

Improved identification of the solution space of aerosol microphysical properties derived from the inversion of profiles of lidar optical data, part 2: simulations with synthetic optical data

ALEXEI KOLGOTIN,¹ DETLEF MÜLLER,^{2,3,*} EDUARD CHEMYAKIN,³ AND ANTON ROMANOV⁴

¹Physics Instrumentation Center, Troitsk, Moscow Region 142190, Russia

²University of Hertfordshire, Hatfield, Hertfordshire AL10 9AB, UK

³Science Systems and Applications, Inc. (SSAI), NASA LaRC, 1 Enterprise Parkway, Hampton, Virginia 23666, USA

⁴The National University of Science and Technology, Leninskii av. 4, Moscow 119049, Russia

*Corresponding author: d.mueller@herts.ac.uk

Received 29 August 2016; revised 12 September 2016; accepted 12 September 2016; posted 14 September 2016 (Doc. ID 266606); published 30 November 2016

We developed a mathematical scheme that allows us to improve retrieval products obtained from the inversion of multiwavelength Raman/HSRL lidar data, commonly dubbed “3 backscatter + 2 extinction” ($3\beta + 2\alpha$) lidar. This scheme works independently of the automated inversion method that is currently being developed in the framework of the Aerosol-Cloud-Ecosystem (ACE) mission and which is successfully applied since 2012 [Atmos. Meas. Tech. 7, 3487 (2014); “Comparison of aerosol optical and microphysical retrievals from HSRL-2 and in-situ measurements during DISCOVER-AQ 2013 (California and Texas),” in International Laser Radar Conference, July 2015, paper PS-C1-14] to data collected with the first airborne multiwavelength $3\beta + 2\alpha$ high spectral resolution lidar (HSRL) developed at NASA Langley Research Center. The mathematical scheme uses gradient correlation relationships we presented in part 1 of our study [Appl. Opt. 55, 9839 (2016)] in which we investigated lidar data products and particle microphysical parameters from one and the same set of optical lidar profiles. For an accurate assessment of regression coefficients that are used in the correlation relationships we specially designed the *proximate analysis* method that allows us to search for a first-estimate solution space of particle microphysical parameters on the basis of a look-up table. The scheme works for any shape of particle size distribution. Simulation studies demonstrate a significant stabilization of the various solution spaces of the investigated aerosol microphysical data products if we apply this gradient correlation method in our traditional regularization technique. Surface-area concentration can be estimated with an uncertainty that is not worse than the measurement error of the underlying extinction coefficients. The retrieval uncertainty of the effective radius is as large as $\pm 0.07 \mu\text{m}$ for fine mode particles and approximately 100% for particle size distributions composed of fine (submicron) and coarse (supermicron) mode particles. The volume concentration uncertainty is defined by the sum of the uncertainty of surface-area concentration and the uncertainty of the effective radius. The uncertainty of number concentration is better than 100% for any radius domain between 0.03 and $10 \mu\text{m}$. For monomodal PSDs, the uncertainties of the real and imaginary parts of the CRI can be restricted to ± 0.1 and ± 0.01 on the domains [1.3; 1.8] and [0; 0.1], respectively.

Published by The Optical Society under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

OCIS codes: (000.3860) Mathematical methods in physics; (010.0280) Remote sensing and sensors; (010.1110) Aerosols; (010.1120) Air pollution monitoring; (010.3640) Lidar.

<http://dx.doi.org/10.1364/AO.55.009850>

1. INTRODUCTION

Our recent publication [1] demonstrates significant progress in the use of multiwavelength lidar systems that measure backscatter coefficients at wavelength 355, 532, and 1064 nm and

extinction coefficients at 355 and 532 nm, commonly dubbed “ $3\beta + 2\alpha$.” Such systems require software capable of processing the optical aerosol profiles into profiles of particle microphysical parameters. In this context, we develop an unsupervised

automated inversion method as part of a project study within NASA's Aerosol-Cloud-Ecosystems (ACE) mission. This inversion method has been successfully applied since 2012 to data collected with the first airborne multiwavelength $3\beta + 2\alpha$ High-Spectral-Resolution Lidar (HSRL-2) developed at NASA Langley Research Center [1,2].

First results retrieved with this automated inversion are encouraging because they demonstrate good agreement to *in situ* data taken aboard an aircraft that flew next to an aircraft that carried HSRL-2 [1,2]. An in-depth analysis of the first results from this automated inversion shows that the main source of disagreement between the results from the automated inversion and the *in situ* data are related to the retrieved number and surface-area concentrations. In many cases, these outliers can be relatively easily identified by visual inspection of the retrieved profiles of the microphysical particle parameters.

In part 1 of our series of papers [3] we presented first results on a mathematical scheme that uses *gradient correlation relationships* among different particle microphysical properties of particle size distributions. We analyzed optical profiles from lidar for our investigations of particle microphysical properties [3]. We investigated how we can use the relationships found in that study to identify outliers of particle microphysical properties that result from the inversion of optical lidar data with the automated, unsupervised inversion algorithm mode.

We developed what we denote as the *proximate analysis method*. This method allows us not only to assess more accurately the regression coefficients that we found from our study of correlation relationships, see Ref. [3]. This method also allows us to search for a first-estimate solution space of particle microphysical parameters on the basis of a look-up table (LUT). The basis of that LUT is a synthetic optical data (SOD) bank that we generated and that we introduced in Section 3.1 in Ref. [3]. The SOD includes 63869 $3\beta + 2\alpha$ data sets computed from logarithmic-normal size distributions that have different mean radii, mean widths, and complex refractive indices.

We describe in Section 2 the gradient correlation method. We explain how correlation relationships are used as a constraint. In Section 3 we show numerical examples with synthetic optical data and compare retrieval results that we obtained with different approaches. Section 4 summarizes our results.

2. METHODOLOGY

A. Regression Equations for Gradient Correlation Method

Present-day, state-of-the-art multiwavelength Raman and HSRL multiwavelength lidar measures backscatter coefficients (β) at the wavelengths 355, 532, and 1064 nm, and extinction coefficients (α) at 355 and 532 nm. This combination of optical data (OD) is usually denoted as $3\beta + 2\alpha$, or simply $3 + 2$. In Ref. [3] we showed that profiles of optical data and particle microphysical parameters (PMP) are interdependent. We find correlation relationships that describe this interdependence. The correlations have a comparably high degree of confidence from the mathematical point of view. The correlation coefficients we obtain from regression analysis are close to $R^2 = 1$ for the most important intensive parameters p (IP) we are interested in. We find that surface-area concentration ($p = s$), the ratio of volume

concentration to effective radius ($p = v/r_{\text{eff}}$) and the product $p = n(r_{\text{mean}}^2 + \sigma^2)$ are linearly correlated with the extinction coefficient $\alpha(\lambda)$ measured at wavelength $\lambda = 355$ nm, i.e.,

$$p = a_p \alpha(355) + b_p, \quad (1a)$$

$$p = s, v/r_{\text{eff}}, \quad n(r_{\text{mean}}^2 + \sigma^2) \text{ in } \mu\text{m}^2 \text{ cm}^{-3} = \text{Mm}^{-1}. \quad (1b)$$

The parameters n , r_{mean} , and σ are number concentration, mean radius, and standard deviation, respectively, of a particle size distribution $f(r)$.

The regression coefficients (RCs) a_p and b_p in the correlation relationships [Eq. (1a)] fulfill the conditions

$$a_s \in [1.3; 1.9], \quad a_v \in [0.4; 0.6], \quad a_n \in [0.11; 0.15], \quad (2a)$$

$$b_p = 0, \quad p = s, v/r_{\text{eff}}, \quad n(r_{\text{mean}}^2 + \sigma^2), \quad (2b)$$

for any type of particle size distribution (PSD) as for example mono- and bimodal PSDs. Details can be found in part 1 of our series of papers [3].

The linear correlations hold true for effective radius r_{eff} versus the extinction-related Ångström exponent \hat{a}_α (EAE)

$$r_{\text{eff}} = a_r \hat{a}_\alpha + b_r, \quad (3)$$

where the RCs fulfill

$$a_r \in [-10; -0.038], \quad b_r \in [0.18; 0.5] \text{ in } \mu\text{m}. \quad (4)$$

However, variations of \hat{a}_α need to be small, for example ± 0.05 for $\hat{a}_\alpha \in [1; 2]$, and the correlation coefficient is lower, i.e., $R^2 \in [0.7; 1]$. More explanations can be found in part 1 [3].

In Ref. [3] we also investigated if there are more correlation relationships for other PMPs. We find linear correlations of mean radius $p = r_{\text{mean}}$ and standard deviation $p = \sigma$ with effective radius r_{eff}

$$p = a_p r_{\text{eff}} + b_p, \quad (5a)$$

$$p = r_{\text{mean}}, \sigma \text{ in } \mu\text{m}, \quad (5b)$$

where the RCs fulfill

$$a_m \in [0.17; 0.84], \quad a_\sigma \in [0.21; 0.31], \quad (6a)$$

$$b_p \approx 0, \quad p = r_{\text{mean}}, \sigma. \quad (6b)$$

The correlation coefficient R^2 is close to 1 in Eq. (5a) if aerosol properties do not change significantly between successive height bins of an aerosol profile. In terms of lidar data products, it means profiles of the lidar ratio α/β and the EAE at these height bins do not oscillate, and their variations are within ± 25 sr and ± 0.25 , respectively.

We note that the linear correlations of Eqs. (3) and (5a) are valid both for mono- and bi-modal PSDs. However, in contrast to the conditions of Eq. (2) the conditions in Eqs. (4) and (6) were obtained only for monomodal PSDs and change significantly in the case of bimodal PSDs.

Finally, we found in Ref. [3] that for particles of any mode fraction (fine versus coarse mode) of the PSD the complex refractive index $m = m_R - im_I$ (CRI) is uniquely defined by the extinction- (\hat{a}_α) and backscatter- (\hat{a}_β) (BAE) related Ångström exponents, the lidar ratios α/β (LR) at two wavelengths, and

the effective radius r_{eff} . We can use that result in the case of monomodal PSDs for locating the CRI in our LUT.

To locate the CRI, we collect the LUT elements that are closest to the data products \hat{a}_α , \hat{a}_β , and $\alpha(\lambda)/\beta(\lambda)$ which can be measured with lidar, and r_{eff} which can be retrieved with our regularization algorithm.

As a measure of closeness, we use in Ref. [3] the discrepancy

$$\rho_{\text{LUT}} = \frac{1}{5} \sum_p \frac{|p_{\text{LUT}} - p_{\text{experiment}}|}{p_{\text{LUT}}},$$

$$p = \hat{a}_\alpha, \hat{a}_{\beta(355), \beta(532)}, \alpha(355)/\beta(355), \alpha(532)/\beta(532), r_{\text{eff}}. \quad (7)$$

We consider only those LUT elements for which the discrepancy ρ_{LUT} lies in the vicinity of the minimal value $\rho_{\text{LUT}, \text{min}}$. The CRIs collected in this way and which fulfill the conditions of Eq. (7) are possible solutions of our search problem.

Equation (7), six correlation relationships, described by Eqs. (1), (3), and (5) are linked to eight important PMPs (r_{eff} , r_{mean} , σ , n , s , v , m_R , and m_I) which are commonly used for the description of atmospheric aerosol particles. They can be used *a priori* for constraining the solution space of aerosol microphysical properties in our underlying inversion problem (see Eq. (1) in Ref. [3]).

B. Constraints of the Solution Space

The lidar measurements deliver profiles of optical data. Thus, we want to derive the microphysical particle parameters not only for one but for several height layers. We therefore rewrite Eq. (1) in Ref. [3] as

$$\int_{r_{\text{min}}^{(l)}}^{r_{\text{max}}^{(l)}} K_g(\lambda, m^{(l)}, r) f^{(l)}(r) dr = g^{(l)}(\lambda),$$

$$l = 1, \dots, N_L; g = \alpha, \beta. \quad (8)$$

The superscript l indicates the number of the height bins we use to describe the OD profiles $g^{(l)}(\lambda)$ as well as the parameters $r_{\text{min}}^{(l)}$, $r_{\text{max}}^{(l)}$, $m^{(l)}$, and the function $f^{(l)}(r)$ in each height bin l . The total number of height bins is N_L . The kernel function K is known and its definition can be found in Ref. [4] for the case of spherical particle geometry.

Equation (8) can be solved, for example, with regularization which allows us to find a solution space $F^{(l)} = \{f_1^{(l)}(r), \dots, f_{N_{\text{sol}}}^{(l)}(r)\}$ for each height bin l [5–8] where N_{sol} is the total number of individual solutions. If we know $F^{(l)}$ we can also obtain the solution spaces for any bulk parameter $P^{(l)} = \{p_1^{(l)}, \dots, p_{N_{\text{sol}}}^{(l)}\}$ for each height bin l , where $P^{(l)} = R_{\text{mean}}^{(l)}, R_{\text{eff}}^{(l)}, \Sigma^{(l)}, N^{(l)}, S^{(l)}$, and $V^{(l)}$ denote the solution space of r_{mean} , r_{eff} , σ , n , s , and v , respectively, see Ref. [5]. The collection of the solution spaces for all N_L heights forms the profiles of the PMPs. The final solution p_{av} at height l is defined by averaging all individual parameters $p_i^{(l)}$, $i \leq N_{\text{sol}}$, over a prescribed interval of the discrepancy range $[\rho_{\text{min}}^{(l)}, \rho_{\text{av}}^{(l)}]$ [6,8].

To find physically meaningful solutions we need to apply different physical constraints, as for example, the discrepancy range $[\rho_{\text{min}}^{(l)}, \rho_{\text{av}}^{(l)}]$, the radius range $[r_{\text{min}}^{(l)}, r_{\text{max}}^{(l)}]$, and the complex refractive index $m^{(l)}$. In that way we can exclude individual solutions

$p_i^{(l)}$ which for example deviate too much from the average value p_{av} . In that case, we insert a threshold δp that allows us to take into consideration only those individual solutions that fulfill the condition

$$|p_i^{(l)} - p_{\text{av}}| < \delta p. \quad (9)$$

In our previous strategy of data inversion we used this averaging procedure for each solution space (in each height bin) independently of each other, i.e., we did not take into consideration that results of PMPs of successive height bins may be correlated to each other.

We developed a mathematical framework that allows us to consider the correlation in a very specific way that is described by Kolgotin and Müller [8] and Müller *et al.* [9]. However, that method is complex, time consuming, and will require more effort to understand how it can be efficiently used for the automated, unsupervised inversion method we are developing for HSRL-2. In the method presented here we provide an elegant, efficient way of using the profile information from lidar without including complications from the mathematical point of view as discussed by Kolgotin and Müller [8].

The profiles of the ODs contain information regarding the variation of the profiles of the PMPs. In other words, we can predict the PMP behavior if we know the law(s) (or correlations) that describe the interdependence of the OD and the PMPs.

In that regard, the linear correlation as described by Eq. (3) can be used as an extra constraint if we want to find the solution space for $R_{\text{eff}}^{(l)}$. In fact, if we measure $\hat{a}_\alpha^{(l)}$, the correlation presented in Eq. (3) allows us to estimate $r_{\text{eff}}^{(l)*}$ according to

$$r_{\text{eff}}^{(l)*} = a_r \hat{a}_\alpha^{(l)} + b_r. \quad (10)$$

If we introduce the threshold δr_{eff} we can rewrite Eq. (9) as

$$|r_{\text{eff}, i}^{(l)} - r_{\text{eff}}^{(l)*}| < \delta r_{\text{eff}}. \quad (11)$$

The linear correlations as described by Eq. (1) can be used as extra constraints to identify the solution spaces $R_{\text{mean}}^{(l)} \cup \Sigma^{(l)} \cup N^{(l)}$, $V^{(l)} \cup R_{\text{eff}}^{(l)}$, and $S^{(l)}$ as well; the symbol \cup means the space conjunction. Again, introducing the threshold δp the condition in Eq. (11) can be rewritten in a more general form as

$$|p_i^{(l)} - p^{(l)*}| < \delta p \quad \text{and} \quad p = s, v/r_{\text{eff}}, n(r_{\text{mean}}^2 + \sigma^2). \quad (12)$$

Here $p^{(l)*}$ is defined by Eq. (1) where all $b_p = 0$ and $a_s = 1.6$, $a_v = 0.5$, and $a_n = 0.13$ with threshold $\delta p = 20\%$. It is clear that the RCs can be made more exact with the conditions described in Eq. (2) depending on the particle properties, e.g., particle radius (size) and CRI. The use of the constraints as described by Eqs. (11) and (12) in the averaging procedure [6,8] is at the heart of the gradient correlation method (GCM).

3. NUMERICAL SIMULATIONS

A. Proximate Analysis of Particle Parameters

The conclusions of the previous section naturally lead us to the step in which we want to develop a practical method that allows us to identify the parameters of fine-mode particles even if the investigated PSDs may contain coarse mode particles (bimodal size distribution). This situation until now has not been tested in simulation studies carried out with the automated software.

HSRL-2 and multiwavelength Raman lidars can measure the extinction coefficients $\alpha(\lambda)$ at $\lambda = 355$ and 532 nm. As a result, the EAE is available from Eq. (19) in [3]. From our own measurements and literature we know that the EAE often varies between $\hat{a}_\alpha = 1$ and 2 except in cases of mineral dust, volcanic ash, and aged smoke. These aerosol types are characterized by comparably large particles. We find the following results:

A) the effective radius is proportional to \hat{a}_α with threshold $\delta r_{\text{eff}} \leq 0.07 \mu\text{m}$ (see Fig. 1(a) in Ref. [3]) and it can be determined from Eq. (10) as

$$r_{\text{eff}} \approx -0.08\hat{a}_\alpha + 0.26, \tag{13}$$

B) the surface-area concentration can be estimated from Eq. (3) as well. For example, if $a_s = 1.6$ we find

$$s \approx 1.6\alpha(355), \tag{14}$$

C) volume concentration is defined on the basis of Eq. (8), see Ref. [3]

$$v \approx 1.6\alpha(355) \times (-0.08\hat{a} + 0.26)/3. \tag{15}$$

If we assume that the standard deviation is not too large so that $\sigma \ll r_{\text{mean}}/\sqrt{3}$, see Eq. (29) in Ref. [3], we find that $r_{\text{eff}} \approx r_{\text{mean}}$. For example, if we select from our LUT optical data such that $\hat{a} > 1$ we find $r_{\text{mean}} < 0.25 \mu\text{m}$. That means σ should be less than 0.15 . In contrast, the maximal value is $\sigma_{\text{max}} < 0.09$ for all elements in the LUT bank if $\hat{a} > 1$. In that case, we obtain the number concentration from Eq. (15) in Ref. [3] according to

$$n \approx \frac{1.6\alpha(355)}{4\pi(-0.08\hat{a} + 0.26)^2}. \tag{16}$$

Equations (13)–(16) allow us to find the most important PSD parameters very easily and fast, and their uncertainties can be estimated if we use as threshold the value $\delta r_{\text{eff}} = \pm 0.07 \mu\text{m}$ and Eq. (2a) at $p = s$. However, in this approach we must be sure that we deal with monomodal PSDs. Otherwise, Eq. (3) does not work with the condition $\delta r_{\text{eff}} \leq 0.07 \mu\text{m}$, see for example the stars in Fig. 1(a) in Ref. [3].

For that reason we must adapt Eqs. (13)–(16) such that they hold true in the case of bimodal PSDs. To achieve this goal, we split the optical coefficients into the fine-mode contribution $g_f(\lambda)$ and the coarse-mode contribution $g_c(\lambda)$, i.e., we write

$$g(\lambda) = g_f(\lambda) + g_c(\lambda) = \varphi_g g(\lambda) + (1 - \varphi_g)g(\lambda), \quad g = \alpha, \beta, \tag{17}$$

with the fractions φ_g (fine mode) and $1 - \varphi_g$ (coarse mode), respectively.

Apparently, the ratio of the fine mode to the coarse mode, expressed in terms of the optical coefficients is

$$\frac{g_c(\lambda)}{g_f(\lambda)} = \frac{g(\lambda)}{g_f(\lambda)} - 1 = \frac{1}{\varphi_g} - 1 \quad g = a, b. \tag{18}$$

If we use the fine-mode and coarse-mode extinction coefficients measured at 532 nm, i.e., $\alpha_f(532)$ and $\alpha_c(532)$ we can find the ratio of the extinction coefficients at 532 and 355 nm, i.e.,

$$\frac{\alpha_f(532) + \alpha_c(532)}{\alpha_f(355) + \alpha_c(355)} = \frac{\alpha(532)}{\alpha(355)}, \tag{19}$$

and from Eq. (19) follows

$$\alpha_f(532) = \frac{\alpha(532)}{\alpha(355)}[\alpha_f(355) + \alpha_c(355)] - \alpha_c(532). \tag{20}$$

If we express $\alpha_f(355)$ in Eq. (18) in terms of $\varphi_{\alpha(355)}$ and $\alpha_c(355)$ we obtain the ratio of the extinction coefficients at 532 and 355 nm of the fine mode fractions, i.e.,

$$\frac{\alpha_f(532)}{\alpha_f(355)} = \frac{1}{\varphi_{\alpha(355)}} \left[\frac{\alpha(532)}{\alpha(355)} + d_c(\varphi_{\alpha(355)} - 1) \right], \tag{21}$$

where the parameter d_c is defined as

$$d_c = \alpha_c(532)/\alpha_c(355). \tag{22}$$

As we discussed in Ref. [3], we find $d_c \in [1; 1.07]$ for large particles whose effective radius is larger than $1.1 \mu\text{m}$.

According to the definition in Eq. (19) in Ref. [3], we can write for the EAE which describes the fine mode

$$\hat{a}_{a,f} = \frac{\ln \left\{ \frac{1}{\varphi_{\alpha(355)}} \left[\frac{\alpha(532)}{\alpha(355)} + d_c(\varphi_{\alpha(355)} - 1) \right] \right\}}{\ln \frac{355}{532}} \geq \hat{a}_a. \tag{23}$$

Now we can find all microphysical parameters of the fine mode fraction of a bimodal particle size distribution. In fact, in view of Eqs. (13)–(15) the effective radius, surface-area, and volume concentrations for the fine mode are

$$r_{\text{eff},f} \approx -0.08\hat{a}_{a,f} + 0.26, \tag{24a}$$

$$s_f \approx 1.6\alpha_f(355) = 1.6\varphi_{\alpha(355)}\alpha(355), \tag{24b}$$

$$v_f \approx s_f r_{\text{eff},f} / 3. \tag{24c}$$

In the case of a fine-mode PSD, the inequality $r_{\text{eff}} < 3r_{\text{mean}}$ works quite well. Then the maximum value of the number concentration can be obtained from Eq. (15) in Ref. [3] and the inequality [Eq. (31b)] presented in Ref. [3] is

$$n_{f,\text{max}} = \frac{1}{4\pi} \frac{s_f}{r_{\text{mean},f}^2 + \sigma_f^2} \leq \frac{1}{2\pi} \frac{s_f}{r_{\text{eff},f}^2}. \tag{24d}$$

The minimum level can be obtained from Eq. (15) in Ref. [3], and the inequality shown in Eq. (31c) in Ref. [3] becomes

$$n_{f,\text{min}} = \frac{1}{4\pi} \frac{s_f}{r_{\text{mean},f}^2 + \sigma_f^2} \geq \frac{1}{4\pi} \frac{s_f}{r_{\text{eff},f}^2}. \tag{24e}$$

According to these equations, the fine mode fraction varies from $\varphi_{\alpha(355)} = 1$ to some minimal value $\varphi_{\alpha(355)} = \varphi_{\alpha(355)\text{min}} \approx 0.5$ at which $r_{\text{eff},f} \leq 0$ and the results become unphysical.

If $\varphi_{\alpha(355)} = 1$, we assume that the particles are distributed according to a monomodal distribution law. On the contrary, if $\varphi_{\alpha(355)} < 1$ there is a pronounced coarse mode.

If we want to estimate the CRI of the fine mode we need to find at least three other fractions $\varphi_{\beta(355)}$, $\varphi_{\alpha(532)}$, and $\varphi_{\beta(532)}$, and extract optical data spectra $\alpha_f(\lambda)$ and $\beta_f(\lambda)$ which describe the fraction of small particles, as we discussed in Section 3.2 in Ref. [3]. We find all fractions more or less accurately if we take into account our LUT. If we predefine $\varphi_{\alpha(355)}$ and d_c in

Eq. (23) and if we consider that $\alpha_c = (1 - \varphi_\alpha)\alpha$ we automatically find from Eq. (22)

$$\varphi_{\alpha(532)} = 1 - (1 - \varphi_{\alpha(355)})d_c \frac{\alpha(355)}{\alpha(532)}. \quad (25)$$

If we subsequently select from our LUT the LR's defined by $\alpha_f(355)/\beta_f(355)$ and $\alpha_f(532)/\beta_f(532)$ we can define another pair of fractions from the equations, i.e.,

$$\varphi_{\beta(355)} = \varphi_{\alpha(355)} \frac{\alpha(355)\beta_f(355)}{\beta(355)\alpha_f(355)}, \quad (26)$$

and

$$\varphi_{\beta(532)} = \varphi_{\alpha(532)} \frac{\alpha(532)\beta_f(532)}{\beta(532)\alpha_f(532)}. \quad (27)$$

These values $\alpha_f(355)/\beta_f(355)$ and $\alpha_f(532)/\beta_f(532)$ are selected so that the respective BAE is equal to

$$\hat{a}_{\beta(355),\beta(532),f} = \hat{a}_{\beta(355),\beta(532)} + \ln \frac{\varphi_{\beta(532)}}{\varphi_{\beta(355)}} \ln^{-1} \frac{355}{532}. \quad (28)$$

Since lidar delivers $\beta(1064)$ and thus also the BAE at the wavelength pair of 532 and 1064 nm we can estimate a fifth unknown fraction with the help of our LUT, i.e.,

$$\varphi_{\beta(1064)} = \varphi_{\beta(532)} \left[\frac{532}{1064} \right]^{\hat{a}_{\beta(532),\beta(1064),f} - \hat{a}_{\beta(532),\beta(1064)}}. \quad (29)$$

The four equations [Eq. (23)] and [Eqs. (26)–(28)] describe the four unknown fractions and thus ensure the unique definition of $\varphi_{\alpha(355)}$, $\varphi_{\beta(355)}$, $\varphi_{\alpha(532)}$, and $\varphi_{\beta(532)}$ for any predefined values of $\varphi_{\alpha(355)}$ and d_c . The parameter d_c is quite strictly constrained by the interval [1; 1.07], whereas the fraction $\varphi_{\alpha(355)}$ can vary between 0.4 and 1. In spite of the fact that we apparently have the constraints

$$\begin{aligned} 0 \leq \varphi_{\alpha(532)} < \varphi_{\alpha(355)} \leq 1, \\ 0 \leq \varphi_{\beta(1064)} < \varphi_{\beta(532)} < \varphi_{\beta(355)} \leq 1, \quad d_c \in [1; 1.07]. \end{aligned} \quad (30)$$

These constraints are still not sufficient for the unambiguous estimation of the unknown fractions.

We suggest to test the quality of the unknown fractions by comparing the derived LR's, BAE's, and EAE of the coarse mode, which is described by $g_c = (1 - \varphi_g)g$, to the respective parameters LR's, BAE's, and EAE of our LUT. We modify the discrepancy expressed by Eq. (7) to

$$\begin{aligned} \rho_{\text{LUT},c} &= \frac{1}{5} \sum_p \frac{|p_{\text{LUT}} - p_c|}{p_{\text{LUT}}}, \\ p &= \hat{a}_\omega, \hat{a}_{\beta(355),\beta(532)}, \hat{a}_{\beta(532),\beta(1064)}, \alpha(355)/\beta(355), \\ &\quad \alpha(532)/\beta(532), \end{aligned} \quad (31)$$

which serves as a measure of the quality, or accuracy of describing the remaining part of the total spectrum $g(\lambda)$. The smaller the discrepancy, which is described by Eq. (31) the better is the quality of $\varphi_{\alpha(355)}$, $\varphi_{\beta(355)}$, $\varphi_{\alpha(532)}$, $\varphi_{\beta(532)}$, and $\varphi_{\beta(1064)}$. Our approach can be used if the LUT includes, at minimum, all elements that allow us to cover all possible combinations for r_{eff} . Thus, we need to extend in the future the LUT with synthetic data calculated for PSDs with $r_{\text{eff}} > 3 \mu\text{m}$.

We can test this approach in the case of the bimodal PSDs considered in Section 3.1 in Ref. [3]. The test can be easily carried out, for example with the help of a table written in Excel to make the calculations of the formulas shown in Eqs. (23) and (24) faster. The example is presented in Table 1 [see also Fig. 1, asterisks].

Table 1 contains the columns $l = 1, \dots, N_L = 6$ which describe different bimodal PSDs or, in other words, height bins that describe a profile of optical data taken by lidar. The actual fraction $\varphi_{\alpha(355)}$ of the fine mode particles decreases with height from 1 to 0.49. According to Eq. (17), this fraction describes the fine-mode fraction in terms of the extinction coefficient at 355 nm. If we use Eq. (7) in Ref. [3], we see that in fact we describe the fine-mode fraction of the surface-area concentration. All actual (true) microphysical parameters of the fine mode of the PSDs are fixed. We show them just for reference.

The profile of the synthetic extinction coefficient $\alpha(355)$ increases with height from 0.093 to 0.190 Mm^{-1} . The EAE decreases from 1.64 to 0.61 because the contribution of coarse mode particles increases with height. We use these data as input (see green symbols in Table 1) for the calculation of the particle microphysical parameters according to Eq. (24).

The first step of our computations is to estimate the fine mode fraction $\varphi_{\alpha(355),\text{fit}}$ by means of a fitting procedure (see red numbers in Table 1). Let us assume there is no coarse mode at the lowest height bin $l = 1$, i.e., $\varphi_{\alpha(355),\text{fit}} = 1$. In that case, the EAE of the fine mode can be found from Eq. (23). Table 1 shows that $\hat{a}_{\alpha,f}^{(1)} = 1.64$ which coincides with the actual value $\hat{a}_{\alpha,f}^{(1)}$. If, in addition, we assume that the fine-mode effective radius does not change with height we can select $\varphi_{\alpha(355),\text{fit}}^{(l)}$ in Eq. (23), which leads to $\hat{a}_{\alpha,f}^{(l)} \approx 1.64$ for $l = 2, \dots, N_L$. The result is that the numbers $\varphi_{\alpha(355),\text{fit}}^{(l)}$ reproduce the true values $\varphi_{\alpha(355)}^{(l)}$ rather well.

We emphasize that the effective radius as determined by Eq. (24a) is equal to $\sim 0.05 \mu\text{m}$ if we significantly decrease $\varphi_{\alpha(355),\text{fit}}$ to 0.77, i.e., we assume that the coarse mode exists at $l = 1$. This value of the effective radius is very small. Given the available measurement wavelengths, such a small value cannot be retrieved in a reliable way with this automated inversion software. That means the fine mode fraction $\varphi_{\alpha(355),\text{fit}}$ cannot be varied over a wide range unless we want to produce unreasonable values of the microphysical particle parameters.

After we fit the profile $\varphi_{\alpha(355),\text{fit}}^{(l)}$ according to the condition $\hat{a}_{\alpha,f}^{(l)} \approx \text{const}$ all microphysical parameters of the fine mode of the investigated PSD can be calculated with the help of the formulas given in Eq. (24). As seen from Table 1 (red symbols), the parameters do not change with height, and the values are close to the true values. We find a maximum relative error of 30% for volume concentration.

If we take into account the threshold $\delta r_{\text{eff}} = \pm 0.07 \mu\text{m}$ for effective radius we can estimate the uncertainties of all parameters of the fine mode particles. In that case the maximum level of uncertainty is $0.2 \mu\text{m}$ for effective radius, and it is $v_{f,\text{max}} = 0.01 \mu\text{m}^3 \text{cm}^{-3}$ for volume concentration. If the PSD shows a fine mode of particles, we can estimate the maximum level of the number concentration from Eq. (24d), i.e., $n_{f,\text{max}} = 1.5 \text{cm}^{-3}$.

Table 1. Example of PA of Particle Microphysical Parameters of the Fine Mode Fraction of a PSD

<i>l</i>	1	2	3	4	5	6	Remark
For reference							
$\Phi_{\beta(355)}$	1.00	0.91	0.84	0.78	0.73	0.35	
$\Phi_{\beta(532)}$	1.00	0.82	0.69	0.60	0.53	0.19	
$\Phi_{\beta(1064)}$	1.00	0.61	0.44	0.35	0.28	0.07	
$\Phi_{\alpha(355)}$	1.00	0.95	0.91	0.87	0.83	0.49	
$\Phi_{\alpha(532)}$	1.00	0.90	0.83	0.76	0.70	0.32	
$\hat{a}_{\beta(532;1064)}$	1.09	0.67	0.44	0.29	0.19	-0.24	
n_f, cm^{-3}	1.0						
$r_{\text{eff},f}, \mu\text{m}$	0.15						
$s_f, \mu\text{m}^2\text{cm}^{-3}$	0.17						
$v_f, \mu\text{m}^3\text{cm}^{-3}$	0.0086						
m_R	1.45						
m_I	0.005						
Input							
$\alpha(355), \text{Mm}^{-1}$	0.093	0.098	0.103	0.108	0.113	0.190	Input for Eq. (24)
$\alpha(532), \text{Mm}^{-1}$	0.048	0.053	0.058	0.063	0.068	0.149	
\hat{a}_{α}	1.64	1.52	1.42	1.33	1.24	0.61	
Fitted							
$\Phi_{\alpha(355),\text{fit}}$	1.00	0.95	0.91	0.87	0.83	0.49	Input for Eq. (23)
Output							
$\hat{a}_{\alpha,f}$	1.64	1.64	1.63	1.63	1.63	1.62	Eq. (23)
$r_{\text{eff},f}, \mu\text{m}$	0.13						Eq. (24a)
$s_f, \mu\text{m}^2\text{cm}^{-3}$	0.15						Eq. (24b)
$v_f, \mu\text{m}^3\text{cm}^{-3}$	0.006						Eq. (24c)
$n_{f,\text{min}} \text{cm}^{-3}$	0.75						Eq. (24e)
$n_{f,\text{max}}, \text{cm}^{-3}$	1.5						Eq. (24d)
$m_{R,f}$	1.425–1.7						LUT
$m_{I,f}$	0.003–0.025						LUT

If we repeat that approach with the assumption that the coarse mode is absent, i.e., $\varphi_{\alpha(355),\text{fit}}^{(l)} = 1$ we obtain a similar result for layers $l = 1, \dots, 5$ within the uncertainties $\delta p^{(1)}$. With regard to layer $l = 6$, the surface-area and volume concentrations are overestimated significantly. We stress that we do not make any assumption about monomodality of the PSD if the BAE is very close to 0 or even negative, and the EAE is above 0.5 simultaneously.

Let us estimate the CRI by using the correlation relationships. Since we assume that the PSD is monomodal at $l = 1$, i.e., $g_f^{(1)}(\lambda) = g^{(1)}(\lambda)$ and $r_{\text{eff},f}^{(1)} = r_{\text{eff}}^{(1)}$, we can calculate the lidar data products BAEs, EAE, and LRs at the measurement wavelengths, and the discrepancy ρ_{LUT} , see Eq. (7). The

minimum discrepancy $\rho_{\text{LUT},\text{min}} \approx 2.8\%$ is obtained from the LUT for the following set of values:

- $\hat{a}_{\alpha,f} = 1.66$, $\hat{a}_{\beta(355),\beta(532),f} = 1.45$, $\hat{a}_{\beta(532),\beta(1064),f} = 1.04$,
- lidar ratios 76 sr and 70 sr at $\lambda = 355$ and 532 nm, respectively,
- $r_{\text{eff}} = 0.12 \mu\text{m}$, and
- $m = 1.7 - i0.025$.

The interval $[\rho_{\text{LUT},\text{min}}, \rho_{\text{LUT}}] = [2.8\%; 5\%]$ includes 12 sets of parameters for which effective radius $r_{\text{eff}} = 0.12\text{--}0.15 \mu\text{m}$, the real part of the CRI $m_R = 1.425\text{--}1.7$ and the imaginary part $m_I = 0.002\text{--}0.025$. We note that the true effective radius (almost $0.15 \mu\text{m}$) lies between $r_{\text{eff}} = 0.13$ and $0.15 \mu\text{m}$. For

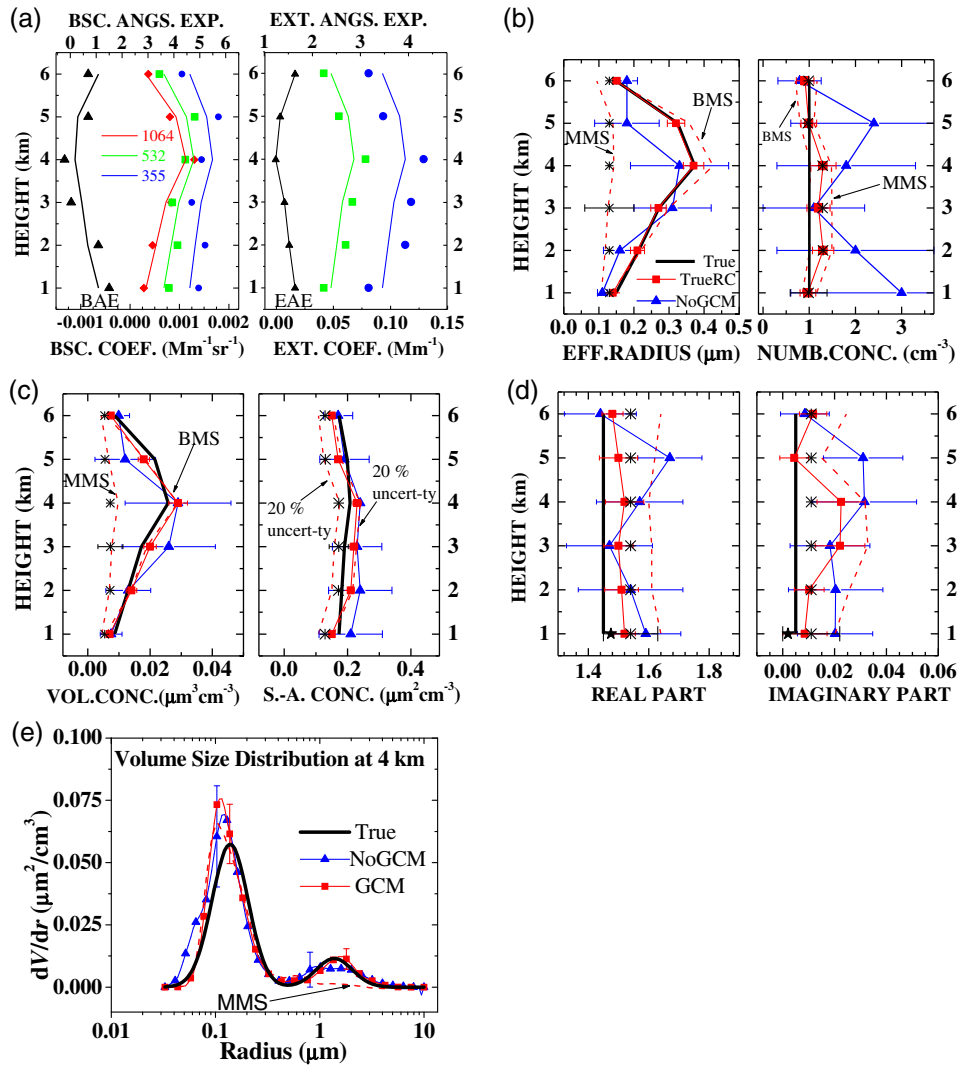


Fig. 1. Numerical example: vertical profile of type 1. Shown are the retrieval results for the case of a $3\beta + 2\alpha$ data set, i.e., (a) true input data (solid line) and distorted data (closed symbols) of the backscatter and extinction coefficients. Also shown are the respective BAE for the wavelength pair 532/1064 nm and the EAE for the wavelength pair 355/532 nm. (b)–(e): true (thick line) and retrieved (line + symbol) results for PMP, CRI, and PSD. The retrieval results were obtained when GCM was not used (triangles), when GCM was used with the true RCs (squares), and when GCM was used with incorrect RCs (dashed line). We assumed that the true PSD was either monomodal (MMS) or bimodal (BMS) and that $a_s = 1.6$. The asterisks describe the PA of the fine mode parameters. Stars denote the retrieval results of the CRI of the fine mode particles that follow from the combination of PA with the results of an effective radius that is obtained with GCM.

that range of effective radii, we find $m = 1.475 - i0.005$ and $1.425 - i0.003$, respectively. However, we see that quite small variations of the effective radius result in a wide variation of the CRI.

We can use this approach for all other height bins $l = 2, \dots, 6$ under the assumption that the PSD is monomodal. Since the uncertainty δr_{eff} does not decrease with height, we do not expect an improvement of the estimated values of the CRI at $l = 2, \dots, 6$.

Now we consider the general case in which we do not make any assumption about the fraction $\varphi_{\alpha(355)}$. For that purpose, we analyze the data at height bin $l = 6$. As we discussed before, the SOD bank does not contain an optical data set that is consistent with $g^{(6)}(\lambda)$. Therefore, we use our idea to extract the spectrum $g_f^{(6)}(\lambda)$ of the fine mode PSD with the help of the

fractions $\varphi_{g(\lambda)}$. We consider 13 initial values for the fraction $\varphi_{\alpha(355)}$ from 0.4 up to 1 with step size 0.05. For each of these 13 initial values we find from Eqs. (23) and (26)–(29) the 13 discrepancies which are defined by Eq. (31).

The minimum discrepancy $\rho_{\text{LUT},c} = 6\%$ is obtained for $\varphi_{\alpha(355)} = 0.5$. The other discrepancies are as large as $\rho_{\text{LUT},c} = 21\%$ if $\varphi_{\alpha(355)}$ increases to 0.9 and as large as $\rho_{\text{LUT},c} = 10\%$ if $\varphi_{\alpha(355)}$ decreases to 0.4 (see Table 2). For the fraction values $\varphi_{\alpha(355)} = 0.95$ and 1, there are either no solutions that fulfill Eq. (30) or there are no solutions that have a low discrepancy. We note that $\rho_{\text{LUT},c} = 4\%$ for the true value $\varphi_{\alpha(355)} = 0.49$. The results in Table 2 are obtained at $d_c \approx 1.03$ which is about in the center of the interval [1; 1.07]. Other d_c -values increase all discrepancies $\rho_{\text{LUT},c}(\varphi_{\alpha(355)})$ similarly, but they do not change the trend for $d_c \approx 1.03$.

Table 2. Particle Fine Mode Fractions Found with PA for the Example of Height Bin $l = 6$ in Table 1

$\varphi_{\alpha(355)}$	$\varphi_{\alpha(532)}$	$\varphi_{\beta(355)}$	$\varphi_{\beta(532)}$	$\varphi_{\beta(1064)}$	$\rho_{LUT,c}$
0.4	0.21	0.39	0.22	0.04	10%
0.45	0.27	0.39	0.22	0.06	8%
0.49	0.33	0.45	0.25	0.07	4%
0.5	0.34	0.49	0.26	0.08	6%
0.55	0.40	0.44	0.26	0.08	9%
0.6	0.47	0.36	0.18	0.11	7%
0.65	0.54	0.48	0.30	0.11	7%
0.7	0.60	0.30	0.20	0.09	15%
0.75	0.67	0.40	0.27	0.09	12%
0.8	0.74	0.65	0.42	0.17	16%
0.85	0.80	0.56	0.35	0.11	16%
0.9	0.87	0.53	0.37	0.17	21%

We analyzed the fractions $\varphi_{g(\lambda)}$ we found in this way in the vicinity of the minimum discrepancy, i.e., from $\rho_{LUT,c} = 4\%$ to 10% (see green lines in Table 2, last column). This vicinity covers the range of values from 0.4 to 0.65 for $\varphi_{\alpha(355)}$ and from 0.04 to 0.11 for $\varphi_{\beta(1064)}$ (see Table 2). At the point $\varphi_{\alpha(355)} = 0.49$, we retrieve $\varphi_{\beta(355)} = 0.45$, whereas the true value is $\varphi_{\beta(355)} = 0.35$ (see Table 1). The uncertainty is 29% at the point $\varphi_{\alpha(355)} = 0.49$, which is maximal compared with all other fractions (see Table 1 for reference). If we average the fractions $\varphi_{g(\lambda)}$ in the vicinity $\rho_{LUT,c} = 4\%$ to 10% we obtain the following values for the height bin $l = 6$:

$$\begin{aligned} \varphi_{\alpha(355)}^{(6)} &= 0.52, & \varphi_{\alpha(532)}^{(6)} &= 0.37, & \varphi_{\beta(355)}^{(6)} &= 0.43, \\ \varphi_{\beta(532)}^{(6)} &= 0.24, & \varphi_{\beta(1064)}^{(6)} &= 0.08. \end{aligned}$$

For this set of fractions, the EAE of the fine mode is 1.48. If we use $\hat{a}_{\alpha,f} = 1.48$ and $\varphi_{\alpha(355)}^{(6)} = 0.52$ in Eqs. (24a)–(24e) we find values of the PMPs that are close to the values mentioned in Table 1. We also obtain the range of the CRI of the fine-mode PSD on the basis of the analysis of the lidar data products and the discrepancy ρ_{LUT} [see Eq. (7)]. The results in the height bins $l = 1$ and $l = 6$ are similar, too.

Our approach for estimating the fine mode parameters of a PSD are summarized into the following steps:

A) We make an analysis regarding the number of modes of the PSD, i.e., whether it is monomodal (1 mode) or bimodal (2 modes) by means of a comparison of the measured properties and the lidar data products in the LUT. If the comparison is consistent (see, for example, case #2 in Table 1 in Ref. [3]), the PSD contains a coarse mode which can be “cut” by extracting the fine mode spectrum $g_f(\lambda) = \varphi_g g(\lambda)$. The fractions φ_g are estimated with Eqs. (23) and (26)–(29) and we minimize the discrepancy according to Eq. (31).

B) We calculate the EAE with Eq. (23).

C) We calculate the effective radius from the system of Eq. (33) in Ref. [3]. The selection of Eq. (33a), or Eq. (33b), or Eq. (33c) depends on the range of values of the EAE. The uncertainty is $\delta r_{\text{eff}} = \pm(0.05\text{--}0.08) \mu\text{m}$ and depends on the EAE range as well.

D) We calculate the surface-area concentration from Eq. (24b). The uncertainty is $\delta s = \pm 20\%$.

E) We calculate the volume concentration from Eq. (24c). The uncertainty δv can be derived from the “worst” possible combination of δr_{eff} and δs .

F) We calculate the limits of number concentration from Eqs. (24d) and (24e). The minimum and maximum limits take into account the uncertainties δr_{eff} and δs .

G) We estimate the CRI on the basis of the analysis of the lidar data products according to the spectrum $g_f(\lambda)$ and the discrepancy ρ_{LUT} [see Eq. (7)].

Our proximate analysis (PA) is useful for preliminary estimations of microphysical parameters of particles in the fine mode of the PSD without applying the method of data inversion and the calculation of kernel functions. PA also allows for the more robust initial assessment of RCs used in the correlation relationships, see Eq. (10). Besides that, a preliminary estimation of the microphysical parameters can be employed as a constraint for more complicated methods of optical data analysis with inversion methods.

B. Retrieval Example for Synthetic Data

1. Numerical Example: Vertical Profile Type 1

In this example, we investigate in more detail the case considered in Section 3.A. The only difference is that we replace layer #6 by layer #1 and change the order of layer #4 and layer #5 to keep the optical data profile more realistic. Besides these changes, we distort the optical data with 15% extreme error. Extreme error in the context of this study means that each of the $3\beta + 2\alpha$ channels is independently distorted with an error of fixed magnitude of $\pm 15\%$. Figure 1(a) shows the

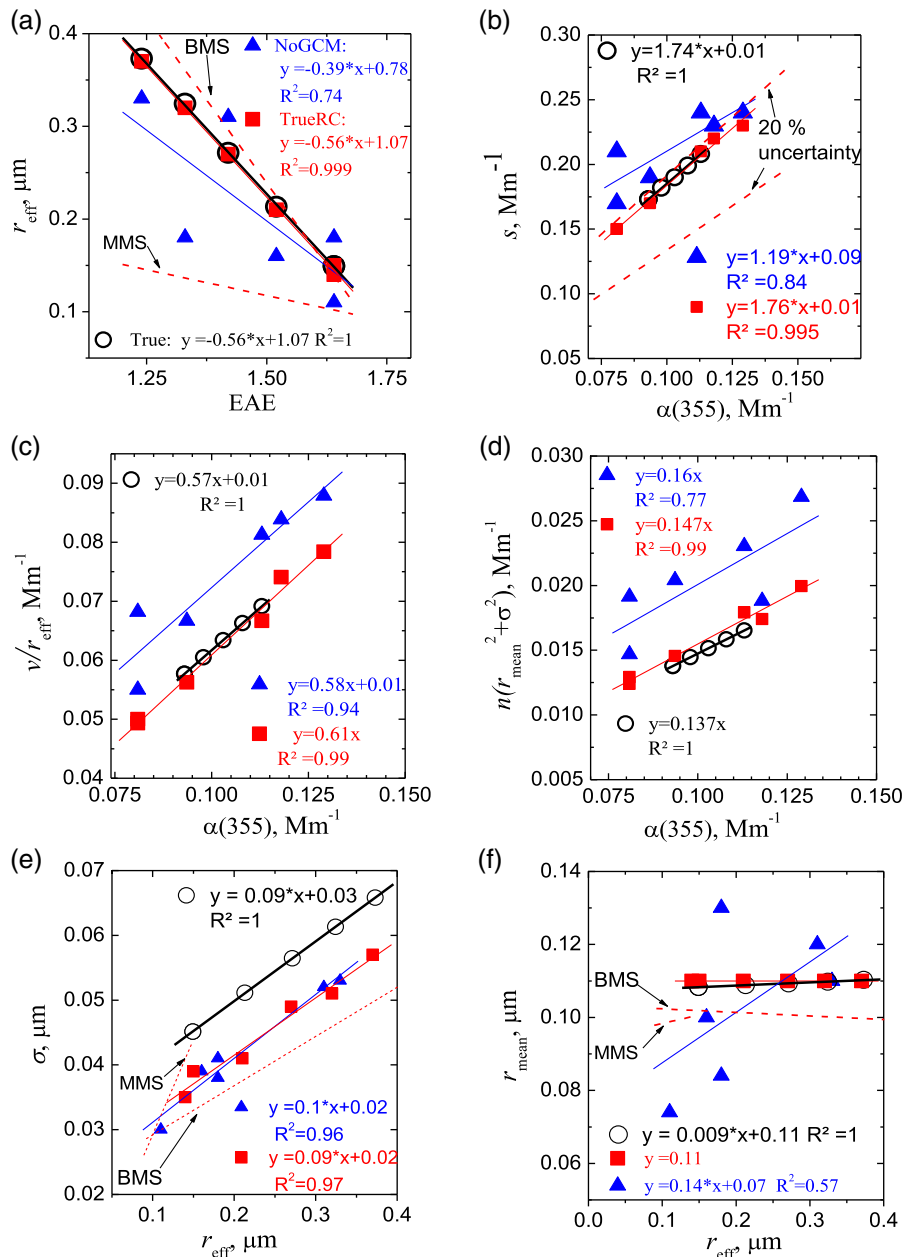


Fig. 2. Numerical example: vertical profile of type 1. The statistics are shown for the true (circles) and the retrieved values obtained for the case that we do not use GCM (triangles), the case that we use GCM with the true RCs (squares), and the case that we use GCM with incorrect RCs (dashed lines). We assumed that the PSD is either monomodal (MMS) or bimodal (BMS), and that $a_s = 1.6$. The solid lines describe the correlation trends according to the equation $y = ax + b$ and R^2 . The regression equations are given in the legends.

profiles of the synthetic (solid line) and distorted (symbols) optical data at 355, 532, and 1064 nm, as well as the Ångström exponents BAE for the wavelength pair 532/1064 nm and the EAE for the wavelength pair 355/532 nm. The profiles of the true PMPs [Figs. 1(b)–1(d)] and the true PSD [Fig. 1(e)] at bin $l = 4$ (at height 4 km) are shown as thick solid black curves.

The statistics and correlations for the true parameters are presented in Fig. 2 by open black circles and black lines in the same way as in Fig. 1(a) in [3]. The regression equations for all parameters are shown in the legends. In spite of the bimodality of the PSD, we see a very “strong” correlation

between the individual height bins of the profiles because the correlation coefficients are equal to $R^2 = 1$ in all cases.

We used distorted optical data profiles as input in our traditional regularization technique [6] to retrieve the solution space $F^{(l)}$, $l = 1, \dots, 6$. Each solution space $F^{(l)}$ contains about 1×10^5 individual solutions which are retrieved for 150 inversion windows that are located in the radius range from 0.03 to 10 μm . We used 20 equidistant values of the real part, i.e., $m_R \in [1.325; 1.8]$ and 30 equidistant values of the imaginary part, i.e., $m_I \in [0.0; 0.1]$ in the retrieval. We note that these ranges for the inversion windows and the CRI do not

create any constraints on the solution spaces, and we do not use them in practice.

In the first step we post-processed the solution spaces without using GCM (NoGCM) and without using any constraints for the parameters r_{eff} , n , $r \in [0.03; 10]$, $m_R \in [1.325; 1.8]$, and $m_l \in [0.0; 0.1]$ in the automated mode [1]. The results are shown as blue curves with triangles in Figs. 1(b)–1(e). The retrieved profiles of effective radius, surface-area and volume concentrations reproduce the true profiles quite well. The largest disagreement of 45% from the true value is observed for effective radius at bin $l = 5$ (at 5 km height). However, the quality of the retrieved profile of the number concentration is unsatisfactory. There are outliers with 100%–200% error. Just two profile bins, the ones at 3 and 6 km, respectively, coincide with the true values.

Figure 2 shows the statistics of the parameters retrieved with NoGCM (blue lines, triangle). We see that the results are not correlated. The correlation coefficient changes from $R^2 = 0.57$ for r_{mean} versus r_{eff} [Fig. 2(f)] to $R^2 = 0.96$ for σ versus r_{eff} [Fig. 2(e)]. The RC for surface-area concentration is $a_s = 1.19$ [Fig. 2(b)]. We find $a_n = 0.16$ for number concentration [Fig. 2(d)]. These values are outside the ranges defined by Eq. (2). The low correlation coefficients and values of the RCs that do not fulfill the condition in Eq. (2) mean that we deal with outliers in our results.

In the next step, we applied GCM in an attempt to suppress these outliers and postprocess the solution spaces for the case that the true RCs are known (TrueRC). Table 3 shows the RCs for all IPs. We applied the algorithm described in Section 2 and used $\delta p = 10\%$ – 25% . Figure 1 shows the results as red curves with squares. We see that all retrieved and true profiles match each other significantly better. The maximum error of 30% occurs for the number concentration at height bin $l = 2$ and $l = 4$.

The result of applying GCM is that we find strong correlations of $R^2 \geq 0.97$ for the IPs. Figure 2 shows these correlations as red lines with squares. We note that the actual RCs shown in the legends of Fig. 2 do not coincide with the true values (see the TrueRC of Table 3) after we used GCM for the postprocessing. Particularly the RC a_n for $n(r_{\text{mean}}^2 + \sigma^2)$ is 0.147, whereas the true value is 0.137. The true and actual mean widths of the RC b_σ are 0.03 and 0.02, respectively.

The disagreement between the actual (retrieved) and the true RCs can be explained, as we already discussed, by the fact that it is not mandatory that Eq. (12) is simultaneously fulfilled for the individual and the averaged $n(r_{\text{mean}}^2 + \sigma^2)$. If we reduce

the threshold δp to 1%–5% we could further stabilize the actual a_n (or b_σ), but we deal with measurement and mathematical errors. These two error sources do not allow us to find any individual solution $f_i^{(j)}(r)$, $i = 1, \dots, N_{\text{sol}}$, if δp is too small. This is the reason why we have to set $\delta p = 10\%$ – 25% in this case.

The true values of the RCs are not known in practice. If the RCs for our IPs, i.e., s , v/r_{eff} , and $n(r_{\text{mean}}^2 + \sigma^2)$ vary inside small intervals, i.e., $\pm 20\%$ in Eq. (2) the spread of the RCs for r_{eff} , r_{mean} , and σ are very wide, see for example the conditions in Eq. (32) in Ref. [3]. The choice of the RCs depends on many factors such as aerosol type, measurement case, and the experience of the data operator. To assess the capabilities of GCM we consider two different (opposite) retrieval strategies.

One strategy assumes that the PSD is monomodal (MMS). The other strategy assumes that the PSD is bimodal (BMS). The constraints for the RCs of s , v/r_{eff} , and $n(r_{\text{mean}}^2 + \sigma^2)$ can be described by the same equations for both strategies. The RCs are $a_s = 1.6$, $a_v = 1.6/3 \approx 0.53$, and $a_n = 1.6/(4\pi) \approx 0.127$ (see Table 3). The numbers are in the center of the respective ranges of values of these RCs. At this step (1st iteration) the profile of the number concentration is stabilized if we apply GCM for s , v/r_{eff} , and $n(r_{\text{mean}}^2 + \sigma^2)$ without using the constraints for r_{eff} , r_{mean} , and σ .

However, after the 1st iteration the profile of effective radius still contains the outliers that we also find if we do not apply GCM (case of NoGCM). As a result, the effective radius and EAE are only slightly correlated with $R^2 = 0.80$ (results are not shown here). To increase the correlation, we add a new constraint in the case of the effective radius. We use the RCs $a_r = -0.08$ and $b_r = 0.26$ in the case of MMS, see Eq. (33b) in Ref. [3]. We use $a_r = -0.65$ and $b_r = 1.2$ in case of BMS; see the equation that describes the regression line of the stars in Fig. 1(a) in Ref. [3].

We simultaneously use GCM with the constraint applied to r_{eff} , as described in the previous paragraph (1st iteration), to s , to v/r_{eff} , and to $n(r_{\text{mean}}^2 + \sigma^2)$. This step stabilizes (smoothes) the profile of r_{eff} . For the case of BMS, we find a high correlation $R^2 = 0.97$ between r_{eff} and \tilde{a}_α . In contrast, we find a lower correlation of $R^2 = 0.79$ between r_{eff} and \tilde{a} in the case of MMS (results are not shown here). Furthermore, the discrepancy between input and backcalculated EAE for the case of MMS starts to increase to values of 20% and more in bins $l = 3$ – 5 .

After the 2nd iteration in which we used the constraints simultaneously with regard to s , v/r_{eff} , $n(r_{\text{mean}}^2 + \sigma^2)$, and r_{eff} the profile of σ becomes well correlated with r_{eff} , and we obtain $R^2 > 0.92$ from both strategies (not shown here).

Table 3. Vertical Profile of Type 1: Regression Coefficients and Thresholds that Can Be Used as Constraints in the Next Version of the Postprocessing of the Solution Space Provided by the Automated, Unsupervised Inversion Algorithm

<i>P</i>	<i>a_p</i>	<i>b_p</i>	δp (%)	<i>a_p</i>	<i>b_p</i>	δp	<i>a_p</i>	<i>b_p</i>	δp
	True RC			MMS			BMS		
<i>s</i>	1.74	0.01	10	1.6	0	5%	1.6	0	5%
<i>v/r</i>	0.57	0.01	10	0.53	0	5%	0.56	0	5%
$n(r_{\text{mean}}^2 + \sigma^2)$	0.137	0.00	10	0.127	0	5%	0.5	0	5%
<i>r_{eff}</i>	-0.56	1.07	15	-0.08	0.26	0.05 μm	-0.65	1.2	0.05 μm
σ	0.09	0.03	25	0.31	0	25%	0.09	0.02	25%
<i>r_{mean}</i>	0.01	0.11	25	0.2	0.075	10%	0.01	0.10	20%

With regard to r_{mean} , the correlation with r_{eff} remains quite small, i.e., $R^2 < 0.15$ (results are not shown). To improve the correlation of r_{mean} and r_{eff} we use all available constraints for s , v/r_{eff} , $n(r_{\text{mean}}^2 + \sigma^2)$, r_{eff} , σ , and r_{mean} in the 3rd iteration. In addition, with regard to r_{mean} and σ , we use the RC values we obtained from the previous (2nd) iteration, i.e., we use $a_\sigma = 0.31$, $b_\sigma = 0$, $a_m = 0.2$, and $b_m = 0.075$ for MMS, and we use $a_\sigma = 0.09$, $b_\sigma = 0.02$, $a_m = 0.01$, and $b_m = 0.1$ for BMS (see Table 3). Figures 1 and 2 show the final results after the last, 3rd iteration as dashed curves. The results we obtain with GCM for the case of TrueRC and BMS are close to each other. In contrast, the results for MMS underestimate effective radius and volume concentration. We can define this deviation (underestimation) as the uncertainty of GCM.

We also apply the PA to the analysis of the fine mode PMPs and the fractions φ_g in the case of the distorted OD profiles (see Fig. 1, asterisks). This approach provides us with similar results that we obtain in Section 3.A for error free OD (see Table 1, red symbols). The effective radii for the distorted and the error free OD profiles coincide since the EAEs of both OD profiles are the same. The respective values for number, surface-area, and volume concentrations deviate from each other proportionally to the distortion of the OD.

In the case of the distorted OD, we find even less uncertainty for the CRI, i.e., in the vicinity of $\rho_{\text{LUT}} \approx 5\%$ at $l = 1$ we find that real parts are between 1.45 and 1.625. We find a similar uncertainty for the imaginary part, i.e., it varies between 0.001 and 0.02 regardless whether the OD are distorted or error free. Again, this uncertainty could be reduced to the resolution of the real part (± 0.025) in the LUT, and from 0.001 to 0.0075 for the imaginary part if we knew the effective radius to at least $\pm 0.01 \mu\text{m}$. With regard to the fine mode fractions at $l = 4$, the PA results in

$$\begin{aligned} \varphi_{\alpha(355)}^{(4)} &= 0.88, & \varphi_{\alpha(532)}^{(4)} &= 0.79, & \varphi_{\beta(355)}^{(4)} &= 0.80, \\ \varphi_{\beta(532)}^{(4)} &= 0.60, & \varphi_{\beta(1064)}^{(4)} &= 0.31. \end{aligned}$$

The discrepancy stays below 12% for $\varphi_{\beta(532)}$ (see Table 1 for reference at $l = 5$). Apparently, we can use this approach to estimate the effective radius of the coarse mode by taking into account the values of the LUT. As another result, the volume concentration of the coarse mode can be assessed, too, like what can be done for v_f by using Eq. (24c). In turn, if we know the total volume and surface-area concentrations ($p = p_f + p_c$) we may derive the total effective radius which is very important for the initial assumptions we need to make for the RCs a_r and b_r in Eq. (10). In that way, we obtain the RCs we use for the BMS. However, we do not further speculate on this result because we need to carry out further investigations regarding this approach of using the LUT; we want to consider all possible variations of the coarse mode effective radius from 1 to approximately $5 \mu\text{m}$, which to our opinion covers most of the potential aerosol scenarios.

We note that the PMPs derived from the MMS and PA described in Section 3.A converge (see Fig. 1). For example, the effective radius retrieved with both approaches is equal to $\sim 0.13 \mu\text{m}$ at $l = 1, \dots, 6$. All other parameters retrieved with MMS and PA at all heights are close to each other within the uncertainty of PA. We conclude that even if we make a

wrong assumption, in our numerical examples this is the case for MMS, we are still able to retrieve the PMPs of the fine mode of the size distribution with good accuracy.

In summary, we recommend the following strategy that allows us to select properly the RCs:

- A) The actual RCs a_s , a_v , and a_n belong to the intervals shown in Eq. (2).
- B) The preliminary RCs a_r and b_r are assessed with PA on the basis of MMS and/or BMS.
- C) The discrepancies of the particle extinction coefficients and/or the EAEs do not exceed the measurement errors significantly.
- D) The correlation coefficients are close to 1.
- E) The predefined number of individual solutions that need to be averaged and the actual number of individual solutions that fulfill all correlation constraints should be similar. It is not acceptable if there are height bins in which the solutions do not fulfill all constraints.

2. Numerical Example: Vertical Profile Type 2

We computed $N_L = 6$ optical sets of the $3\beta + 2\alpha$ type on the basis of Eq. (8). From these data sets we constructed OD profiles which are shown in Fig. 3(a) as solid curves. The true PMPs and CRIs are shown in Figs. 3(b)–3(d) as thick lines. For example, the real part and mean width do not change with height and are equal to $m_R = 1.5$ and $\sigma_G = 2.1$, respectively. The imaginary part increases with height, i.e., from $m_I = 0$ (height bins 1–2) to 0.01 (height bins 3–4) to 0.03 (height bin 5) to 0.05 (height bin 6).

With regard to the CRI values this example could for example describe a situation of a comparably clean marine boundary layer that contains a mixture of marine particles (sea salt) with low light-absorbing anthropogenic pollution and a strongly light-absorbing pollution layer aloft. We stress that the purpose of the synthetic profiles is to test the robustness of our post-processing scheme. We will show results of the analysis of experimental profiles in the third part of our study which is currently in preparation.

The true OD profiles were distorted with 15% extreme error and processed with the inversion algorithm. The circles in Fig. 4 show the statistics of the true data. The RCs of the regression equations (black lines) are within the respective ranges shown in Eqs. (2) and (4). For convenience, we summarize the RCs of the SOD in Table 4.

First, we classify the OD profiles according to Table (1) in Ref. [3]. The profiles contain OD sets from case #8 at $l = 1$, from case #1 at $l = 2, 3, 4$, and 6, and from case #4 at $l = 5$. That means we can approximate our results by monomodal PSDs and use our LUT for the PA. However, at $l = 1$ the particle size is comparably small, i.e., $r_{\text{eff}} \leq 0.29 \mu\text{m}$, and $m_I \leq 0.045$ and $m_R \geq 1.4$. In contrast, at $l = 2 - 6$ the cases #1 and #4 permit for the “full” ranges of effective radius ($r_{\text{eff}} \geq 0.2 \mu\text{m}$) and CRI ($m_I \geq 0.0$ and $m_R \geq 1.3$).

In contrast to the numerical example in Section 3.A (see layer $l = 6$) where BAE is equal to -0.24 the BAE is close to 2 in this example. In view of Eq. (33c) in Ref. [3], and Eq. (24) and $\varphi^{(1)} = 1$ we find (true values are in brackets):

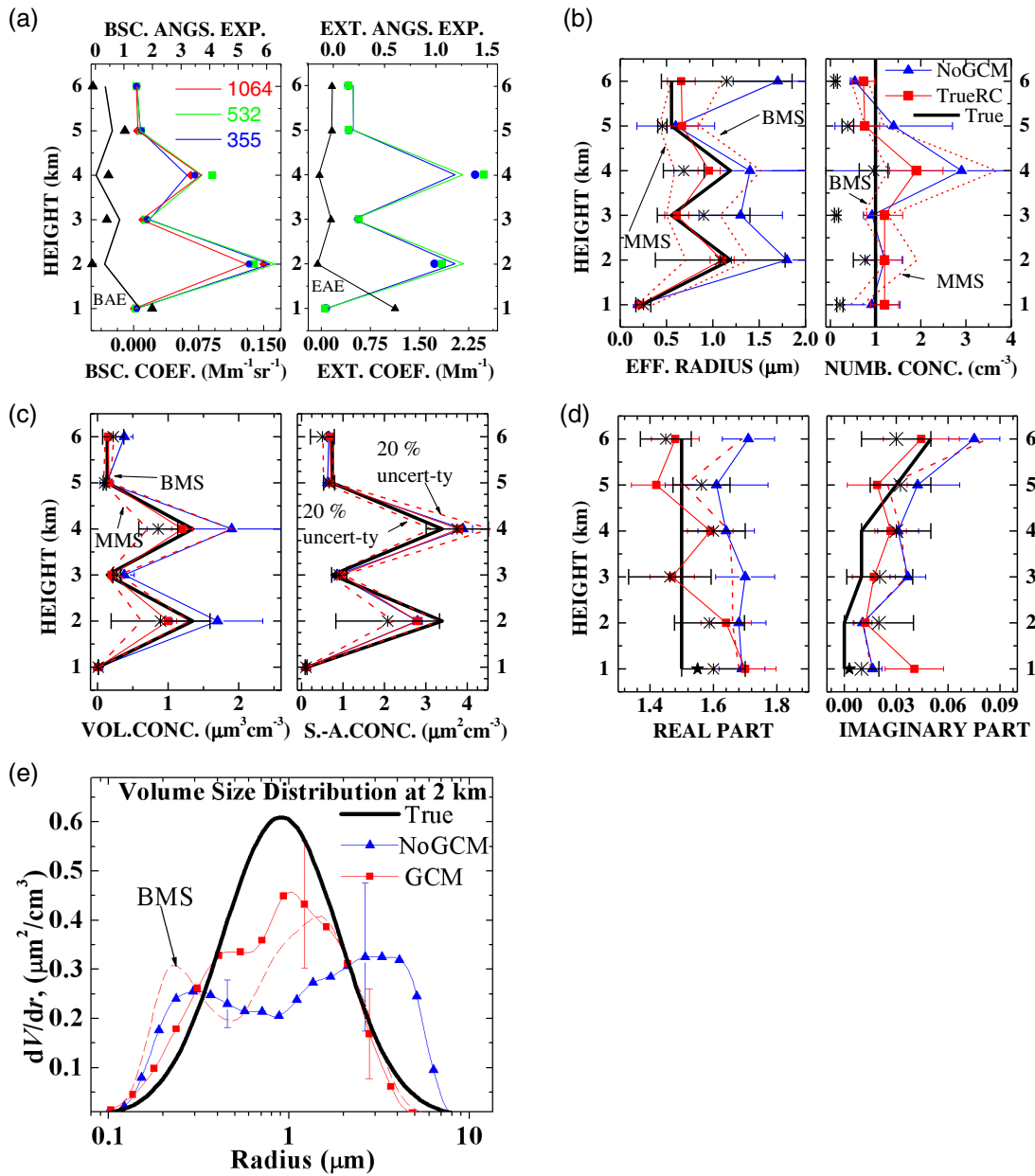


Fig. 3. Numerical example: vertical profile of type 2. The meaning of the lines, symbols, and colors is the same as in Fig. 1.

$$r_{\text{eff},f}^{(1)} \approx -0.193 \times 0.6 + 0.37$$

$$= 0.25 \pm 0.08 \mu\text{m} (0.24 \mu\text{m})$$

$$s_f^{(1)} \approx 1.6 \times 1 \times 0.07$$

$$= 0.11 \pm 20\% \mu\text{m}^2 \text{cm}^{-3} (0.136 \mu\text{m}^2 \text{cm}^{-3})$$

$$v_f^{(1)} \approx 0.11 \times 0.25/3$$

$$= 0.009 \pm 20\% \mu\text{m}^3 \text{cm}^{-3} (0.011 \mu\text{m}^3 \text{cm}^{-3})$$

$$n_{f,\text{min}}^{(1)} \approx 0.11 / (4 \times 3.14 \times 0.25^2) = 0.14 \text{cm}^{-3} (1 \text{cm}^{-3})$$

$$n_{f,\text{max}}^{(1)} \approx 0.11 / [(2 \times 3.14 \times 0.25^2)] = 0.28 \text{cm}^{-3} (1 \text{cm}^{-3})$$

Our PA gives us reasonable results within the respective uncertainty of δp , though the maximum value of the number

concentration does not include the true value. That result can be explained by the fact that the true PSD is not “narrow” ($\sigma_G = 2.1$), i.e., the PSD includes particles both of the fine and the coarse modes, and the numbers in Eq. (24d) are only a rough estimate at $l = 1$. Besides that, we neither took into account the 20% uncertainty of our estimation of s nor did we consider the measurement error of 15% of the extinction coefficients.

We carried out our PA of the CRI in the vicinity of $\rho_{\text{LUT}} \leq 15\%$ [see Eq. (7)]. The minimal discrepancy $\rho_{\text{LUT},\text{min}} \approx 7\%$ corresponds to the set of data in the LUT for which $m = 1.55 - i0.003$. In that vicinity we find another 118 sets with real parts $m_R \in [1.5; 1.7]$ and imaginary parts $m_I \in [0; 0.02]$. We again observe that in spite of erroneous

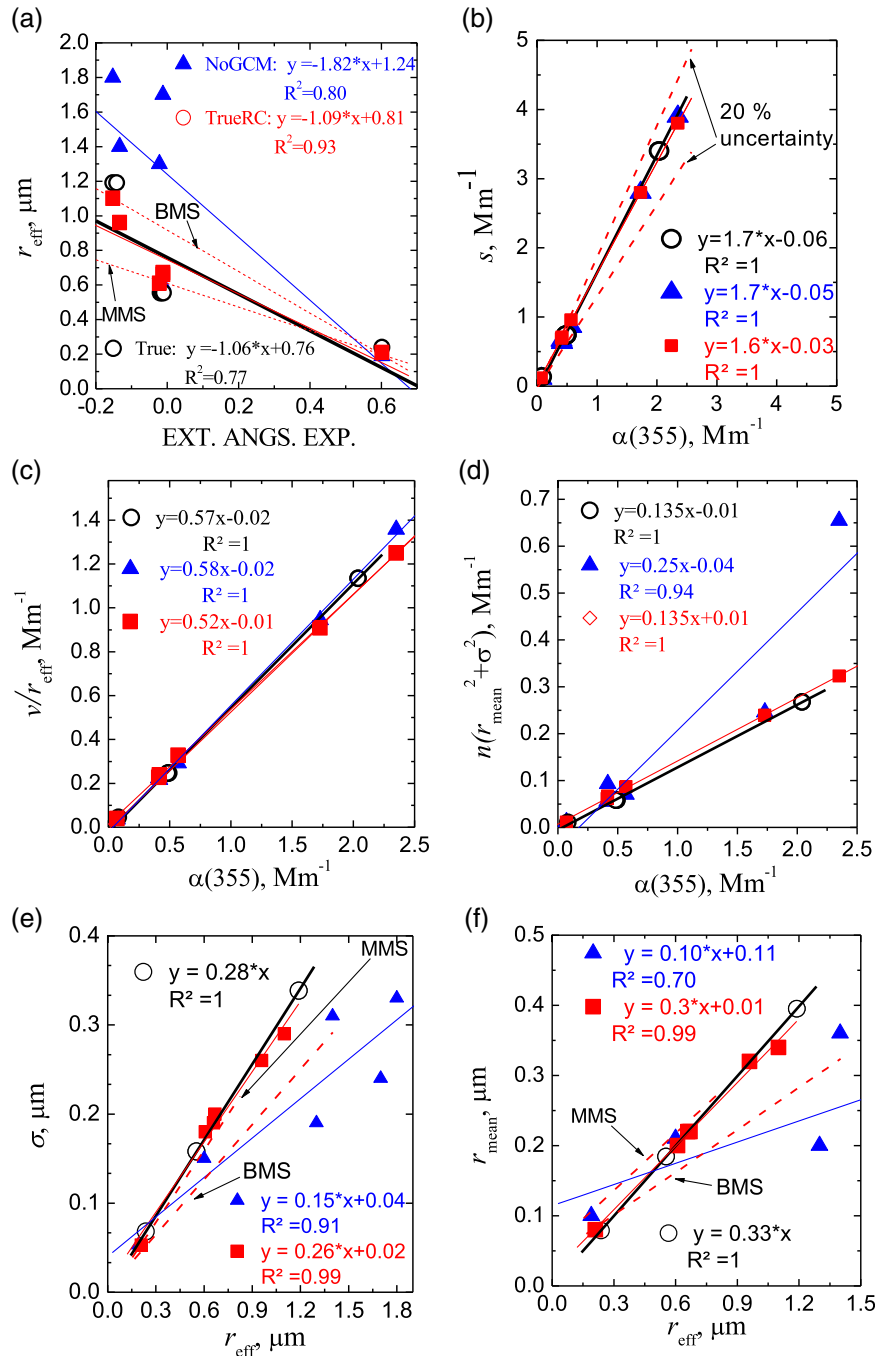


Fig. 4. Numerical example: vertical profile of type 2. The meaning of the lines, symbols, and colors is the same as in Fig. 2.

Table 4. Variations of the RCs of the SOD Bank

P	a_p	b_p
s	[1.33; 1.78]	[-0.09; 0.03]
v/r	[0.40; 0.60]	[-0.03; 0.01]
$n(r_{\text{mean}}^2 + \sigma^2)$	[0.11; 0.14]	[-0.05; 0.002]
r_{eff}	[-10; -0.038]	[0.1; 0.5]
σ	[0.21; 0.31]	0
r_{mean}	[0.17; 0.84]	0

OD the LUT data sets that have effective radii closest to the true value ($0.24 \mu\text{m}$) result in real parts $m_R = 1.5\text{--}1.55$ and imaginary parts $m_I = 0\text{--}0.003$.

Since the EAE is too low and changes from $\hat{a}_\alpha = -0.15$ to $\hat{a}_\alpha = -0.02$ in the upper height bins ($l = 2, \dots, 6$) the PA of the fine-mode fraction of the particle size distribution will result in an unacceptably high uncertainty δr_{eff} of effective radius [see Fig. 1(a) in [3] at $\hat{a}_\alpha \approx 0$]. Actually, this example is one of the most difficult scenarios for our retrieval scheme. However, we can further simplify the PA of the effective radius, not by using Eq. (24a) but by using the LUT in the same way

we use it to find the CRI. In this case, we rewrite the discrepancy of Eq. (7) as

$$\rho_{LUT} = \frac{1}{5} \sum_p \frac{|p_{LUT} - p_{\text{experiment}}|}{p_{LUT}},$$

$$p = \hat{a}_\alpha \hat{a}_{\beta(355), \beta(532)}, \hat{a}_{\beta(532), \beta(1064)},$$

$$\alpha(355)/\beta(355), \alpha(532)/\beta(532), \quad (32)$$

i.e., we take into consideration all lidar data products (we add $p = \hat{a}_{\beta(532), \beta(1064)}$) apart from our PA of effective radius ($p \neq r_{\text{eff}}$).

We used the simplified PA of the PMPs for the remaining heights ($l = 2-6$). The results are shown as asterisks in Fig. 3. We collected the LUT values of r_{eff} and m for discrepancies ρ_{LUT} up to 25% because of the high level of the OD distortion (15%). This distortion level can permit for even higher discrepancies of the lidar data products. If we derive effective radius from our LUT on the basis of Eq. (32), and furthermore derive surface-area concentration from Eq. (14) we find volume and number concentrations with Eqs. (24c), (24d), and (24e), respectively.

Unfortunately, the minimal discrepancies ρ_{LUT} at height 2 and 6 km ($l = 2, 6$) exceed 25%. In practice, it may mean that the error of the OD measurements is very high and/or the PSD is not monomodal. We tested the latter assumption and estimated the fine mode fractions as described in Section 3.A, i.e., we find

$$\begin{aligned} \varphi_{\alpha(355)}^{(2)} &= 0.5, & \varphi_{\alpha(532)}^{(2)} &= 0.5, & \varphi_{\beta(355)}^{(2)} &= 0.3, \\ \varphi_{\beta(532)}^{(2)} &= 0.2, & \varphi_{\beta(1064)}^{(2)} &= 0.06, & \varphi_{\alpha(355)}^{(6)} &= 0.5, \\ \varphi_{\alpha(532)}^{(6)} &= 0.5, & \varphi_{\beta(355)}^{(6)} &= 0.6, & \varphi_{\beta(532)}^{(6)} &= 0.5, \\ \varphi_{\beta(1064)}^{(6)} &= 0.4. \end{aligned}$$

As a result the minimum discrepancies defined by Eq. (31) decrease to $\rho_{LUT,c}^{(2)} = 6\%$ and $\rho_{LUT,c}^{(6)} = 9\%$, respectively. If we take into account the LUT values for effective radii $r_{\text{eff},f}$ and $r_{\text{eff},c}$, and the fraction $\varphi_{\alpha(355)}$, we can find the surface-area, volume and number concentrations of the fine and coarse modes [see Eqs. (24b)–(24e)], respectively.

Figure 3 (asterisk) shows that the PA provides the PMPs and their error bars and that the results include the true values of the PMPs. The only exception is number concentration. We already discussed that Eq. (24d) does not work properly for the estimation of $n_{f,\text{max}}$ if the investigated PSD includes large particles. The parameters \hat{a}_α and $\hat{a}_{\beta(532;1064)}$ are close to 0 and

even negative, simultaneously, which confirms our assumption. We note the large uncertainty of r_{eff} at 2 and 6 km. As a result, we obtain a large uncertainty of v in these same heights. Such large uncertainty arises because the erroneous OD can be reproduced by a bimodal PSD quite well. Our PA demonstrates this possibility. Besides that, we note that the $r_{\text{eff}}^{(4)}$ value at 4 km does not cover the true value, even if we include the uncertainty bar. We can explain this behavior by the fact that the minimum discrepancy $\rho_{LUT}^{(4)}$ is 20%. If we consider the discrepancy range from 20% to approximately 25% we can avoid using wrong parameter sets from the LUT, i.e., parameters for which the OD are too “far” off from the input OD. In that case, the derived set of LUT parameters contains an effective radius that varies only slightly. In such a case we would extend the discrepancy vicinity or consider the strategy of extracting the fine and coarse mode fractions.

We use the same approach in this example as in Section 3.B.1, i.e., we identify the solution space $F^{(l)}$, $l = 1, \dots, 6$, [6] and we could postprocess it with the automated, unsupervised software for the following situations

- A) GCM is not used, i.e., the case of NoGCM is applied,
- B) GCM is used for the case of True RC,
- C) GCM is used with incorrect RCs and the assumption that the PSD is monomodal (MMS) or bimodal (BMS, i.e., presence of coarse-mode particles), and $a_s = 1.6$, $a_v \approx 0.53$, and $a_n \approx 0.127$.

Case A: Figures 3(b)–3(e) show all retrieval results. We see that in the case of NoGCM (triangles) effective radius has large retrieval errors of up to 300% in three of the six height bins. The error of number concentration is up to 200% in the height bin $l = 4$. In particular, the outliers of r_{eff} are caused by a strong coarse mode which is absent in the true PSD, see the blue curve with triangles in Fig. 3(e).

Case B: The solution space can be significantly improved if we require that the correlation relationships are known and fulfilled (True RC). In that case, the retrieval errors of number concentration are below 90%. The error of effective radius is 20%. We obtained these results by using the RCs shown in Table 5 (True RC).

Case C: We assess the GCM uncertainty by assuming MMS and BMS. In both cases the RCs a_s , a_v , and a_n are located in the middle of the respective intervals (see Table 5). The RCs for effective radius are defined on the basis of the analysis of our LUT. Figure 1(a) in Ref. [3] shows that in the case of $\hat{a}_\alpha \in [0; 0.6]$ values for r_{eff} are minimal if we select

Table 5. Vertical Profile Type 2: Regression Coefficients and Thresholds that Can Be Used as Constraints in the Future Postprocessing of the Solution Space Provided by the Automated, Unsupervised Inversion Algorithm

p	a_p	b_p	δp	a_p	b_p	δp	a_p	b_p	δp
	True RC			MMS			BMS		
s	1.7	0	5%	1.6	0	5%	1.6	0	5%
v/r	0.57	0	5%	0.53	0	5%	0.56	0	5%
$n(r_{\text{mean}}^2 + \sigma^2)$	0.135	0	5%	0.127	0	5%	0.5	0	5%
r_{eff}	-1.06	0.76	0.3 μm	-0.6	0.6	0.2 μm	-1.25	0.9	0.5 μm
σ	0.285	0	10%	0.25	0.01	20%	0.21	0.0	20%
r_{mean}	0.332	0	20%	0.27	0.06	20%	0.2	0.05	20%

$a_r = -0.6$ and $b_r = 0.6$. The values for a_r and b_r fulfill the conditions required by MMS. On the contrary, we select the values $a_r = -1.25$ and $b_r = 0.9$ which allow us to maximize r_{eff} in the case of BMS. In that case, we can take into account the condition that the correlation trends in both strategies should intersect the vicinity of the coordinate point ($\hat{a}_\alpha^{(1)} = 0.6$; $r_{\text{eff},f}^{(1)} \approx 0.25 \mu\text{m}$) that we found from our PA. Furthermore, we make the thresholds δr_{eff} large enough (0.2–0.5 μm) since we deal with a wide spread of the effective radius when $\hat{a}_\alpha < 0.5$, see Table 5 (MMS, BMS).

After a few iterations we can estimate the RCs for our IPs σ and r_{mean} . The values are $a_\sigma = 0.25$, $b_\sigma = 0.01$, $a_m = 0.27$, and $b_m = 0.06$ for MMS. The values are $a_\sigma = 0.21$, $b_\sigma = 0.0$, $a_m = 0.2$, and $b_m = 0.05$ for BMS. Table 5 summarizes these numbers for the cases of MMS and BMS. The final results are shown in Fig. 3 as dashed curves.

The use of GCM causes quite large variations of effective radius and number concentration in the cases MMS and BMS. The uncertainties are 100% and 300% for effective radius and number concentration, respectively. That level of uncertainty could be suppressed if information about the modality (monomodal or bimodal) of the PSD is available. Unfortunately, the solution subspaces for the monomodal and the bimodal PSDs are acceptable with regard to $F^{(l)}$ (no matter if NoGCM or GCM is used) and to PA. For the moment there still is ambiguity of the results when both BAE and EAE are comparably small, e.g., if they are close to 0. In this case, erroneous OD cannot be identified as “monomodal” because of large measurement errors. At the same time, the “bimodal” interpretation yields a good discrepancy but a wrong coarse mode. One of the next steps in our research work could be to decrease the uncertainty of the retrieved effective radius that is provided by MMS and BMS.

Any further improvement of the GCM results depends on the data operator’s experience and *a priori* information. Nonetheless, we stress that even at this stage of our work GCM already smoothes the profiles of r_{eff} and v and thus provides us with significantly better consistency between successive aerosol layers compared with what we obtain from the method of NoGCM.

Figure 4 presents the statistics for the vertical profile of type 2 for all results derived with NoGCM (triangle), TrueRC (square), and MMS and BMS (dashed line). The legend in Fig. 4 also shows the regression equations that we obtained after postprocessing the solution space $F^{(l)}$ in the case of NoGCM and the case of True RC. We also show the regression equations for the true data (circles).

We see that the IPs s and v/r_{eff} that are retrieved with NoGCM and TrueRC are linearly correlated with the particle extinction coefficient [Figs. 4(b) and 4(c)]. The results for MMS and BMS show the same behavior (not shown here). We find outliers in the profiles of n if we use NoGCM (see Fig. 3), and as a result the actual RC $a_n = 0.25$ [Fig. 4(d), triangle] is outside the interval defined by Eq. (17) in Ref. [3]. The statistics for NoGCM for the IPs r_{eff} , σ , and r_{mean} also are “de-correlated,” i.e., R^2 is low. The numbers decrease from 0.91 to 0.7 [see Figs. 4(a), 4(e), and 4(f)].

If we employ the method TrueRC and use the RCs from Table 5 we find the following numbers for the RCs:

$a_r = 1.09$ ($R^2 = 0.93$), $a_n = 0.135$ ($R^2 = 1$), $a_\sigma = 0.26$ ($R^2 = 0.99$), and $a_m = 0.3$ ($R^2 = 0.99$), see the squares in Figs. 4(a) and 4(d)–4(f). If we use MMS and BMS (dashed lines) we obtain RCs close to the ones shown in Table 5. The correlation coefficients are $R^2 > 0.98$. We note once more that the measurement errors are 15%.

We complete the analysis of both numerical examples with 15% measurement error by considering the results of the CRI retrieval. If we know the true RCs, the real and imaginary parts converge significantly better to the true values compared to the convergence we obtain from the method of NoGCM. The results are shown in Figs. 1(d) and 3(d) as red solid lines with squares. The errors of m_R and m_I are on average approximately 0.05–0.10 and $i0.010 - i0.015$, respectively. There is only one outlier at 1 km (see vertical profile type 2). The analysis of the solution space at this height shows that it contains individual solutions with $m_R > 1.7$ and $m_I > 0.05$.

The results for MMS and BMS almost coincide [see Figs. 1(d) and 3(d), red dashed line] and overestimate the true CRI (bias). The results for the CRI are also close to the NoGCM results [see Figs. 1(d) and 3(d), blue solid line with triangles]. Therefore, we cannot expect an improvement of the CRI retrieval if we apply our new GCM only, compared to our traditional approach [6]. However, we can use GCM and combine it with our PA. We obtain $r_{\text{eff}} = 0.265 \mu\text{m}$ as a mean value from the two values we find from MMS and BMS at the lowest height bin. In that height, a monomodal PSD is highly likely. We obtain from the LUT values of CRI and effective radius ($m = 1.55 - i0.003$ and $r_{\text{eff}} = 0.24 \mu\text{m}$), if we use a discrepancy [Eq. (7)] that is equal to $\rho_{\text{LUT},\text{min}} \approx 8\%$. If we extend the value of the discrepancy to $\rho_{\text{LUT}} = 16\%$ (twice as large as the minimum value) the LUT values for the CRI and the effective radius stay below 1.6 (m_R), 0.01 (m_I), and 0.24 μm (r_{eff}). The results we obtain for the CRI from the combination of PA and GCM are shown as stars in Figs. 1(d) and 3(d). We are considering the option of including this new method in the next generation version of the automated inversion software that we are developing for NASA Langley’s HSRL-2 instrument.

4. CONCLUSION

We presented the novel method of GCM which allows us to stabilize the solution space of particle microphysical parameters (PMPs) that we retrieve with our data inversion algorithm from profiles of optical data taken with multiwavelength Raman/HSRL (3 backscatter + 2 extinction) lidar. GCM uses correlation relationships between particle bulk parameters and optical properties.

These relationships allow us to use additional constraints during the postprocessing of the solution spaces. As a result, the IPs [$s, v, n(r_{\text{mean}}^2 + \sigma^2)$] can be estimated with an uncertainty that does not exceed the measurement error of the extinction coefficient.

We specially designed the PA method that allows us to obtain from a LUT a first-estimate solution space of the PMPs. We obtain a robust initial estimation of the RCs on the basis of the correlation analysis presented in Ref. [3]. PA itself provides us with estimated values of the PMPs of the fine-mode

fraction of particle size distributions. The accuracy that we obtain for the PMPs with this approach is not worse than the accuracy for the PMPs obtained with our traditional inversion method. However, one of our next steps in fact is that we use the results from PA as *a priori* information in the second step of the data analysis procedure, i.e., we apply this PA in combination with the traditional inversion algorithm. We are of the opinion that the combination of the two methods will further improve our inversion data products, most notably the complex refractive index. That step of combining the two methods will allow us to derive trustworthy profiles of absorption coefficients in future versions of our software development.

The PA method can be generalized for the estimation of coarse-mode parameters if we extend our LUT by data that describe particle size distributions with effective radii exceeding 3 μm .

Our comparisons between the results obtained with GCM and our traditional inversion-with-regularization technique shows that the methodology presented in this contribution leads to a significant stabilization of the retrieved profiles of aerosol microphysical properties. That means the uncertainty of

- surface-area concentration does not exceed the sum of mathematical ($\pm 20\%$) and measurement ($\pm 15\%$ for modern lidar systems) errors, or 35% in total;
- effective radius is $\pm 0.07 \mu\text{m}$ for fine mode particles and about 100% for particle size distributions composed of fine (submicron) and coarse (supermicron) mode particles;
- volume concentration is defined by the sum of the uncertainty of surface-area concentration and the uncertainty of effective radius;
- the uncertainty of number concentration is less than 100% for any particle radius domain from 0.03 to 10 μm .

Aside from these results, we find that the uncertainties of the real and imaginary parts of the CRI of monomodal PSDs can be restricted to ± 0.1 and ± 0.01 on the domains [1.3;1.8] and [0;0.1], respectively, if we combine PA and GCM. This will allow us in the future to investigate bimodal particle size distributions in the sense that we can fully analyze the fine mode fraction of particle size distributions without making critical assumptions on properties of the coarse mode fraction. We can thus focus our future work steps on developing a light-scattering model for coarse mode particles which to a large part consist of nonspherical (dust) particles and for which to date, to the best of our knowledge, there is no practical model

that can describe light-scattering at 180°-scattering direction (backscatter).

We will continue our research work with numerical simulations and the analysis of case studies to further assess the potential of this new approach.

We are currently preparing part 3 of our series of papers. There, we will use experimental data and test our method for retrieving PMPs from optical profiles measured with lidar.

Funding. NASA Langley Research Center; University of Hertfordshire (UH); Science Systems and Applications.

REFERENCES

1. D. Müller, C. A. Hostetler, R. A. Ferrare, S. P. Burton, E. Chemyakin, A. Kolgotin, J. W. Hair, A. L. Cook, D. B. Harper, R. R. Rogers, R. W. Hare, C. S. Cleckner, M. D. Obland, J. Tomlinson, L. K. Berg, and B. Schmid, "Airborne multiwavelength high spectral resolution lidar (HSRL-2) observations during TCAP 2012: vertical profiles of optical and microphysical properties of a smoke/urban haze plume over the northeastern coast of the US," *Atmos. Meas. Tech.* **7**, 3487–3496 (2014).
2. P. Sawamura, D. Müller, S. Burton, E. Chemyakin, C. Hostetler, R. Ferrare, A. Kolgotin, L. Ziemba, A. Beyersdorf, and B. Anderson, "Comparison of aerosol optical and microphysical retrievals from HSRL-2 and in-situ measurements during DISCOVER-AQ 2013 (California and Texas)," in *International Laser Radar Conference*, July 2015, paper PS-C1-14.
3. A. Kolgotin, D. Müller, E. Chemyakin, and A. Romanov, "Improved identification of the solution space of aerosol microphysical properties derived from the inversion of profiles of lidar optical data, part 1: theory," *Appl. Opt.* **55**, 9839–9849 (2016).
4. C. F. Bohren and D. R. Huffman, eds., *Absorption and Scattering of Light by Small Particles* (Wiley, 1983).
5. D. Müller, U. Wandinger, and A. Ansmann, "Microphysical particle parameters from extinction and backscatter lidar data by inversion with regularization: theory," *Appl. Opt.* **38**, 2346–2357 (1999).
6. I. Veselovskii, A. Kolgotin, V. Griaznov, D. Müller, U. Wandinger, and D. Whiteman, "Inversion with regularization for the retrieval of tropospheric aerosol parameters from multiwavelength lidar sounding," *Appl. Opt.* **41**, 3685–3699 (2002).
7. C. Böckmann, I. Miranova, D. Müller, L. Scheidenbach, and R. Nessler, "Microphysical aerosol parameters from multiwavelength lidar," *J. Opt. Soc. Am. A* **22**, 518–528 (2005).
8. A. Kolgotin and D. Müller, "Theory of inversion with two-dimensional regularization: profiles of microphysical particle properties derived from multiwavelength lidar measurements," *Appl. Opt.* **47**, 4472–4490 (2008).
9. D. Müller, A. Kolgotin, I. Mattis, A. Petzold, and A. Stohl, "Vertical profiles of microphysical particle properties derived from inversion with two-dimensional regularization of multiwavelength Raman lidar data: experiment," *Appl. Opt.* **50**, 2069–2079 (2011).