particle composition, and hence, the refractive index, around the PSC type Ib temperature threshold (Larsen, 2000; Carslaw et al., 1997) instead of assuming a constant refractive index whatever the environmental conditions. In addition, the retrieval algorithm is only strictly valid for spherical particles (e.g. liquid ternary solution PSCs). For this type of PSCs, the absorption part of the refractive index can be neglected in optical scattering calculations (Beyer et al., 1996).



**Fig. 2.** Colour ratios at 355 ( $CR_{355\text{nm}}$ ) and 1064 nm ( $CR_{1064\text{nm}}$ ) versus  $r_m$  (a) and  $\sigma$  (b). Reference values for the non varying parameters are,  $\sigma = 1.45$  (a),  $r_m = 0.29 \mu\text{m}$  (b),  $N_o = 0.1$  and  $10 \text{ cm}^{-3}$  in both (a) and (b).

### 3.2 Backscatter modelling

The particle backscatter coefficient is simulated using Mie theory. It is strictly valid for spherical particles only. Consequently, our size distribution algorithm can only be applied to lidar measurements of spherical particles such as supercooled sulphuric acid aerosol particles or type Ib PSC. The model-simulated size backscatter coefficient,  $\beta_{\text{Mie}}$  ( $\text{m}^{-1} \cdot \text{sr}^{-1}$ ), is expressed as:

$$\beta_{\text{Mie},\lambda} = \int_0^{\infty} n(r) \frac{d\sigma_b}{d\Omega}(r, \lambda, m) \cdot dr. \quad (2)$$

where  $\lambda$  is the incident wavelength,  $n(r)$  is the size distribution, the number of particles at the radius  $r$  between  $r$  and  $r+dr$ ,  $m$  the refractive index and  $d\sigma_b/d\Omega$  the Mie particle backscattering differential cross section.

The other optical quantity used in the retrieval algorithm is the colour ratio  $CR_\lambda$  (or  $CR$ ) which is the  $BC$  at the wavelength  $\lambda$  normalised by the  $BC$  at the most sensitive wavelength of the lidar system, 532 nm:

$$CR_\lambda = \frac{\beta_{\text{Mie},\lambda}}{\beta_{\text{Mie},532\text{nm}}}. \quad (3)$$

### 3.3 Size distribution retrieval methodology

An optical module coupled to a size-resolving aerosol model is used to calculate the three backscatter coefficients ( $\beta_{355\text{nm}}$ ,  $\beta_{532\text{nm}}$  and  $\beta_{1064\text{nm}}$ ) and the two associated colour ratios ( $CR_{355\text{nm}}$  and  $CR_{1064\text{nm}}$ ) as a function of the size distribution parameters. In principle, the  $CR$  should not depend on  $N_o$  because  $N_o$  vanishes when  $CR$  are formed. But, in our size distribution algorithm, that is not the case. In the look-up table, the refractive index also varies with  $N_o$  because  $N_o$ , along with  $\sigma$  and  $r_m$ , determines the aerosol volume concentration and, hence, the condensed mass of  $\text{HNO}_3$  and  $\text{H}_2\text{O}$ . As the total  $\text{HNO}_3$  and  $\text{H}_2\text{O}$  is fixed, the partitioning between gas phase and condensed phase as well as the aerosol composition depend on aerosol volume concentration. For example,

if  $N_o$  is very small (or large), most of the  $\text{HNO}_3$  is in the gas-phase (or condensed phase), and so, the particle equilibrium composition would correspond to high (or low) partial pressures of gaseous  $\text{HNO}_3$ . As a result, even when  $\sigma$  and  $r_m$  are kept constant, different  $N_o$  give different aerosol compositions, different refractive indices, and hence, different  $CR$ .

The influence of  $N_o$  on modelling the refractive index appears in Fig. 2, which displays sample plots of  $CR_{355\text{nm}}$  and  $CR_{1064\text{nm}}$ . The standard deviation is fixed in Fig. 2a ( $CR = f_i(r_m)_{\sigma=1.45}$ ) while the mean radius is fixed in Fig. 2b ( $CR = f_i(\sigma)_{r_m=0.29\mu\text{m}}$ ). Two sets of curves are plotted. They correspond to  $N_o$  being equal to either 0.1 or  $10 \text{ cm}^{-3}$ . The reference values of  $r_m$  and  $\sigma$  come from the validation case described in Sect. 4.3. For high aerosol volumes (high values of  $\sigma$  or  $r_m$ ) the curves of  $CR(r_m, \sigma)$  for  $N_o = 0.1$  and for  $N_o = 10 \text{ cm}^{-3}$  start to differ, illustrating the influence of  $N_o$  on  $CR$ . The non-linearity of the retrieval problem is also highlighted in this figure, in that the colour ratios are not in univalent relationship with the size distribution parameters in Fig. 2.

For the environmental conditions (temperature, pressure, mixing ratios of total  $\text{HNO}_3$  and  $\text{H}_2\text{O}$ ) of each lidar data point, a look-up table of model-calculated  $BC$  and  $CR$  is generated as a function of  $N_o$ ,  $r_m$  and  $\sigma$ . The resolution step is typically  $0.1 \text{ cm}^{-3}$  for  $N_o$ ,  $0.01 \mu\text{m}$  for  $r_m$  and  $0.01$  for  $\sigma$ . The influence of the look-up table resolution on the retrieval is checked and increased till such influence is not noticeable anymore.

The size distribution retrieval algorithm then searches the look-up table for the model-simulated  $BC$  ( $\beta_{i,\text{model}}$  with  $i=355, 532, 1064 \text{ nm}$ ) and  $CR$  ( $CR_{j,\text{model}}$  with  $j=355/532, 1064/532$ ) that correspond to the measurements ( $\beta_{i,\text{lidar}}$ ,  $CR_{j,\text{lidar}}$ ). As  $CR$  also depend on  $N_o$  in the look-up table, the search for the optimal values of  $N_o$ ,  $\sigma$  and  $r_m$  is done in one step by fitting the 5 optical quantities ( $= 3 BC + 2 CR$ ). If the refractive index had been assumed constant,  $CR$  would have been independent of  $N_o$ . We could have first searched for  $\sigma$  and  $r_m$  by fitting the 2  $CR$  and then derived  $N_o$  by fitting the 3  $BC$  (as in Blum et al., 2006 and Baumgarten et al., 2007). This two-step procedure would have been less demanding in terms of computing time than the simultaneous search of  $N_o$ ,  $r_m$  and  $\sigma$  adopted here.

In order to estimate the errors in the size distribution retrieval, the lidar measurement errors ( $\Delta\beta_{i,\text{lidar}}$ ,  $\Delta CR_{j,\text{lidar}}$ ) are taken into account by searching the  $\beta_i$  and  $CR_j$  that are within the measurement intervals ( $\beta_{i,\text{lidar}} - \Delta\beta_{i,\text{lidar}}$ ,  $\beta_{i,\text{lidar}} + \Delta\beta_{i,\text{lidar}}$ ) and ( $CR_{j,\text{lidar}} - \Delta CR_{j,\text{lidar}}$ ,  $CR_{j,\text{lidar}} + \Delta CR_{j,\text{lidar}}$ ) respectively. Any ( $N_o$ ,  $r_m$ ,  $\sigma$ ) combination of the look-up table whose associated  $\beta_{i,\text{model}}$  and  $CR_{j,\text{model}}$  belong to the measurement intervals is taken as a possible solution. Obviously, the larger the lidar errors, the wider the pool of possible solutions. Note that, in our algorithm, the model errors are ignored. For instance, the size distribution retrieval algorithm does not account for errors in modelling the particle refractive

index. In other words, the model is assumed to be perfect. The additional size distribution retrieval errors originating from possible model deficiencies are discussed in the results and conclusion sections.

Several approaches are possible for determining the model solution ( $N_o$ ,  $r_m$ ,  $\sigma$ ). It is possible to simply look for the best (according to the least-square criterion) match between model simulations and measurements by minimizing the following scalar function:

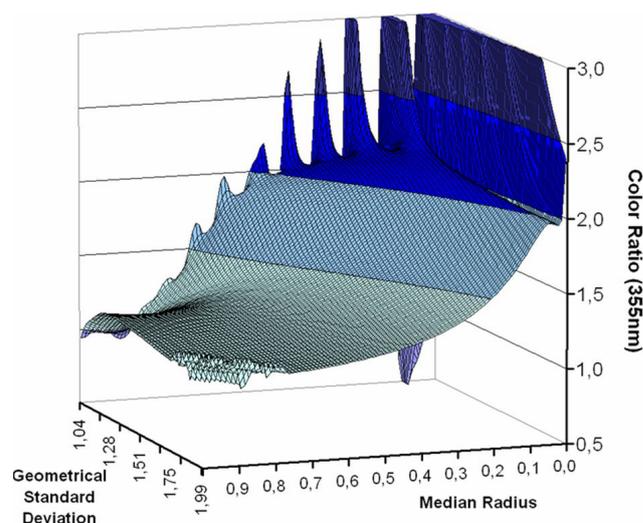
$$J(N_o, r, \sigma) = \sum_{i=1}^3 \frac{(\beta_{i, \text{model}} - \beta_{i, \text{lidar}})^2}{\Delta \beta_{i, \text{lidar}}^2} + \sum_{j=1}^2 \frac{(CR_{j, \text{model}} - CR_{j, \text{lidar}})^2}{\Delta CR_{j, \text{lidar}}^2}. \quad (4)$$

with  $\beta_i$  or  $CR_{j, \text{model}} = \mathbf{M}(N_o, \sigma, r_m, \lambda)$  where  $\mathbf{M}$  is the model operator.

This type of quadratic function that quantifies the misfit between model and data is variously referred as cost function, distance function, objective function, or penalty function in data assimilation (Elbern and Schmidt, 2001). This simple approach does not even require identifying the pool of possible model solutions, i.e. the combinations ( $N_o, \sigma, r_m$ ) such as  $X_{k, \text{lidar}} - \Delta X_{k, \text{lidar}} < (N_o, \sigma, r_m, \lambda) < X_{k, \text{lidar}} + \Delta X_{k, \text{lidar}}$  (with  $X=BC$  or  $CR$ , and  $k=1$  to 5). The best match model solution is indeed identified independently from the statistical properties of the cluster of possible solutions. This approach gives the most likely solution when the model is quasi-linear and the errors are Gaussian.

A second approach is to look for the most likely solution using the statistical properties of the cluster of possible solutions, given the lidar uncertainties. According to estimation theory, the most likely solution is expected to be found in the densest part of the solution cluster, around the maximum of the probability density function (i.e. maximum likelihood estimation).

The best match solution (i.e. solution corresponding to the minimization of the model-data misfit) and the most probable solution should be very close if the model  $\mathbf{M}(N_o, \sigma, r_m)$  was linear and the errors were Gaussian. However, our model is highly non-linear. To illustrate the non-linearity of the model, Fig. 3 presents typical examples of  $CR_{355 \text{ nm}}$  evolution as a function of  $r_m$  and  $\sigma$  for  $N_o = 0.01 \text{ cm}^{-3}$  and the environmental conditions corresponding to the first validation case (see Sect. 4.1). Depending on the value of the  $CR_{355 \text{ nm}}$  measurement, there are several possible solutions that can be quite scattered all over the  $r_m$  and  $\sigma$  domain. Figure 3 clearly indicates that the solutions to a given value of  $CR_{355 \text{ nm}}$  do not necessarily form a tight cluster in the size distribution parameter space. This is due to the non-linearity of the model and hence, to the high sensitivity of the calculated  $BC$  to the input size distribution parameters. In the same way, the other model-simulated optical properties,  $BC$  or  $CR_{1064 \text{ nm}}$ , can also have highly non-linear dependencies on the size distribution parameters. As shown by Eq. (1) and Eq. (2), the only possible linear relationship at first sight is the dependency of  $BC$  on  $N_o$ .  $BC$  has a somewhat exponential dependency on  $r_m$  and  $\sigma$  but this dependency varies



**Fig. 3.** Colour ratio at 355 nm plotted versus median radius  $r_m$  and standard deviation  $\sigma$ , calculated from Mie theory.  $r_m$  ranges from 0.01 to  $1 \mu\text{m}$  and  $\sigma$  from 1.01 to 2.  $N_o$  is set to  $10^{-2} \text{ cm}^{-3}$ . Refractive index calculations are made at  $T=186 \text{ K}$ , 4.5 ppmv  $\text{H}_2\text{O}$  and 10 ppbv  $\text{HNO}_3$ . Altitude is 22.5 km.

in a complex way in the size distribution parameter space. When  $\sigma$  gets close to 1, the particles tend to have their radius tightly scattered around the median radius. This narrow size distribution enhances the behaviour of the Mie differential backscattering cross section as a function of  $r$ , and, when  $\sigma$  is close to 1, its oscillations drive the backscatter as can be seen on Fig. 3 (Bohren and Huffman, 1983). The final cluster of possible model solutions is the intersection of 5 different solution clusters corresponding to the 5 optical quantities ( $=3 BC + 2 CR$ ). The resulting 3-D surface of possible model solutions can be very convoluted.

Multiple sensitivity tests have shown that, in several cases, the sole best match approach leads to somewhat unrealistic values of  $N_o$ ,  $r_m$  and  $\sigma$  with, in particular,  $\sigma$  being close to 1. Indeed, the surface of  $J=f(N_o, r_m, \sigma)$  sometimes exhibits a deep and extremely localised minimum, but the realistic solutions are mostly found in a very broad but shallower minimum. When looking at the cluster of possible model solutions in these cases, the best match solution (i.e. minimum of  $J=f(N_o, r_m, \sigma)$ ) is found on the edge or even completely disconnected from the cloud of possible solutions in the size distribution parameters space. This is confirmed by other numerical experiments using model-simulated  $BC$  as measurements. In this setup, the synthetic measurements are produced by adding random errors to the model-calculated  $BC$ . The true solution is the set of size distribution parameters taken as input to the model. In several cases, depending on the amplitude of the added random errors, the best match solution can be localised in a deep minimum area far away from the true solution. Therefore, on its own, the best match criterion guarantees neither unicity nor optimality of the solution because of the high non-linearity of the model.

In order to look for a model solution in the densest part of the possible solution cluster, the cluster is filtered based on its spread. The standard deviations  $\sigma_{N_o}$ ,  $\sigma_{r_m}$ , and  $\sigma_{\sigma}$  on the 3-D solution cluster are first calculated. Then, any possible solutions outside the  $\pm\sigma=(\sigma_{N_o}, \sigma_{r_m}, \sigma_{\sigma})$  envelope around the cluster centre are discarded. The coordinates of the cluster centre are taken as those of the median values of  $N_o$ ,  $r_m$  and  $\sigma$  because the median values are less sensitive to anomalous solution points than the mean values. Overall, this data filtering ensures removal of most unrealistic possible solutions. As expected, test simulations indicate that stable model solutions are associated with a large number of possible solutions within the clusters. Therefore, we also check the number of possible solutions left in the filtered cluster and proceed to the next step only when there is a significant fraction of the look-up table present in the solution cluster (typically one hundred points). This issue is of course linked to the amplitude of lidar  $BC$  errors and to the resolution of the look-up table.

The model solution is finally searched within the filtered cluster of possible solutions. The 3-D size of the initial solution cluster and, hence, the filtered cluster are determined by the amplitude of the lidar errors that are not known accurately. We prefer to use the best match approach (i.e. minimum of the model-measurement, as quantified by the scalar function  $J$ ) to determine the final model solution instead of an average of the solution cluster, because the best match solution is not sensitive to the size of the solution cluster (except when the errors are assumed to be excessively small and so the realistic model solution may not be found within the searched domain in some particular cases), whereas the mean or median values of the cluster can fluctuate with the cluster size. Overall, the uncertainties on the lidar errors are not critical for the best match model solution but they do influence the estimated errors in the retrieved size distribution parameters. Indeed, the standard deviations on the filtered cluster are taken as estimates of the size distribution retrieval errors. These estimates are possibly lower limits on the retrieval errors because the size of the cluster is reduced by the filtering and model errors are ignored. Note that the points within the cluster are not exactly normally distributed. Therefore, we could have estimated the retrieval error with percentiles that do not require any assumptions on the shape of the solution PDF. However, simple tests have shown that the use of the standard deviation or the 75 percentile does not have a very significant influence on the final estimation.

If the inverse problem was properly posed, the best match solution and the cluster centre should be very close. In order to be consistent with estimation theory, a final and slight adjustment is carried out using the lidar error as an adjustable parameter. Indeed, lidar errors are notoriously difficult to estimate (ASAP, 2006). The estimated errors in the measured  $BC$  are uncertain by at least 20%. Therefore, the distance between the best match solution and the cluster centre (defined by the median values) in the size distribution parameter space

is minimized by varying the  $BC$  errors  $\Delta\beta_{\text{lidar}}$  within their uncertainties. The distance is expressed as a scalar function:

$$D(\Delta\beta_{\text{lidar}}) = \frac{(N_{o,\text{bestmatch}} - N_{o,\text{center}})^2}{\sigma_{N_o,\text{cluster}}^2} + \frac{(r_{m,\text{bestmatch}} - r_{m,\text{center}})^2}{\sigma_{r_m,\text{cluster}}^2} + \frac{(\sigma_{\text{bestmatch}} - \sigma_{\text{center}})^2}{\sigma_{\sigma,\text{cluster}}^2}. \quad (5)$$

where  $X_{\text{best-match}}$  and  $X_{\text{center}}$  are the size distribution parameter  $X$  of the best match solution and of the cluster centre (defined as the median value) respectively;  $\sigma_{X,\text{cluster}}$  is the standard deviation of the size distribution parameter  $X$  on the solution cluster.

The uncertainties on the  $BC$  errors are assumed to be 20% typically. Therefore, if the  $BC$  error is 12.5% for example, the distance  $D$  is minimized by varying the  $BC$  error between 10% and 15%. The main objective of the whole procedure is to ensure that the retrieval algorithm produces realistic and stable model solutions (i.e. size distribution parameters) that are weakly sensitive to the estimated  $BC$  errors as well as being consistent with the estimation theory.

## 4 Evaluation against size-resolved measurements

The first evaluation of the retrieval algorithm is based on more or less coincident lidar and balloonborne size-resolved measurements of a stratospheric thick cloud observed above ALOMAR and described in Deshler et al. (2000). The lidar measurements provide the input data for the retrieval algorithm and the size-resolved measurements provide the reference values for an independent evaluation of the retrieved size distribution.

### 4.1 PSC Observations

On 23 January 1996 a thick cloud was detected by lidar above ALOMAR. This cloud was present between 19 and 26 km. The lidar signal was depolarized between 23 and 25 km. The cloud was also probed at about the same time with an Optical Particle Counter (OPC) instrument that provided size distribution measurements. European Centre for Meteorological Weather Forecasts (ECMWF) analyses indicated favourable conditions for PSC formation. A detailed description of the temporal evolution of the lidar measurements and of the cloud structure can be found in Hansen and Hoppe (1997). Measurements were carried out at several altitudes within the broad cloud layer, both in the depolarizing and non-depolarizing regions of the cloud. These sets of lidar and size-resolving measurements appear to be excellent opportunities to validate our size distribution retrieval methodology. Around 22.5 km, in a non-depolarizing (spherical particles) region of the cloud, Deshler et al. (2000) identified a type Ib PSC whose measured size distribution could be fitted with a unimodal lognormal distribution. This type Ib PSC represents an ideal case of validation because it fulfils the conditions of sphericity for the particles and unimodality for the size distribution. The size-resolved measurements of the other PSC layers indicated bimodal distributions with some

being non-depolarizing. In the depolarizing (non-spherical particles) layers, type Ia or even type II PSC were identified. To check the influence of the sphericity and unimodality conditions on the results, we also perform size distribution retrievals on these cases where the retrieval system is a priori expected to perform poorly. One of the aims is to identify the size distribution parameter whose retrieval is most affected by the non-respect of these conditions. All the log-normal parameters of the measured (reference) and retrieved size distributions are gathered in Table 1.

#### 4.2 Lidar data

Lidar signals are averaged between 17:50 and 18:30 UTC in order to be coincident with the in situ balloon-borne size-resolving measurements, as described in Deshler et al. (2000). It is worth pointing out that, even if the lidar and OPC measurements were truly coincident, a *BC* derived from lidar data averaged over 40 min is not expected to be equivalent to a *BC* calculated from the corresponding particle size distribution averaged over the same time interval unless the particle distribution does not vary temporally. The *BC* at 355 and 532 nm are retrieved using the vibrational Raman  $N_2$  scattering detected at 387 and 607 nm respectively. As the association of elastic and inelastic scattering allows the determination of the particle extinction, there is no assumption on the value of the lidar ratio (extinction over backscatter) in the inversion of the equation linking the *BC* and the lidar signal. The lidar inversion still requires the choice of a reference altitude at which the *BC* is supposed to be known (clear-air assumption) (Ansmann et al., 1990, 1992a, b). Figure 4 shows the backscatter ratios profiles at 355, 532 and 1064 nm, defined as the total backscatter divided by the Rayleigh (molecular) contribution. These profiles can be compared to those obtained in Deshler et al. (2000) who used the Klett method for the three wavelengths assuming a linear relation between extinction and backscatter (Klett, 1981, 1985). There is still an excellent agreement between the two inversion techniques. We estimate the errors in our determination of the backscatter coefficient to be about 10% at 355 and 532 nm, and 20% at 1064 nm.

#### 4.3 Non-depolarizing unimodal size distribution

The unimodal type Ib PSC was observed at about 22.5 km. The temperature was about 186 K according to the measurements. This temperature is used here as input to the microphysical model. The volume mixing ratios in total  $HNO_3$  and  $H_2O$  required by the model are set to the typical values of 10 ppbv and 4.5 ppmv respectively. The ranges of size distribution parameters considered by the retrieval algorithm are as follows:  $10^{-3} < N_o < 20 \text{ cm}^{-3}$ ,  $10^{-3} < r_m < 3 \mu\text{m}$  and  $1.01 < \sigma < 2$ . After applying the retrieval algorithm to the *BC* and *CR* data, we obtain  $N_o = 11.4 \text{ cm}^{-3}$  (9%),  $r_m = 0.30 \mu\text{m}$  (5%) and  $\sigma = 1.44$  (2%) with the numbers in

**Table 1.** Reference and retrieved size distribution parameters on the PSC case observed on 23 January 1996. On the left part, OPC size distribution reference values in Deshler et al. (2000). On the right part, retrieved values in case of liquid (a and b) and solid particles (c). The associated retrieval errors in the estimated parameters are bracketed.

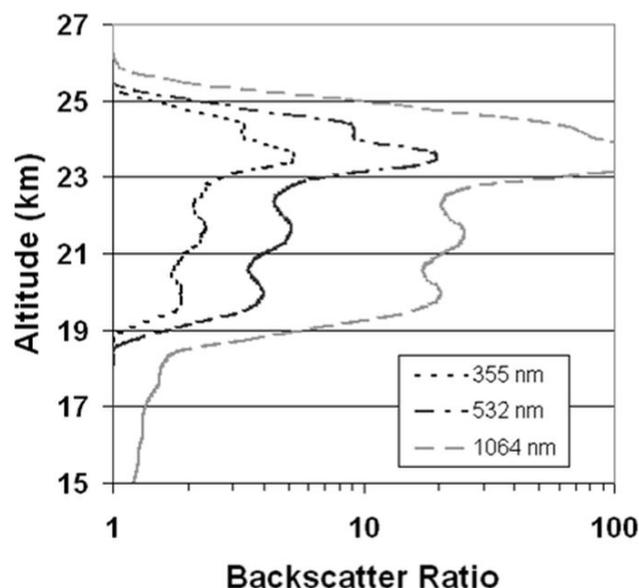
	Reference values		Retrieved values	
	Mode 1	Mode 2	m imposed	
Altitude: 22.45 km	Non-depolarizing region (a)			
$N_o$ ( $\text{cm}^{-3}$ ):	7.71	-	11.4 (9%)	6.67 (22%)
$r_m$ ( $\mu\text{m}$ ):	0.29	-	0.30 (5%)	0.32 (6%)
$\sigma$ :	1.45	-	1.44 (2%)	1.38 (2%)
$A$ ( $\mu\text{m}^2 \cdot \text{cm}^{-3}$ ):	10.4	-	16.8 (20%)	10.6 (40%)
$V$ ( $\mu\text{m}^3 \cdot \text{cm}^{-3}$ ):	1.4	-	2.4 (30%)	1.5 (50%)
Altitude: 21.83 km	Non-depolarizing region (b)			
$N_o$ ( $\text{cm}^{-3}$ ):	5.63	0.005	7.3 (11%)	
$r_m$ ( $\mu\text{m}$ ):	0.4	1.74	0.41 (6%)	
$\sigma$ :	1.26	1.33	1.27 (6%)	
$A$ ( $\mu\text{m}^2 \cdot \text{cm}^{-3}$ ):	12.63	0.22	17.1 (30%)	
$V$ ( $\mu\text{m}^3 \cdot \text{cm}^{-3}$ ):	1.92	0.16	2.7 (40%)	
Altitude: 23.98 km	Depolarizing region (c)			
$N_o$ ( $\text{cm}^{-3}$ ):	8.23	0.005	9.0 (17%)	
$r_m$ ( $\mu\text{m}$ ):	0.26	1.68	0.62 (7%)	
$\sigma$ :	1.39	1.34	1.29 (8%)	
$A$ ( $\mu\text{m}^2 \cdot \text{cm}^{-3}$ ):	8.71	0.21	49.4 (50%)	
$V$ ( $\mu\text{m}^3 \cdot \text{cm}^{-3}$ ):	0.99	0.15	12 (70%)	

brackets being the estimated retrieval errors. The retrieved values can be compared with the OPC measurements:  $N_o = 7.71 \text{ cm}^{-3}$ ,  $r_m = 0.29 \mu\text{m}$  and  $\sigma = 1.45$ . The  $r_m$  and  $\sigma$  values agree extremely well, around a few percent, which is within the error bars. In contrast, the  $N_o$  values differ by about 50% which is a factor 5 higher than the estimated retrieval error in  $N_o$  (9%). As stated before, *BC* and *CR* are generally less sensitive to  $N_o$  than  $r_m$  and  $\sigma$ . Therefore, the retrieval of  $N_o$  is expected to be less accurate.

From the values of  $N_o$ ,  $r_m$  and  $\sigma$ , one can calculate higher order moments of the size distribution such as the total surface area density  $A$  and volume density  $V$  that are key quantities for heterogeneous chemistry and, hence, ozone loss:

$$A = N_o \cdot 4\pi \cdot r_m^2 \cdot e^{[2 \cdot \ln^2 \sigma]} \text{ and } V = N_o \cdot \frac{4\pi}{3} \cdot r_m^3 \cdot e^{[1.5 \cdot \ln^2 \sigma]}. \quad (6)$$

For our retrieved size distribution parameters,  $A = 16.8 \mu\text{m}^2 \cdot \text{cm}^{-3}$  (20%) and  $V = 2.4 \mu\text{m}^3 \cdot \text{cm}^{-3}$  (30%). For the OPC reference measurements,  $A = 10.4 \mu\text{m}^2 \cdot \text{cm}^{-3}$  and  $V = 1.4 \mu\text{m}^3 \cdot \text{cm}^{-3}$  with uncertainties of around 30%, with 20% originating from the OPC measurement errors and 10% from errors in the refractive index (Deshler et al., 2000). Our  $A$  and  $V$  values differ by about 60% from the reference values, which is almost greater than the total (= size distribution retrieval + OPC) errors. Although  $A$  and  $V$  are most sensitive to  $r_m$  (power law dependency), the 60% bias on  $A$  and  $V$  mostly originates from the bias in  $N_o$ . Indeed, the respective 3% and 1% biases on  $r_m$  and  $\sigma$  only lead to a 5% error in  $A$  and  $V$ . It is not totally surprising that



**Fig. 4.** Backscatter ratio profiles associated with lidar measurements performed at ALOMAR on 23 January 1996. Lidar signals are averaged between 17:50 and 18:30 UTC.

$BC$  are less sensitive to  $N_o$  and hence, that  $N_o$  is the most difficult parameter to retrieve. The Mie scattering kernels in the spectral range of interest (from 0.385 to 1.02  $\mu\text{m}$ ) show that lidar data are most sensitive to particles in the 0.1–1  $\mu\text{m}$  range (Fig. 4.28 in ASAP, 2006). As a result, the moments that depend on smaller particles (such as the  $N_o$  concentration) are less likely to be accurately retrieved.

To check the stability of the solution, the retrieval is performed with  $BC$  errors ranging from 5% to 50%. The solution does not change from around 10% to 50% lidar errors. When lidar errors are lower than 10%, another solution ( $N_o=11\text{ cm}^{-3}$ ,  $r_m=0.32\text{ }\mu\text{m}$  and  $\sigma=1.39$ ) is identified. However, this solution has to be discarded because of the very low number of points present in the solution cluster. Besides, lidar errors are unlikely to be lower than 10%. Another test is carried out with a look-up table that does not cover the range of the reference OPC values. As expected, no solutions are identified. Overall, it can be concluded that the algorithm appears to perform well, especially when retrieving the size distribution parameters  $r_m$  and  $\sigma$  of unimodal liquid PSC distributions. The overall errors on the integrated quantities  $A$  and  $V$  as well as the  $N_o$  parameter appear to be higher than errors derived theoretically for methodologies using regularization techniques (Müller et al., 1999; Böckmann and Kirsche, 2006). Such studies use extinction coefficients on the top of backscattering coefficients at tropospheric altitudes. In our case, the extinction coefficient is not retrieved with sufficient accuracy at stratospheric altitudes to be used as an additional constraint in the retrieval procedure, thus leading to higher errors on the final estima-

tion. The total number of particles  $N_o$  still appears to be the most difficult parameter to retrieve accurately (Müller et al., 1999; Böckmann and Kirsche, 2006).

#### 4.4 Assumptions of unimodality and sphericity

##### 4.4.1 Bimodal size distribution

A non-depolarizing PSC layer with a bimodal size distribution was detected around 21.8 km (Deshler et al., 2000). The values of the size distribution parameters of the two modes are given in Table 1b. Using them as inputs to the optical model, the simulated  $BC$  of the small particle mode is found to be more than a hundred times higher than the  $BC$  of the large particle mode at all the considered wavelengths. The retrieval algorithm is applied at the 21.8 km altitude for the corresponding pressure and temperature and for the same conditions as above (total  $\text{HNO}_3$  and  $\text{H}_2\text{O}$ , considered ranges of size distribution parameters and lidar errors). The retrieved values are  $N_o=7.3\text{ cm}^{-3}$  (11%),  $r_m=0.41\text{ }\mu\text{m}$  (6%) and  $\sigma=1.27$  (6%), with  $A=17.1\text{ }\mu\text{m}^2\cdot\text{cm}^{-3}$  (30%) and  $V=2.7\text{ }\mu\text{m}^3\cdot\text{cm}^{-3}$  (40%). As expected, the retrieved parameters are found to reflect the optically dominant mode of the small particles. They can be compared to the OPC-derived reference parameters of the small particles mode ( $N_o=5.63\text{ cm}^{-3}$ ,  $r_m=0.4\text{ }\mu\text{m}$  and  $\sigma=1.26$  with  $A=12.6\text{ }\mu\text{m}^2\cdot\text{cm}^{-3}$  and  $V=1.9\text{ }\mu\text{m}^3\cdot\text{cm}^{-3}$ ). As for the previous case featuring an unimodal distribution, the agreement with the reference values is excellent for  $r_m$  and  $\sigma$  and the largest discrepancy is found for  $N_o$ . The bias in  $N_o$  results in retrieved values of  $A$  and  $V$  that differ from the reference values by about 40%. Again, the retrieval is found to be stable regarding the specified lidar errors. The results show that the algorithm can be applied to multimodal size distributions, but that the retrieved size distribution parameters are only relevant to the optically dominant mode.

##### 4.4.2 Spherical particles

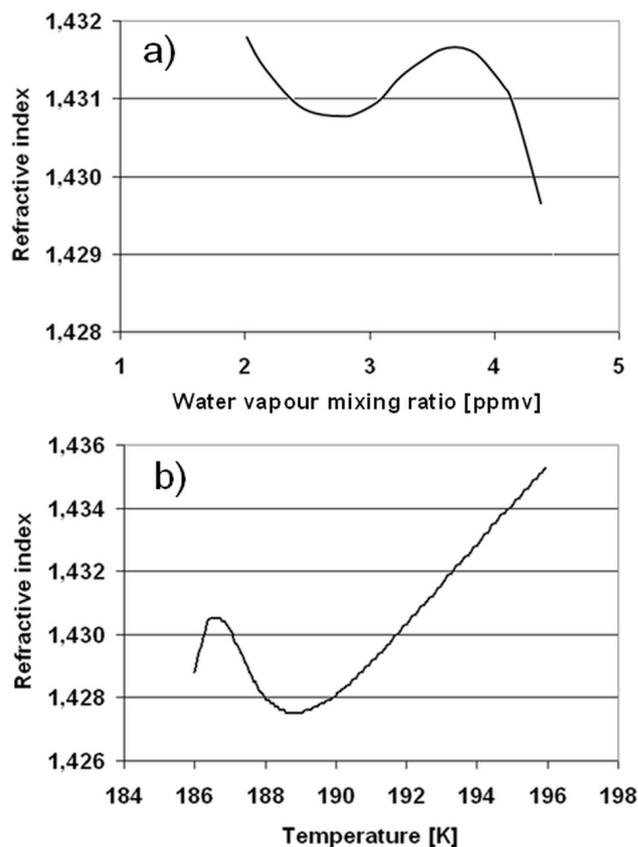
A depolarizing PSC layer was observed around 24 km. Its size distribution was bimodal. The exact type of this solid PSC, or at least, PSC containing a mode of solid particles, has not been clearly defined. Based on morphology and colour index, Hansen and Hoppe (1997) concluded that it was a type II (ice) PSC whereas Deshler et al. (2000) analysed the measured size distribution and the associated particle volume to arrive at a type Ia (Nitric Acid Trihydrate) PSC. As for the case of the liquid bimodal distribution previously considered, optical model calculations indicate that the optical signal of the small particle mode is largely dominant over the signal of the large particle mode. As solid PSC are composed of non-spherical particles, this case is used to test the influence of the presence of non-spherical particles on the quality of the retrieval. The retrieved values are  $N_o=9.0\text{ cm}^{-3}$ ,  $r_m=0.62\text{ }\mu\text{m}$  and  $\sigma=1.29$  with

$A \approx 49 \mu\text{m}^2 \cdot \text{cm}^{-3}$  and  $V \approx 12 \mu\text{m}^3 \cdot \text{cm}^{-3}$ . The OPC-derived parameters for the small particle mode were  $N_o = 8.2 \text{ cm}^{-3}$ ,  $r_m = 0.26 \mu\text{m}$  and  $\sigma = 1.39$  with  $A \approx 9.0 \mu\text{m}^2 \cdot \text{cm}^{-3}$  and  $V \approx 1.0 \mu\text{m}^3 \cdot \text{cm}^{-3}$ . In contrast to the previous results, the agreement is very poor for  $r_m$  and  $\sigma$  but it is satisfactory for  $N_o$ . The very large discrepancies on  $r_m$  and  $\sigma$  lead to retrieved values of  $A$  and  $V$  differing from the OPC-derived values by a factor 5–10. Note that the refractive index value was calculated for STS particles, whose value is close to the NAT value. We also performed a retrieval forcing the refractive index to be equal to the PSC type II refractive index value as inferred in Scarchilli et al. (2005). But the overall results do not change very significantly. The size distribution was found to be  $N_o = 7 \text{ cm}^{-3}$ ,  $r_m = 0.53 \mu\text{m}$  and  $\sigma = 1.75$  with  $A \approx 46 \mu\text{m}^2 \cdot \text{cm}^{-3}$  and  $V \approx 18 \mu\text{m}^3 \cdot \text{cm}^{-3}$ . Accounting for a more accurate refractive index does not improve the retrieval in this case.

As expected, using Mie theory for non-spherical particles results in erroneous retrieved parameters. For the retrieval of solid PSC size distributions, it is necessary to replace the Mie optical model by a T-matrix optical model in the algorithm (Mishchenko and Travis, 1998), possibly making the particle asphericity (for example the aspect ratio) one of the control variables along with  $N_o$ ,  $r_m$  and  $\sigma$ .

#### 4.5 Refractive index

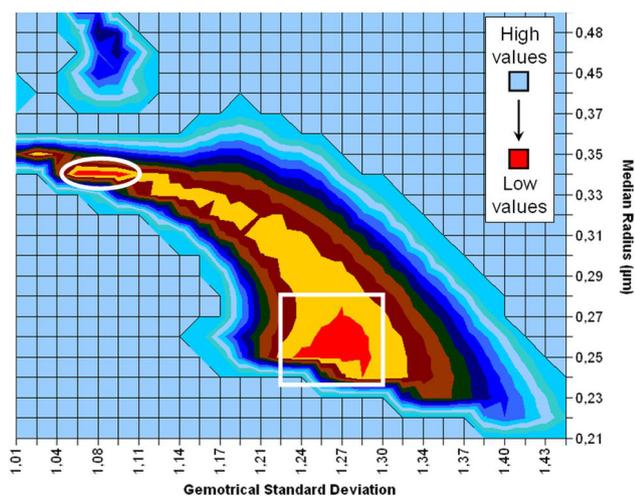
An important issue in the retrieval is connected to the modelling of the refractive index  $m$ . The backscatter coefficient is actually most sensitive to  $m$ , but its possible range of values is relatively limited (1.3–1.6) (Deshler et al., 2000; Luo et al., 1996; Scarchilli et al., 2005). For the case of the non-depolarizing unimodal distribution, the STS equilibrium composition was calculated for an assumed 4.5 ppmv H<sub>2</sub>O and 10 ppbv HNO<sub>3</sub> giving a refractive index of 1.45 at 352 nm and reducing to 1.43 at 1064 nm (Luo et al., 1996). Deshler et al. (2000) inferred an  $m$  value of  $1.47 \pm 0.03$  for the same PSC from a comparison between scattering and size-resolved aerosol concentration measurements. They also concluded that this value was higher than expected on the basis of laboratory experiments, measurements in Antarctica and theory. In order to estimate the impact of this  $m$  uncertainty on the results, another retrieval is performed on the same case, but with the refractive index calculation replaced by setting  $m$  to 1.48, 1.46 and 1.51 at 355, 532 and 1064 nm wavelengths respectively, as provided in Deshler et al. (2000). The retrieved parameters are  $N_o = 6.7 \text{ cm}^{-3}$  (22%),  $r_m = 0.32 \mu\text{m}$  (6%) and  $\sigma = 1.38$  (2%) with  $A = 10.6 \mu\text{m}^2 \cdot \text{cm}^{-3}$  (40%) and  $V = 1.5 \mu\text{m}^3 \cdot \text{cm}^{-3}$  (50%). Let us recall that the OPC reference values were  $N_o = 7.71 \text{ cm}^{-3}$ ,  $r_m = 0.29 \mu\text{m}$  and  $\sigma = 1.45$  with  $A = 10.4 \mu\text{m}^2 \cdot \text{cm}^{-3}$  and  $V = 1.4 \mu\text{m}^3 \cdot \text{cm}^{-3}$ . The quality of the retrieval is slightly degraded for  $r_m$  and  $\sigma$ , but it improves drastically for  $N_o$ . The improvement is even more noticeable for  $A$  and  $V$ ; the retrieved values are now within 1% and 7% respectively of the reference values instead of



**Fig. 5.** Plots of the refractive index versus water vapour mixing ratio (a) and temperature (b). Wavelength is 532 nm, size distribution is as in Table 1a. Water vapour is 4.5 ppmv in plot (a) and temperature is 186 K in plot (b).

the 60% difference initially obtained when the refractive index was calculated by the model (Sect. 4.3). Although the change in  $m$  seems to be very marginal (about 0.03 or 2%), it has a profound effect on the retrieval results, confirming by the way that the backscatter coefficient is highly sensitive to  $m$ . The results obtained by bypassing the modelling of the composition and refractive index suggest that most of the bias in the initial retrieval was due to errors in  $m$  and not to problems in the Monte Carlo approach. This further validates our retrieval methodology.

If the refractive index inferred by Deshler et al. (2000) is taken as a reference value, the bias in  $m$  could originate from either errors in the relationship linking  $m$  to the particle composition (Luo et al., 1996) or uncertainties in modelling the composition. The input variables are the temperature, pressure and total HNO<sub>3</sub> and H<sub>2</sub>O mixing ratios. The temperature is taken from in situ measurements and therefore should not carry a large uncertainty. The volume mixing ratio of total HNO<sub>3</sub> and H<sub>2</sub>O were set to typical values (10 ppbv HNO<sub>3</sub> and 4.5 ppmv H<sub>2</sub>O). If ice clouds had occurred prior to the lidar measurements, the water vapour content could have decreased significantly. Measurements at Sodankylä on



**Fig. 6.** Cost function in the  $(r_m, \sigma)$  space for the size distribution retrieved in Sect. 5. Associated values of  $N_o$  are not displayed. The purely best match solution is found within the encircled area. Taking lidar errors into account, the densest part of the solution cluster is found to be within the squared area.

23 January at 00:00 UTC indicated that  $\text{H}_2\text{O}$  dropped under 3 ppmv between 22 and 24 km (Vömel et al., 1997). Therefore, lower  $\text{H}_2\text{O}$  could have been considered. An additional retrieval is performed with 3 ppmv of total  $\text{H}_2\text{O}$ . The results are very close to those presented in Sect. 4.3. The effect of  $\text{H}_2\text{O}$  and temperature uncertainties on the model-calculated refractive index is further illustrated in Fig. 5. Size distribution parameters and altitude are set to the reference values of the liquid unimodal distribution given in Table 1. The curves show that the uncertainties on the temperature and  $\text{H}_2\text{O}$  can account for, at best, a  $5 \cdot 10^{-3}$  change in  $m$  which is largely insufficient to explain the 0.03 discrepancy with the refractive index derived by Deshler et al. (2000). It is difficult for us to pinpoint the exact reasons for the discrepancy on this particular case because of several possible sources of biases, ranging from the lack of spatio-temporal coincidence between the lidar and size-resolving measurements, errors in the composition model itself (for example, the assumption of thermodynamical equilibrium) or errors in the relationship between composition and  $m$  (Luo et al., 1996; Larsen et al., 2000). In addition, the refractive indices of PSC do not appear to be known with an accuracy better than 0.03 (Scarchilli et al., 2005).

## 5 Evaluation against a simplified estimation methodology

Our results are then compared to values retrieved with another approach that is similar but does not consider the effect of the non-linearity of the model on the likelihood of the solution (Blum et al., 2006). The test case is a liquid PSC observed over northern Scandinavia on 5 January 2005.

Blum et al. (2006) and first searched for  $\sigma$  and  $r_m$  by fitting the 2 *CR* and then derived  $N_o$  by fitting the 3 *BC*. They did not consider the errors in the lidar measurements and retrieved the parameters by simply looking for the minimum of the misfit between model simulations and measurements (i.e. best match solution). The size distributions were found to be very narrow throughout the 5 km thick cloud. Retrieved values of  $r_m$  and  $\sigma$  were more or less constant at  $0.3 \mu\text{m}$  and 1.04 respectively whereas  $N_o$  varied between 2 and  $20 \text{ cm}^{-3}$ . We processed the lidar signal using vibrational Raman detection for determining the *BC* and rotational Raman detection at 529 and 530 nm to get the in situ temperature using the Rotational Raman Technique (Cooney and Pina, 1976; Nedeljkovic et al., 1993). The determination of a simultaneous and collocated temperature profile improved the accuracy of both the *BC* and the chemical composition modelling. The volume mixing ratios of total  $\text{HNO}_3$  and  $\text{H}_2\text{O}$  are set to the typical values of 10 ppbv and 4.5 ppmv respectively. Lidar backscatter errors are increased to 20% at 355 and 532 nm to account for differences in the inversion procedure with the *BC* retrieval of Blum et al. (2006). The full retrieval procedure gives  $N_o=16 \text{ cm}^{-3}$  (20%),  $r_m=0.26 \mu\text{m}$  (7%) and  $\sigma=1.27$  (5%) in the middle of the PSC layer, at 21 km. The difference with the results of Blum et al. for the  $\sigma$  value is well beyond any possible retrieval errors. When the measurement errors are ignored (i.e. skipping the cluster filtering and statistical error analysis), our retrieval becomes a search for the best match solution only. This simplified version of our retrieval algorithm is then applied to the same lidar data, but the cluster size is still used to estimate the retrieval errors. The retrieved parameters become  $N_o=9 \text{ cm}^{-3}$  (47%),  $r_m=0.34 \mu\text{m}$  (18%) and  $\sigma=1.04$  (12%) which is in excellent agreement with the results of Blum et al. (2006).

The results clearly illustrate the importance of the lidar errors and of the model non-linearity in the choice of the best estimator. The best match solution is found to be outside the  $\pm\sigma$  boundary of the initial solution cluster. It corresponds to a very localised and deep minimum in the surface of the model-data misfit function (see Fig. 6). But the highest density of possible solutions (when lidar errors are accounted for) is found in a very broad but shallower minimum in that case. The advantage of the filtering can be further justified by performing a retrieval in which the backscatter at 532 nm is substantially biased on purpose (i.e. increasing  $BC_{532 \text{ nm}}$  by 20%). To be consistent, the error in  $BC_{532 \text{ nm}}$  is also increased by 20%. The retrieval is performed with the biased  $BC_{532 \text{ nm}}$  and the associated biased  $CR_{355 \text{ nm}}$  and  $CR_{1064 \text{ nm}}$  while  $BC_{355 \text{ nm}}$  and  $BC_{1064 \text{ nm}}$  remain unchanged. The retrieved parameters for the purely best match solution (no cluster filtering) are  $N_o=2.9 \text{ cm}^{-3}$  (150%),  $r_m=0.51 \mu\text{m}$  (15%) and  $\sigma=1.04$  (10%) whereas the full retrieval (including cluster filtering) gives  $N_o=11.5 \text{ cm}^{-3}$  (20%),  $r_m=0.31 \mu\text{m}$  (7%) and  $\sigma=1.22$  (5%). The different results of this test are summarized in Table 2. The no-filtering retrieval gives  $N_o$  and  $r_m$  values for the artificially biased  $BC_{532 \text{ nm}}$

**Table 2.** Reference and retrieved size distribution parameters on the PSC case observed on 5 January 2005. Reference values come from Blum et al. (2006). The table gives the size distributions retrieved with (full algorithm) and without (best match algorithm) the cluster filtering step. Errors in the retrieved parameters are bracketed.

	$N_o$ (cm <sup>-3</sup> )	$r_m$ (μm)	$\sigma$
Reference size distribution			
5 January 2005 (Blum et al., 2006)	2–20	0.3	1.04
Size distribution retrieval			
Best match approach:	9 (47%)	0.34 (18%)	1.04 (12%)
Statistical filtering:	16 (20%)	0.26 (7%)	1.27 (5%)
Size distribution retrieval ( $BC_{532}$ biased)			
Best match approach:	3 (150%)	0.51 (15%)	1.04 (10%)
Statistical filtering:	11.5 (20%)	0.31 (7%)	1.22 (5%)

that are completely different, well beyond the retrieval error bars, from those produced with the unbiased  $BC_{532\text{nm}}$ . In contrast, taking into account the retrieval error bars, the results of the full retrieval with the artificially biased  $BC_{532\text{nm}}$  are fully consistent with the full retrieval with the unbiased  $BC_{532\text{nm}}$ . Our retrieval algorithm shows less sensitivity to the specified lidar errors and a much greater convergence and stability when the statistical filtering is applied.

## 6 Summary and conclusions

This paper introduces a particle size distribution retrieval method developed to take advantage of the numerous 3-wavelength lidar setups. It is based on comparing measured and model-simulated lidar backscatter coefficients. The retrieval algorithm minimizes a cost function of the misfit between measurements and model simulations with the control variables being the parameters of the PSC size distribution. In similar approaches, the estimator is often chosen as the sole “best match” estimator (i.e. best estimation corresponding to the best fit between modelled and measured backscatter coefficients). The errors in the lidar backscatter coefficients and associated colour ratios are rarely used as input parameters in the retrieval procedure. This may be critical in that the retrieval problem is highly non linear, and, on its own, the best match criterion guarantees neither unicity nor optimality of the solution. Our approach looks for the most likely estimation, using the statistical properties of the solution cluster given the lidar uncertainties. This solution is expected to be found in the densest part of the solution cluster, around the maximum of the probability density function. With our retrieval procedure, we aim at producing realistic and stable estimators of the size distribution parameters that are weakly sensitive to the specified backscatter coefficient errors that are difficult to estimate accurately.

We use a microphysical model to calculate the refractive index from the particle composition, taking into account sul-

phuric acid aerosol (binary  $\text{H}_2\text{SO}_4/\text{H}_2\text{O}$  solution) and liquid type Ib PSC particles (ternary STS solution). We then generate a look-up table of backscatter coefficients as a function of the particle size distribution parameters. Modelling the composition allows us to account for the rapid variation of the particle composition, and hence, the refractive index around the PSC type Ib temperature threshold. This refractive index modelling requires the local temperature to be known accurately because the temperature remains the key parameter regarding PSC formation.

To be able to resolve the inverse problem from measurements at only three wavelengths, the size distribution is assumed to be unimodal. The particles are also assumed to be spherical, due to the use of Mie theory in the optical modelling. The model errors are ignored. Our methodology was described and validations were performed against size-resolved OPC measurements and against a similar approach. The first validation case is a PSC event observed above ALOMAR on 23 January 1996. OPC size distribution measurements performed by Deshler et al. (2000) on this PSC provided the reference validation dataset. Due to photo-counting statistics, calculations from OPC measurements suffer from an error of about 20%.

In regions where the criteria of unimodality and sphericity are fulfilled, the parameters of size distribution of PSC type Ib particles are mostly correctly retrieved, with a respective 3% and 1% error in the median radius ( $r_m$ ) and standard deviation ( $\sigma$ ) (see Table 1a). The total number of particles  $N_o$  is found to be the most difficult parameter to retrieve accurately, with an error of around 50%. In addition, we ran our retrieval algorithm at other altitudes, where non-depolarizing (spherical particles) bimodal size distributions are observed, to check the influence of the assumption of unimodality. When retrieving the size distribution of this bimodal particle population, the retrieved parameters mainly correspond to the particle mode that dominates the lidar backscatter signal (see Table 1b). Overall, the results appear to be satisfying for spherical particles.

When considering the depolarizing region of the PSC, as expected, the retrieved parameters strongly differ from the OPC reference values. It is clear that the assumption of spherical particles appears to be critical (see Table 1c). Accounting for the particle asphericity in the optical model could allow our algorithm to be applied to solid particles (PSC type Ia or type II).

Model errors are not accounted for in our retrieval algorithm. However, the refractive index is a critical parameter in the calculation of the aerosol optical properties. The model-calculated refractive index of liquid particles in the case observed in 1996 was found to be 0.03 lower than the one inferred in Deshler et al. (2000) at the same altitudes. Using this inferred refractive index value in the optical model, the retrieved parameters are found to be in excellent agreement with the OPC measurements, with both the surface area density and aerosol volume carrying less than 10% error as compared to the reference values. This suggests that most of the errors in the retrieval originate from uncertainties on the refractive index rather than errors in the retrieval methodology. Note that our estimation of the retrieval errors does not take into account the errors involved in the optical modelling such as errors in the chemical composition required in refractive index calculations or errors in the scattering theory.

Our results are then compared to values retrieved with another approach that is similar but does not consider the effect of the non-linearity of the model on the likelihood of the solution: Blum et al. (2006) retrieved the size distribution a liquid PSC observed over northern Scandinavia on 5 January 2005 using comparison between measured and model-simulated colour ratios, the solution being the best match. We performed retrieval tests with and without the cluster filtering approach on the PSC case described in Blum et al. (2006). An excellent agreement is found between their values and values retrieved with our algorithm without any statistical filtering of the solution cluster. However, the value of the geometrical standard deviation appears to be unlikely, being very close to 1 ( $\sigma \approx 1.04$ ). When the full retrieval algorithm is applied (statistical filtering), our results are found to be significantly different with a more realistic value of the standard deviation ( $\sigma \approx 1.27$ ). Additional tests were performed with synthetic lidar measurements generated a priori with the optical model using specific values of size distribution parameters. These synthetic measurements are then used as inputs to the retrieval algorithm. Even with a 20% artificial bias on one of the backscatter coefficients, the full algorithm is able to retrieve the original size distribution parameters with a reasonable accuracy. Our full retrieval procedure does not appear to be very sensitive to the specified lidar errors. Overall, our statistical approach produces reliable estimates of liquid particle size distributions from only three lidar backscatter coefficients.

Edited by: G. Vaughan

## References

- Ansmann, A., Riebesell, M., and Weitkamp, C.: Measurement of atmospheric aerosol extinction profiles with a Raman lidar, *Opt. Lett.*, 15, 746–748, 1990.
- Ansmann, A., Riebesell, M., Wandinger, U., Weitkamp, C., Voss, E., Lahmann, W., and Michaelis, W.: Combined Raman elastic-backscatter lidar for vertical profiling of moisture, aerosol extinction, backscatter, and lidar ratio, *Appl. Phys. B*, 55, 18–28, 1992a.
- Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W.: Independent measurement of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-backscatter lidar, *Appl. Opt.*, 31, 7113–7131, 1992b.
- Assessment of Stratospheric Aerosol Properties (ASAP), WCRP-124, WMO/TD-No. 1295, SPARC Report No. 4, 2006.
- Baumgarten, G., Fiedler, J., and Von Cossart, G.: The size of noctilucent cloud particles above ALOMAR (69°N, 16°E): Optical modelling and method description, *J. Adv. Space Res.*, 40, 772–784, doi:10.1016/j.asr.2007.01.018, 2007.
- Böckmann, C. and Kirsche, A.: Iterative regularization method for lidar remote sensing, *Comp. Phys. Com.*, 174, 607–615, 2006.
- Beyer, D., Ravishankara, A. R., and Lovejoy, E. R.: Measurements of UV refractive indices and densities of H<sub>2</sub>SO<sub>4</sub>/H<sub>2</sub>O and H<sub>2</sub>SO<sub>4</sub>/HNO<sub>3</sub>/H<sub>2</sub>O solutions, *J. Geophys. Res.*, 101(D9), 14519–14524, 1996.
- Beyerle, A., Neuber, R., and Schrems, O.: Multiwavelengths lidar measurements of stratospheric aerosols above Spitzbergen during winter 1992/1993, *Geophys. Res. Lett.*, 21, 57–60, 1994.
- Blum, U., Khosrawi, F., Baumgarten, G., Stebel, K., Müller, R., and Fricke, K. H.: Simultaneous lidar observations of a polar stratospheric cloud on the east and west sides of the Scandinavian mountains and microphysical box model calculations, *Ann. Geophys.*, 24, 3267–3277, 2006, <http://www.ann-geophys.net/24/3267/2006/>.
- Bohren, C. F. and Huffman, D. R.: *Absorption and Scattering of Light by Small Particles*, J. Wiley & Sons, New York, 530 pp., 1983.
- Carslaw, K. S., Peter, T., and Clegg, S. L.: Modelling the composition of liquid stratospheric aerosols, *Rev. Geophys.*, 35, 125–154, 1997.
- Charlson, R. J. and Heintzenberg, J.: *Aerosol Forcing of Climate*, Wiley, Chichester, 416 pp., 1995.
- Cooney, J. and Pina, M.: Laser radar measurements of atmospheric temperature profiles by use of Raman rotational backscatter, *Appl. Opt.*, 15, 602–603, 1976.
- David, C., Godin, S., Mégie, G., Emery, Y., and Flésia, C.: Physical state and composition of Polar Stratospheric Clouds inferred from airborne lidar measurements during SESAME, *J. Atmos. Chem.*, 27, 1–16, 1997.
- Deshler, T., Nardi, B., Adriani, A., Cairo, F., Hansen, G., Fierli, F., Hauchecorne, A., and Pulvirenti, L.: Determining the index of refraction of polar stratospheric clouds above Andoya (69°N) by combining size-resolved concentration and optical scattering measurements, *J. Geophys. Res.*, 105(D3), 3943–3953, doi:10.1029/1999JD900469, 2000.
- Deshler, T., Hervig, M. E., Hofmann, D. J., Rosen, J. M., and Liley, J. B.: Thirty years of in situ stratospheric aerosol size distribution measurements from Laramie, Wyoming (41°N), using balloon-borne instruments, *J. Geophys. Res.*, 108(D5), 108(D5), 4167,

- doi:10.1029/2002JD002514, 2003.
- Elbern, H. and Schmidt, H.: Ozone episode analysis by four-dimensional variational chemistry data assimilation, *J. Geophys. Res.*, 106(D4), 3569–3590, doi:10.1029/2000JD900448, 2001.
- Fiocco, G. and Smullins, L. D.: Detection of scattering layers in the upper atmosphere (60–140 km) by optical radar, *Nature*, 199, 1275–1276, doi:10.1038/1991275a0, 1963.
- Hansen, G. and Hoppe, U. P.: Lidar observations of polar stratospheric clouds and stratospheric temperature in winter 1995/96 over northern Norway, *Geophys. Res. Lett.*, 24(2), 131–134, 1997.
- Hofmann, D. J. and Rosen, J. M.: Stratospheric sulphuric acid layer: Evidence for an anthropogenic component, *Science*, 208, 1368–1370, 1980.
- Hofmann, D. J.: Increase in the stratospheric background sulphuric acid aerosol mass in the past 10 years, *Science*, 248, 996–1000, 1990.
- Junge, C. E., Chagnon, C. W., and Manson, J. E.: Stratospheric aerosols, *J. Meteorol.*, 18, 81–108, 1961.
- Klett, J. D.: Stable analytical inversion solution for processing lidar returns, *Appl. Opt.*, 20, 2, 211–220, 1981.
- Klett, J. D.: Lidar inversion with variable backscatter/extinction ratios, *Appl. Opt.*, 24, 1638, 1638–1643, 1985.
- Krieger, U. K., Mössinger, J. C., Luo, B., Weers, U., and Peters, T.: Measurements of the refractive indices of H<sub>2</sub>SO<sub>4</sub>/HNO<sub>3</sub>/H<sub>2</sub>O solutions to stratospheric temperatures, *Appl. Opt.*, 39(21), 3691–3703, 2000.
- Lahoz, W. A., Geer, A. J., Bekki, S., Bormann, N., Ceccherini, S., Elbern, H., Errera, Q., Eskes, H. J., Fonteyn, D., Jackson, D. R., Khattatov, B., Marchand, M., Massart, S., Peuch, V. H., Rharmili, S., Ridolfi, M., Segers, A., Talagrand, O., Thornton, H. E., Vik, A. F., and von Clarmann, T.: The Assimilation of Envisat data (ASSET) project, *Atmos. Chem. Phys.*, 7, 1773–1796, 2007, <http://www.atmos-chem-phys.net/7/1773/2007/>.
- Larsen, N., Mikkelsen, I. B., Knudsen, B. M., Schreiner, J., Voigt, C., Mauersberger, K., Rosen, J. M., and Kjome, T.: Comparison of chemical and optical in situ measurements of polar stratospheric cloud particles, *J. Geophys. Res.*, 105(D1), 1491–1502, 2000.
- Larsen, N.: Polar stratospheric clouds Microphysical and optical models, Danish Meteorological Institute, scientific report, 00–06, 85 pp., 2000.
- Luo, B., Krieger, U. K., and Peter, T.: Densities and refractive indices of H<sub>2</sub>SO<sub>4</sub>/HNO<sub>3</sub>/H<sub>2</sub>O solutions to stratospheric temperatures, *Geophys. Res. Lett.*, 23, 3707–3710, 1996.
- McCormick, M. P., Hamill, P., Pepin, T. J., Chu, W. P., Swissler, T. J., and McMaster, L. R.: Satellite studies of the stratospheric aerosol, *Bull. Am. Meteorol. Soc.*, 60, 1038–1046, 1979.
- Mehrtens, H., von Zahn, U., Fierli, F., Nardi, N., and Deshler, T.: Type IPSC-particle properties, measurements at ALOMAR from 1995 to 1997, *Geophys. Res. Lett.*, 26(D5), 603–606, 1999.
- Mishchenko, M. I. and Travis, L. D.: Capabilities and limitations of a current FORTRAN implementation of the T-matrix method for randomly oriented rotationally symmetric scatterers, *J. Quant. Spectrosc. Radiat. Transfer*, 60, 309–324, 1998.
- Müller, H. and Quenzel, H.: Information content of multispectral lidar measurements with respect to the aerosol size distribution, *Appl. Opt.*, 24, 648–654, 1985.
- Müller, D., Wandinger, U., Althausen, D., Mattis, I., and Ansmann, A.: Retrieval of physical particle properties from lidar observations of extinction and backscatter at multiple wavelengths, *Appl. Opt.*, 37, 2260–2263, 1998.
- 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, 1999.
- 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, 1999.
- Müller, D., Wagner, F., Wandinger, U., Ansmann, A., Wendisch, M., Althausen, D., and von Hoyningen-Huene, W.: Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: experiment, *Appl. Opt.*, 39, 1879–1892, 2000.
- Nedeljkovic, D., Hauchecorne, A., and Chanin, M. L.: Rotational Raman lidar to measure the atmospheric temperature from the ground to 30 km, *IEEE Trans. Geosci. Remote Sens.*, 31, 90–101, 1993.
- Peter, T.: Microphysics and heterogeneous chemistry of polar stratospheric clouds, *Annu. Rev. Phys. Chem.*, 48, 785–822, 1997.
- Pinnick, R. G., Rosen, J. M., and Hofmann, D. J.: Stratospheric aerosol measurements III: Optical model calculations, *J. Atmos. Sci.*, 33, 304–314, 1976.
- Robock, A.: Comment on “Climate forcing by the volcanic eruption of Mount Pinatubo” by Douglass, D. H. and Knox, R. S., *Geophys. Res. Lett.*, 32, L20711, doi:10.1029/2005GL023287, 2005.
- Scarchilli, C., Adriani, A., Cairo, F., Di Donfrancesco, G., Buontempo, C., Snels, M., Moriconi, M. L., Deshler, T., Larsen, N., Luo, B., Mauersberger, K., Ovarlez, J., Rosen, J., and Schreiner, J.: Determination of polar stratospheric cloud particle refractive indices by use of in situ optical measurements and T-matrix calculations, *Appl. Opt.*, 44, 3302–3311, 2005.
- Tarantola, A.: *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, SIAM, 343 pp., 2005.
- Tikhonov, A. N. and Arsenin, V. Y.: *Solution of Ill-Posed Problems*, Wiley, New York, 258 pp., 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., Kolgotin, A., Müller, D., and Whiteman, D. N.: Information content of multiwavelength lidar data with respect to microphysical particle properties derived from eigenvalue analysis, *Appl. Opt.*, 44, 25, 5292–5303, 2005.
- Vömel, H., Rummukainen, M., Kivi, R., Jarhu, J., Turunen, T., Kyrö, E., Rosen, J., Kjome, N., and Oltmans, S.: Dehydration and sedimentation of ice particles in the Arctic stratospheric vortex, *Geophys. Res. Lett.*, 24, 795–798, 1997.

Von Zahn, U. G., von Cossart, G., Fiedler, J., Fricke, K. H., Nelke, G., Baumgarten, G., Rees, D., Hauchecorne, A., and Adolfsen, K.: The ALOMAR Rayleigh/Mie/Raman lidar: objectives, configuration, and performance, *Ann. Geophysicae*, 18, 815–833, 2000.

WMO (World Meteorological Organization) Scientific Assessment of Ozone Depletion: 2006, Global Ozone Research and Monitoring Project–Report No. 50, 572 pp., Geneva, 2007.