Development and application of an automatic lidar-based aerosol typing algorithm

Der Fakultät für Physik und Geowissenschaften
der Universität Leipzig
genehmigte

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

D O C T O R   R E R U M   N A T U R A L I U M
Dr. rer. nat.

vorgelegt

von Athina Avgousta Floutsi, M.Sc. Applied Earth Sciences (Geoscience and Remote Sensing)
geboren am 21. August 1992 in Korinth

Gutachter: Prof. Dr. Andreas Macke
Prof. Dr. Dimitrios Balis

Tag der Verleihung 26. August 2022

Bibliographische Beschreibung:

Floutsi, Athina Avgousta
Entwicklung und Anwendung eines automatischen Lidar-basierten Algorithmus zur Aerosol-Typisierung
Universität Leipzig, Dissertation
161 Seiten, 143 Literaturzitate, 68 Abbildungen, 18 Tabellen

Referat:

Im Rahmen der vorliegenden Arbeit wurde eine Methode zur Aerosoltypisierung entwickelt, die sowohl für bodengebundene als auch für weltraumgestützte Lidarsysteme anwendbar ist. Das Aerosol-Typisierungsschema wurde auf der Grundlage der "Optimal Estimation Methode"(OEM) entwickelt und ermöglicht die Identifizierung von bis zu vier verschiedenen Aerosolkomponenten in einem Aerosolgemisch, sowie die Quantifizierung ihres Beitrags zum Aerosolgemisch in Bezug auf das relative Volumen. Von den vier Aerosolkomponenten wird angenommen, dass sie physikalisch voneinander getrennt und extern gemischt sind. Zwei Aerosolkomponenten repräsentieren jeweils absorbierende (FSA) und weniger absorbierende (FSNA) feine Partikel, die verbleibenden zwei Komponenten repräsentieren sphärische (CS) und nicht-sphärische (CNS) grobe Partikel. Diese vier Komponenten repräsentieren die am häufigsten beobachteten Aerosolarten in der Atmosphäre: Rauch, Luftverschmutzung, Meersalz und Wüstenstaub.

Die von Lidargrössen abgeleiteten optischen Parameter, die in diesem Typisierungsschema verwendet werden, sind das Lidar-Verhältnis und das lineare Depolarisationsverhältnis der Partikel bei zwei unterschiedlichen Wellenlängen (355 und 532 nm), die "color ratio"der Rückstreuung (532/1064 nm) und der Ångströmexponent (355/532 nm). Diese intensiven optischen Eigenschaften können auf unterschiedliche Weise kombiniert werden, was die Methode flexibel und somit anwendbar auf Lidar-Systeme mit unterschiedlichen Konfigurationen (z. B. mit einer oder mehreren Wellenlängen) macht.

Die Funktionalität des Schemas wurde mittels Anwendung auf Fallstudien mit bekanntem und unbekanntem Aerosolbedingungen überprüft. Der Algorithmus wurde auch auf einen Langzeitdatensatz angewendet, um eine saisonale Charakterisierung der Aerosolverteilung zu erhalten. Es konnte gezeigt werden, dass OEM zur Aerosoltypisierung eingesetzt werden kann. Des Weiteren wurde gezeigt, dass sie zur Unterstützung der bodengestützten Validierung der Datenprodukte und Algorithmen der geplanten Satellitenmission Earth-CARE verwendet werden kann.

Bibliographic Description:

Floutsi, Athina Avgousta
Development and application of an automatic lidar-based aerosol typing algorithm
University of Leipzig, Dissertation
161 Pages, 143 References, 68 Figures, 18 Tables

Abstract:
Within the framework of the present work, an aerosol typing methodology applicable to both ground-based and spaceborne lidar systems has been developed. The novel aerosol typing scheme was developed based on the optimal estimation method (OEM) and allows the identification of up to four different aerosol components present in an aerosol mixture as well as the quantification of their contribution to the aerosol mixture in terms of relative volume. The four aerosol components considered in this typing scheme represent the most commonly observed aerosol particles in nature and are assumed to be physically separated from each other and, therefore, can create external mixtures. Two components represent fine-mode particles, absorbing (FSA) and less absorbing (FSNA), and the remaining two aerosol components represent coarse-mode particles, spherical (CS) and non-spherical (CNS). These components can adequately represent the most frequently observed aerosol types in the atmosphere: combustion- and pollution-related aerosol, sea salt and desert dust, respectively.

The lidar-derived optical parameters used in this typing scheme are the lidar ratio and the particle linear depolarization ratio at two distinctive wavelengths (355 and 532 nm), the backscatter-related color ratio (for the wavelength pair of 532/1064 nm) and the extinction-related Ångström exponent (for the wavelength pair of 355/532 nm). These intensive optical properties can be combined in different ways making the methodology flexible, allowing thus its application to lidar systems with different configurations (e.g., single wavelength or multiwavelength).

The functionality of the typing scheme was demonstrated by its application to case studies of known aerosol conditions as well as to cases of non-characterized aerosol load. The algorithm was also applied to a long-term dataset to provide a seasonal characterization of the aerosol situation over Haifa, Israel. It was shown that the OEM is an effective methodology that can be also applied for aerosol typing purposes, and that it can be used to support the ground-based validation efforts of EarthCARE's products and algorithms.

Contents

CONTENTS

Chapter 1

Introduction

The influence of aerosol on the Earth's climate, by direct, indirect and semi-direct effects, has been studied for several decades. Aerosol particles influence directly the Earth's radiation balance by scattering and absorbing shortwave and longwave radiation (McCormick and Ludwig, 1967) and indirectly by acting as cloud condensation nuclei (CCN) and ice-nucleating particles (INP), altering thus cloud radiative properties and lifetime (Twomey, 1977; Albrecht, 1989). The direct absorption of radiation by aerosol can induce changes in the atmosphere's temperature structure and thus influence cloud formation and convection patterns, which is known as the semi-direct effect (Hansen et al., 1997; Ackerman et al., 2000).

The quantification of these effects continues to be a major scientific topic, due to uncertainties induced on the Earth's radiation budget calculations that greatly exceed uncertainties of any other climate forcing agent (Kaufman et al., 2002; Vardavas and Taylor, 2011; IPCC, 2013). The greatest part of uncertainty in predictions of radiative forcing is attributed to limited knowledge of spatial and temporal distribution of aerosols, their physico-chemical properties and the processes they are involved in (e.g., aerosol-cloud interactions; Boucher et al. (2013)). It becomes obvious that a reliable estimation of the aerosol effects on climate requires detailed information on the aerosol conditions, such as the aerosol vertical distribution, and the optical, microphysical and chemical properties, which are a function of the aerosol type and age.

Continuous aerosol monitoring on global scales can be achieved via coordinated research strategies and infrastructures. Observational ground-based networks, satellites, aircraft and ships are platforms that can facilitate both in-situ and remote-sensing aerosol measurements. The Aerosol Robotic Network (AERONET) established by the National Aeronautics and Space Administration (NASA) and PHOTONS (French National Centre for Scientific Research, CNRS) is a pioneering and one of the first and principal networks of ground-based remote-sensing aerosol observations (Holben et al., 1998). AERONET sites

are distributed across different locations around the world, covering thus diverse aerosol regimes. AERONET provides high-quality data of spectral aerosol optical depth (AOD), inversion products (i.e., microphysical and radiative properties), and precipitable water (Holben et al., 2001). However, sun/sky and lunar photometer measurements can only provide information about the whole atmospheric column.

The vertical aerosol distribution is a key factor for evaluation of the aerosol direct radiative effect, since the lifetime and climate response of aerosols are not only type but also altitude dependent (Hansen et al., 1997) and thus columnar measurements alone are not sufficient. Light detection and ranging (lidar) is an active remote-sensing technique that allows for vertically resolved aerosol measurements. Lidar instruments emit light preferably at different wavelengths, which interacts with molecules and aerosol particles in the atmosphere. The detection of the backscattered light, along with the polarization state of the backscattered light, provides information on the optical and microphysical properties of different aerosol types along with their vertical distribution (Ansmann and Müller, 2005). Several lidar networks have emerged around the globe in the last decades. The NASA Micro-Pulse Lidar Network (MPLNET) has sites mainly across North America (Welton et al., 2002), the Asian Dust and Aerosol Lidar Observation Network (AD-Net) is expanding in East Asia (Sugimoto et al., 2014), and most parts of Europe are covered by the European Aerosol Research Lidar Network (EARLINET; Pappalardo et al. (2014)). These networks operate different lidar systems and therefore have different capabilities. For instance, MPLNET is equipped with an elastic-backscatter lidar that operates at 532 nm. AD-Net is mostly equipped with Raman lidars operational mainly at 532 and 1064 nm. Similarly, EARLINET stations are mostly equipped with multiwavelength (355, 532 and 1064 nm) Raman lidars. Elastic lidar systems have limited capabilities (e.g., with respect to aerosol typing) compared to Raman lidar systems, since two physical quantities, the particle backscatter and extinction coefficients, need to be determined by one measured quantity i.e., the elastic lidar return (Ansmann and Müller, 2005). On the other hand, with the Raman lidar method (Ansmann et al., 1992) the backscatter and extinction coefficients can be determined independently. In addition, spectral information provided by multi-wavelength lidars advance aerosol typing and can be used to derive microphysical particle properties (via inverse modelling). However, despite the widespread of the stations of the aforementioned networks, spatial coverage by ground-based systems is limited.

The limitation of spatial coverage has been partly overcome since the 1980s by space-borne passive sensors, such as the early Advanced Very High Resolution Radiometer and the Total Ozone Mapping Spectrometer (AVHRR and TOMS, respectively). These sensors were primarily designed for other purposes, but provided valuable information and advanced the understanding of the aerosol global distribution (Mishchenko et al., 1999; Herman et al.,

1997; Torres et al., 2002). However, they provided little or no information regarding the vertical distribution of aerosols. Since then, the situation has been greatly improved with the development of more sophisticated and accurate sensors as well as satellite aerosol retrievals. Such a sensor is the Moderate Resolution Imaging Spectroradiometer (MODIS) on board NASA's twin polar-orbiting satellites Terra and Aqua, which, since February 2000, perform near-global observations of atmospheric aerosols, providing one of the most reliable satellite datasets of AOD over ocean and land (Bréon et al., 2011; Nabat et al., 2013).

The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) on board NASA's Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite was the first polarization lidar in space to provide global long-term atmospheric measurements (Winker et al., 2009). For more than 15 years, CALIOP has measured vertical profiles of attenuated backscatter at visible and near infrared wavelengths, along with depolarization in the visible channel. However, as an elastic-backscatter lidar, CALIOP is not able to perform direct extinction measurements. To enable the calculation of the extinction from the backscatter signals, the aerosol extinction-to-backscatter ratio (or lidar ratio) needs to be assumed. Since the lidar ratio depends on the microphysical and chemical properties of the scattering particles present in the atmosphere, an aerosol typing scheme was developed for CALIPSO (Omar et al., 2005, 2009; Kim et al., 2018). CALIPSO's typing algorithm is able to classify and assign typical lidar ratio values to six different aerosol types, which are desert dust, biomass burning, background (clean continental), polluted continental, polluted dust and clean marine (more details regarding the typing scheme are discussed in Ch. 2). It becomes clear that the goodness of the extinction retrievals is always dependent on this typing scheme, even though several quality control procedures are in place (Winker et al., 2009).

Recently (2018), the European Space Agency (ESA) launched the wind lidar mission Aeolus (Stoffelen et al., 2006). The satellite is equipped with a high-spectral-resolution lidar (HSRL; Wandinger (1998)), the Atmospheric Laser Doppler Instrument (ALADIN). ALADIN acquires one-directional horizontal tropospheric and stratospheric wind profiles (mainly west-east), aiming to improve weather forecast, advance atmospheric dynamics research and evaluate climate models (Stoffelen et al., 2006; Straume et al., 2020). Apart from being the first wind lidar in space, ALADIN is also the first lidar in space able to measure extinction coefficients directly and even though aerosol monitoring is not the main scope of Aeolus (coarse resolution), aerosol and cloud optical properties are retrieved with ALADIN as spin-off products (Ansmann et al., 2007; Flamant et al., 2008; Straume et al., 2020). Recently, Baars et al. (2021) demonstrated that spin-off products such as the backscatter coefficient, the extinction coefficient and the lidar ratio derived from ALADIN, even though preliminary (due to the ongoing algorithm improvements and quality assurance updates)

agree very well with ground-based lidar measurements. This is a great step towards the harmonization of multiple satellite instruments, especially when keeping in mind the upcoming Cloud, Aerosol and Radiation Explorer (EarthCARE) joint mission of ESA and the Japanese Aerospace Exploration Agency (JAXA), scheduled for launch in 2023.

EarthCARE's payload consists of four instruments: an ATmospheric LIDar (ATLID), a Cloud Profiling Radar (CPR), a Multi-Spectral Imager (MSI) and a Broad-Band Radiometer (BBR) (Illingworth et al., 2015). ATLID is a 355-nm HSRL that will provide direct cloud and aerosol profile measurements of backscatter and extinction coefficients. Furthermore, ATLID is able to measure the depolarization ratio of the atmospheric particles – an ideal parameter for aerosol typing – as well as ice particle characteristics (Illingworth et al., 2015; do Carmo et al., 2021). The primary goal of EarthCARE is radiative closure, which is aimed to be achieved in a synergistic approach from the two active and two passive instruments. One key element for this goal is a proper aerosol typing scheme, to calculate the aerosol's radiative properties. For this purpose, the Hybrid End-To-End Aerosol Classification (HETEAC) model has been developed (Wandinger et al., 2016a). As the name indicates, the HETEAC model delivers the required theoretical description of aerosol microphysics that is consistent with experimentally derived optical properties (hybrid approach) to close the loop from observations and aerosol microphysics to radiative properties (end-to-end approach). Since the contents of this dissertation are related to the developments of HETEAC, a more complete description is provided in Ch. 2.

Once in orbit, EarthCARE's HSRL will provide accurate measurements of the extinction and backscatter coefficients, which will continue and improve the already existing record of global aerosol lidar measurements initiated by CALIPSO. With HETEAC and a synergistic combination of HSRL and MSI measurements, aerosol classification and quantification of the radiative impact will be available at global and regional scales. As for every satellite mission, ground-based remote-sensing measurements are essential for EarthCARE's product and algorithm validation. To achieve this goal and with a specific focus on the aerosol classification and the radiative impact quantification, an aerosol typing methodology applicable to both ground-based and spaceborne lidar systems has been developed and is presented in this dissertation. The methodology is based on the optimal estimation method (Rodgers, 2000) and allows the identification of aerosol mixtures consisting of up to four different aerosol components from lidar measurements. The methodology is rather flexible, allowing thus its application to lidar systems with different configurations (e.g., single wavelength or multiwavelength). Microphysical and optical properties of the predefined aerosol components are in accordance with HETEAC, thus permitting direct comparisons, algorithm harmonization and further support to the validation activities for EarthCARE.

In Chapter 2, an overview of typing techniques commonly used in previous studies is

provided. Chapter 3 presents an introduction to lidar-derived aerosol-type-dependent parameters (intensive optical properties), along with a collection of intensive aerosol properties from ground-based lidar measurements that serves as a baseline for the development of the aerosol typing scheme. Chapter 4 describes in detail the principles of the newly developed typing algorithm. Chapter 5 is devoted to the application and intensive discussion of the algorithm in selected case studies, while Chapter 6 demonstrates the application to long-term lidar observations obtained from field campaigns. The summary in Chapter 7 illustrates the comprehensive results of this dissertation and concludes with an outlook.

Chapter 2

Lidar-based aerosol typing

This chapter provides an overview of recent developments in the research field of aerosol typing based on intensive (i.e., concentration-independent) aerosol properties derived from lidar. Traditionally, different intensive optical parameters can be combined to classify major aerosol types and the mixtures that they are involved in (Sasano and Browell, 1989; Sugimoto et al., 2002; Ansmann et al., 2002; Müller et al., 2002, 2003, 2005; Mattis et al., 2002, 2004; Tesche et al., 2009b, 2011a; Groß et al., 2011, 2013; Weinzierl et al., 2011; Kumar et al., 2018; Papagiannopoulos et al., 2018; Nicolae et al., 2018). These parameters are the lidar ratio (or the extinction-to-backscatter ratio), the particle linear depolarization ratio, the Ångström exponent and the color ratio (ratio of backscatter coefficients at different wavelengths). The aforementioned intensive optical parameters reveal information about the size, shape and absorption efficiency of the aerosol particles, allowing thus to identify and type them (more details are provided in Ch. 3). In the following, the most important schemes for this study are discussed. Specific focus is given to HETEAC, as it is the basis of the developments presented in this work.

2.1 Dust and non-dust separation – POLIPHON

During the SAMUM–2 campaign (Ansmann et al., 2011a), several lofted aerosol plumes containing mixtures of desert dust and biomass-burning aerosol were observed with lidars at Praia, Cape Verde (Tesche et al., 2009b). This complex mixture was identified mainly by depolarization values that were too high for pure smoke layers and too low for pure desert-dust-containing layers. To allow the assessment of the radiative effects of this aerosol mixture, a methodology for separating the optical properties of desert dust and biomass-burning particles as a function of height was developed by Tesche et al. (2009b).

The methodology was initially applied to separate dust and smoke, but can be used to separate dust from any other non-dust aerosol as well. Starting point is the volume depo-

larization ratio δ_{vol}, which includes contributions of both molecules and aerosol particles,

$$\delta_{\text{vol}} = \frac{\beta_{\text{m}}^{\perp} + \beta_{\text{p}}^{\perp}}{\beta_{\text{m}}^{\parallel} + \beta_{\text{p}}^{\parallel}}, \tag{2.1}$$

where β is the backscatter coefficient and subscripts m and p denote molecules and particles, respectively. For simplicity, the wavelength dependence (λ) has been omitted, but of course β_{m}, β_{p} and δ_{vol} are functions of the wavelength. Then, the molecular and particle contributions are separated by using the following relationships:

$$\beta_{\text{x}} = \beta_{\text{x}}^{\perp} + \beta_{\text{x}}^{\parallel}, \tag{2.2}$$

$$\beta_{\text{x}}^{\parallel} = \frac{\beta_{\text{x}}}{1 + \delta_{\text{x}}}, \tag{2.3}$$

$$\beta_{\text{x}}^{\perp} = \frac{\beta_{\text{x}}\delta_{\text{x}}}{1 + \delta_{\text{x}}}, \tag{2.4}$$

where the subscript x denotes either the molecular or the particle contribution. By using Eq. (2.3) and (2.4) all the terms at the right side of Eq. (2.1) can be replaced, and the following expression is obtained:

$$\delta_{\text{vol}} = \frac{\frac{\beta_{\text{m}}\delta_{\text{m}}}{1+\delta_{\text{m}}} + \frac{\beta_{\text{p}}\delta_{\text{p}}}{1+\delta_{\text{p}}}}{\frac{\beta_{\text{m}}}{1+\delta_{\text{m}}} + \frac{\beta_{\text{p}}}{1+\delta_{\text{p}}}}. \tag{2.5}$$

Simple conversion and rearrangements yield to:

$$\delta_{\text{p}} = \frac{\beta_{\text{m}}(\delta_{\text{vol}}-\delta_{\text{m}}) + \beta_{\text{p}}\delta_{\text{vol}}(1 + \delta_{\text{m}})}{\beta_{\text{m}}(\delta_{\text{m}}-\delta_{\text{vol}}) + \beta_{\text{p}}(1 + \delta_{\text{m}})}. \tag{2.6}$$

Similarly, the same steps from Eq. (2.1) to Eq. (2.6) can be used to separate desert dust and non-dust contributions. The subscripts volume (vol), molecules (m) and particles (p) can be replaced by particles (p), dust (d) and non-dust particles (nd), respectively. Then the following equation is obtained:

$$\delta_p = \frac{\beta_{\text{nd}}\delta_{\text{nd}}(1 + \delta_{\text{d}}) + \beta_{\text{d}}\delta_{\text{d}}(1 + \delta_{\text{nd}})}{\beta_{\text{nd}}(1 + \delta_{\text{nd}}) + \beta_{\text{d}}(1 + \delta_{\text{nd}})}. \tag{2.7}$$

The backscatter coefficient of dust particles is determined after substitution of β_{nd} with β_{p}–β_{d} and calculation of δ_p from Eq. (2.6):

$$\beta_{\text{d}} = \beta_{\text{p}} \frac{(\delta_{\text{p}}-\delta_{\text{nd}})(1 + \delta_{\text{d}})}{(\delta_{\text{d}}-\delta_{\text{nd}})(1 + \delta_{\text{p}})}. \tag{2.8}$$

The dust backscatter coefficient can be multiplied by the respective dust lidar ratio S_{d} to obtain the dust particle extinction coefficient:

$$\alpha_{\mathrm{d}} = S_{\mathrm{d}}\beta_{\mathrm{d}}. \tag{2.9}$$

The estimates of the particle linear depolarization ratio of pure non-dust particles δ_{nd} and pure dust particles δ_{d} needed for the calculations above are key input parameters for this method and were derived from values found during the SAMUM–1 campaign after comparison with already existing climatologies from the literature. Values of 31% for dust and 5% for non-dust (i.e., smoke) were used at 532 nm.

Depending on the lidar and its configuration, this approach can provide data sets of three backscatter (355, 532 and 1064 nm) and two extinction coefficients (355 and 532 nm) for the dust and non-dust components (i.e., 3+2 configuration) as it has been demonstrated in Tesche et al. (2009b) and other studies (Tesche et al., 2011b; Marinou et al., 2017). Further refinements to the methodology (polarization-lidar photometer networking; POLIPHON) have allowed the separation of dust particles into fine and coarse mode (Ansmann et al., 2011b, 2012; Mamouri and Ansmann, 2014) as well as the derivation of CCN- and INP-relevant aerosol parameters (Mamouri and Ansmann, 2016, 2017).

2.2 The CALIPSO aerosol typing scheme

As mentioned in Ch. 1, CALIPSO has its own typing algorithm (Omar et al., 2009) that is able to classify up to six different aerosol types, namely desert dust, biomass burning, background (clean continental), polluted continental, polluted dust and clean marine. This typing scheme is essential for the correct assignment of lidar ratios to features (aerosol or cloud layers). Since CALIOP is a simple elastic-backscatter lidar (Winker et al., 2009; Omar et al., 2009; Young and Vaughan, 2009), the lidar ratio is needed for further data processing, i.e., to derive the backscatter coefficient and make an estimate for the aerosol extinction coefficient. This means CALIOP is not able to conduct direct extinction measurements although this parameter is often used and depends strongly on the a priori typing.

After a feature mask has been applied to the measurements, to identify regions of the same optical properties (i.e., features), the typing scheme categorizes the features into the aforementioned six aerosol types by means of a decision tree. The decision tree takes into account the lidar-derived feature-integrated volume linear depolarization ratio and attenuated backscatter, feature height information as well as external information such as geographical location, underlying surface type and season. Once an aerosol type has been selected from the decision tree, the corresponding lidar ratio (at 532 and 1064 nm) is assigned to the feature.

Behind the aforementioned six aerosol types, there are aerosol models that are obtained from a multiyear AERONET cluster analysis (Omar et al., 2005; Cattrall et al., 2005) of the optical and physical properties of characteristic aerosol types (Omar et al., 2005) and, therefore, are representative for the most frequently observed aerosol mixtures at AERONET sites. In addition, every aerosol model has its own, predefined size distribution and microphysical properties (e.g., refractive indices).

With the release of new updated data versions, several limitations in the classification scheme emerged. Kanitz et al. (2014) compared the CALIPSO typing scheme against aerosol-typing-relevant parameters obtained by a ground-based lidar that was located in the coastal city Punta Arenas, Chile. It was shown that CALIOP's data analysis classified the detected aerosol either as polluted dust, polluted continental aerosol, or smoke, while the ground-based lidar analysis indicated the presence of marine particles. This discrepancy was the result of the surface-dependent CALIPSO typing scheme, which did not allow marine aerosol to be present over land. Another similar surface-dependent limitation is that Asian dust and biomass-burning particles from boreal fires observed above the polar region were missclassified as either clean continental or polluted continental, due to the fact that these were the only aerosol types allowed above surfaces characterized as snow, ice or tundra. Therefore, to address such limitations, a new updated version (version 4) of the algorithm was developed and released by Kim et al. (2018). Major changes in the algorithm include the introduction of four aerosol types for stratospheric aerosol, namely polar stratospheric aerosol (PSA), volcanic ash, sulfate/other and smoke. Regarding tropospheric aerosols, all types are now allowed over polar regions, a new aerosol mixture (dusty marine) was introduced and the polluted continental and smoke types have been renamed to polluted continental/smoke and elevated smoke, respectively. In addition, the existing lidar ratios for clean marine, dust, clean continental, and elevated smoke categories have been revised to reflect the current knowledge state for these types. The aforementioned changes lead to a more accurate feature type identification and aerosol type classification and, thus, more reliable extinction retrievals (Kim et al., 2018).

CALIPSO's typing scheme updates did not resolve the ambiguity regarding the aerosol type naming, which varies heavily in different aerosol studies and applications. Apart from pure aerosol types (e.g., clean marine, desert dust), all the other aerosol mixtures, referred to also as types, are in some degree vague. For example, the aerosol type "clean continental", which is included in CALIPSO's typing scheme, is frequently encountered in many other classification schemes (e.g., in Papagiannopoulos et al. (2018) discussed in Sec. 2.3). The clean continental aerosol is typically described as a mixture of anthropogenic pollution and natural aerosol (e.g., sea salt). Clean continental aerosol layers are high-transmittance layers (low AOD, Ansmann et al. (2001)), with low depolarizing ability ($\delta_{\mathrm{vol}} < 0.07$, Omar

et al. (2009)), lidar ratios ranging between 20 and 53 sr (Omar et al., 2009; Kim et al., 2018) and Ångström exponents above 1 (Ansmann et al., 2001). This high variability in the observations typed as clean continental further suggests that the term itself is very general and that the determination of the different aerosol components present in the mixture is quite hard. For instance, in the latest update of CALIPSO's typing scheme (version 4, Kim et al. (2018)) the lidar ratio for clean continental is set to 53 ± 24 sr at 532 nm, suggesting that the mixing state of this mixture can vary significantly (from less absorbing sea salt particles to highly absorbing smoke particles), without any further quantification. Such classification results are not particularly helpful for the quantification of the radiative effect or for atmospheric models and data assimilation.

2.3   Aerosol typing by utilizing the Mahalanobis distance

Another way of aerosol typing can be performed by separating classes/clusters of measurements that have predefined characteristics and then adding new points/measurements into these classes based on the distance of the measurements from each class center. A distance metric frequently used is the Mahalanobis distance (Mahalanobis, 1936), defined as

$$D_{\mathrm{M}}(\boldsymbol{x} - \overline{\boldsymbol{x}}) = \sqrt{(x - \overline{x})^T \boldsymbol{S}^{-1}(x - \overline{x})}. \qquad (2.10)$$

In Eq. (2.10), $\boldsymbol{x}$ represents the multidimentional observation vector, $\overline{\boldsymbol{x}}$ the mean class, $\mathbf{S}$ the class covariance matrix and D_{M} the ellipsoids, centered around the mean of each class. As Eq. (2.10) suggests, the Mahalanobis distance metric takes into account the variance in each variable and their corresponding covariance making it, thus, an ideal metric for classification purposes. Once the clusters have been defined, any new measurement serves as a new point and its Mahalanobis distance to each of the clusters is calculated. Then, this new point is assigned to the closest class (smallest Mahalanobis distance).

Burton et al. (2012) were the first to apply the Mahalanobis distance to classify aerosol observations into their respective aerosol types. Based on specific samples of known aerosol types, eight aerosol classes were constructed (pure dust, dusty mix, maritime, polluted maritime, urban, fresh smoke, smoke and ice) and HSRL measurements of aerosol intensive parameters were used for the classification procedure. The parameters used were the particle linear depolarization ratio (at 532 nm), the lidar ratio (at 532 nm), the backscatter-related color ratio (532-to-1064 nm) and the ratio of particle depolarization ratios (spectral depolarization ratio, at 1064-to-532 nm). The relative discriminatory power of each intensive parameter (Wilks' overall and partial lambda statistic, Hill and Lewicki (2007)) was also explored by this study, and it was found that the depolarization ratio had the most weight in the classification procedure, followed closely by the depolarization spectral ratio

and the lidar ratio. The backscatter color ratio was found to be the parameter with the least discriminatory power.

A similar approach has been applied by Papagiannopoulos et al. (2018) aiming at building a classification tool able to provide near-real-time aerosol typing information for EARLINET. Based on EARLINET classified data, eight aerosol classes were defined initially (clean continental, polluted continental, dust, mixed dust, polluted dust, mixed marine, smoke and volcanic ash). Four classifying parameters were used: the particle linear depolarization ratio (532 nm), the lidar ratio (532 nm), the Ångström exponent (355/1064 nm), and the ratio of lidar ratios (532-to-355 nm). The aerosol classes represent the most frequently observed aerosol types over the EARLINET stations and, as in every supervised learning technique, are crucial to the overall predictive performance of the algorithm. The impact of the number of aerosol classes to the classification performance was also investigated in this study. For this purpose, the classes number varied between eight and four. The error rate in the classification was found to decrease monotonically with decreasing number of aerosol classes (from 39% to 10%). The prediction rate of the automatic classification showed positive results when compared against the already classified data. More specifically, the less the classes the highest the learning success of the algorithm but with the cost of a much coarser classification. Nevertheless, the algorithm is flexible (available observed variables), automated and the application to such a large (and continuing) dataset of great importance for the scientific community.

2.4   Neural network aerosol typing algorithm based on lidar data

In a recent study, Nicolae et al. (2018) explored the potential of an Artificial Neural Network (ANN) for the identification of the most probable aerosol type, using multispectral lidar data. The Neural Network Aerosol Typing Algorithm Based on Lidar Data (NATALI) uses ANNs with different learning rules that were trained to classify aerosols using only a combination of mean layer intensive optical parameters (i.e., lidar ratios, Ångström exponent, color ratios) from the provided aerosol backscatter and extinction coefficient profiles of lidar systems, without any additional information. The ANNs were trained using synthetic data based on an aerosol model that was developed for this reason. A spheroidal shape representation is considered for particles based on OPAC-defined internal mixtures (Optical Properties of Aerosols and Clouds, Hess et al. (1998)). The aerosol microphysical properties were derived from the Global Aerosol Data Set (GADS, Koepke et al. (1997)). To account for the natural variability of the optical properties, different mixing ratios and relative humidity (RH) values were considered. The uncertainty of the observations was

also taken into account.

The input data required are typical EARLINET data products, namely backscatter coefficient (355, 532, 1064 nm), extinction coefficient (355, 532 nm) and linear particle depolarization (532 nm) if available (so-called $3\beta + 2\alpha + 1\delta$ configuration). As a first step, NATALI identifies the geometric boundaries of the aerosol layers by utilizing the gradient method for the 1064-nm backscatter coefficient profile (Belegante et al., 2014). Then, for each detected aerosol layer, mean layer intensive optical parameters and their corresponding uncertainties are calculated at all available wavelengths and wavelength combinations (e.g., in the case of backscatter-related color ratios). As a first quality assurance procedure, the calculated intensive optical properties are checked and only the layers with values that fall within acceptable range limits (e.g., lidar ratio values between 5 and 200 sr) are considered for further evaluation. In the second step, three ANN perform simultaneously and provide the most probable aerosol type that is then assigned to the feature represented by the input data. The identification of the most probable aerosol type is realized through a voting procedure that takes into account the confidence level of the ANN output and the overall retrieval stability (i.e., within uncertainty range).

The algorithm has two different classification modes, based on the information available, and the quality of the output product is dependent on the quality of the optical parameters provided as input. The first one is the high-resolution mode, which includes depolarization ratio information and allows the identification of 14 aerosol mixtures. When no depolarization ratio is available, the low-resolution mode allows the identification of five predominant aerosol types. The 14 aerosol mixtures included in the high-resolution mode are: continental, continental polluted, dust, maritime, smoke, volcanic, coastal, coastal polluted, continental dust, continental smoke, dust polluted, maritime mineral, mixed dust and mixed smoke. In the low-resolution mode the aerosol mixtures include only continental, continental polluted, smoke, dust, maritime and volcanic aerosol. Finally, the output includes the intensive optical parameters within each identified aerosol layer, along with their mean values and corresponding uncertainty, and the most probable aerosol type.

An extensive performance analysis of the methodology and the ANN was also carried out. In particular, the aerosol model outputs were compared with literature-based intensive optical parameters for each aerosol type considered, ANN's performance during the learning phase was compared to the synthetic data and finally NATALI retrieval performance was tested against the EARLINET-CALIPSO classification retrieved aerosol types. NATALI's overall performance was good (more detailed information in Nicolae et al. (2018)), being able to correctly classify most aerosol types in both high-resolution and low-resolution mode. However, the depolarization ratio (high-resolution mode) provided considerably higher confidence retrieval results, especially for aerosol mixtures. Pure aerosol types were

recognized by NATALI for all values of RH. However, ANNs showed a decreased performance with an increase in RH only for continental smoke and continental dust for high-resolution typing and for continental smoke and mixed smoke for low-resolution typing. On the other hand, for coastal polluted, RH increase results in an increase of typing performance.

2.5   Hybrid End-To-End Aerosol Classification model

HETEAC is an aerosol classification model that was developed for the EarthCARE mission (Wandinger et al., 2016a). The model serves as a common baseline for development, evaluation and implementation procedures within all the development activities for Earth-CARE. Data synergy between EarthCARE's ATLID, MSI and BBR is further supported with HETEAC as consistency between their aerosol products is ensured. The approach is hybrid, meaning that the theoretical description of the aerosol microphysics is consistent with the aerosol optical properties obtained by observations. In addition, the end-to-end model approach ensures a uniform representation of the different aerosol types in terms of microphysical, optical and radiative properties.

The first step in the HETEAC model development is the definition of the basic aerosol components. Based on an experimental basis of optical properties from ground-based lidars at EarthCARE's wavelength (355 nm), four components were defined as representatives of the aerosol types observed in nature. These components comprise of two fine modes, one strongly absorbing and one weakly absorbing, and two coarse modes, one with spherical particles and one with non-spherical particles. A detailed description and insights of the aerosol components is provided in Sec. 4.1.3.

For each of these components, a mono-modal particle size distribution and a wavelength-dependent complex refractive index is assigned to obtain their microphysical description, based on ESA's Climate Change Initiative (CCI) project Aerosol_cci (Holzer-Popp et al., 2013). The mode radii and refractive indexes are obtained from AERONET and are considered typical for the aerosol components. To describe the scattering of the non-spherical particles, two models were examined. One is the spheroid model introduced by Dubovik et al. (2006) and the other one is proposed by Gasteiger et al. (2011). While both models are widely used in several applications (e.g., Dubovik et al. (2006) in Aerosol_cci and Gasteiger et al. (2011) in OPAC), the model and spheroid distribution chosen for HETEAC was the one of Gasteiger et al. (2011), mainly due to the degree of reproducibility of the observations.

To account for aerosol mixtures of two or more modes, a multimodal representation is achieved by mixing rules. Each mode has specific scattering properties per unit particle volume, which are then used, in combination with the relative volume contribution of each

mode, to derive the optical properties of the aerosol mixture. This procedure results in lookup tables (LUT) of the optical and radiative properties for the different mixing ratios between the aerosol modes at 355 nm. Even though HETEAC is optimized for the 355 nm wavelength, LUT have been constructed for other wavelengths as well. Once EarthCARE is in orbit, the LUT can be used in reverse to reveal the mixing ratio of the four different aerosol components from ATLID's measurements (at 355 nm only).

2.6   Discussion and research questions

The complexity and necessity of accurate aerosol typing is well revealed by the different approaches and methodologies developed solely for that purpose. The scientific efforts towards resolving the aerosol classification issue are translated, among others, into methodology developments that are spanning from rather simple and straightforward threshold-based techniques to advanced ANN algorithms.

Nevertheless, the issue remains highly subjective in several aspects. For instance, aerosol type naming varies heavily depending on the studies, leading to ambiguous assumptions on the atmospheric aerosol loading and subsequent conclusions, as discussed already for the CALIPSO typing methodology (Sec. 2.2). The naming of the different aerosol types is usually driven by the aerosol's source origin (e.g., marine, desert dust, continental), providing thus not sufficient information on the microphysical properties of the aerosol particles (i.e., size, shape).

Optical parameter values, assigned to aerosol types and used in several aerosol classification schemes as well as in atmospheric modeling, are another issue that polarizes the scientific community. These values, their thresholds and uncertainties to some extent depend on the instrument performing the measurements and, thus, generalizations are required (e.g., a global lidar ratio for desert dust, regardless of its source origin). Several campaigns aimed to improve the knowledge on certain aerosol types, e.g., the SAMUM–1 and SAMUM–2 field experiments (Ansmann et al., 2011a), which focused on the investigation of the relationship between chemical composition, shape morphology, size distribution and optical effects of the dust particles originating from the Saharan desert. Another location-focused campaign was the Central Asian Dust Experiment (CADEX, Hofer et al. (2017, 2020)), which focused on long-term observations of the optical and microphysical properties of Central Asian mineral dust. It was the first time that ground-based lidar observations were conducted in Central Asia at Dushanbe, Tajikistan. While CADEX lasted two years (2014–2016), it resulted in major findings in terms of optical properties of Central Asian dust, establishing it thus as a separate aerosol type. In addition, the campaign pointed out the necessity of a permanent ground-based station, which was later on established as part of EARLINET. It is

obvious that campaigns aiming at monitoring a specific aerosol type, in a specific location where it is found in abundance, are very important in the global efforts of aerosol type standardization. In addition, global satellite observations also help shed light on previously unexplored areas and observational gaps. However, till nowadays there is no commonly accepted aerosol-type-linked range of optical properties.

As mentioned already, lidar-based aerosol classification and typing methodologies are based mainly on lidar-derived aerosol optical properties (Sec. 3.1). Different combinations of these aerosol properties have been used in aerosol classification studies (e.g., the lidar ratio, particle linear depolarization ratio, backscatter color ratio and depolarization spectral ratio in Burton et al. (2012)). Most studies conclude that the particle linear depolarization ratio along with the lidar ratio are really crucial for classification and typing purposes (e.g., Sassen, 1991; Tesche et al., 2009a,b; Groß et al., 2011; Burton et al., 2012; Papagiannopoulos et al., 2018; Nicolae et al., 2018). However, there is still no clear suggestion on the minimum number of optical parameters required for classification purposes. While considering many optical parameters might be necessary sometimes, it could also lead to an overclassification of the aerosol.

Table 2.1: Summary of the lidar-based aerosol typing schemes presented in Ch. 2.

Typing scheme	Ground-based	Space-borne	Height-resolved	Layer-mean values	Quantification
POLIPHON	✓		✓		✓
CALIPSO		✓		✓	
Mahalanobis	✓			✓	
NATALI	✓			✓	
HETEAC		✓		✓	✓

A summary of the key aspects of the aerosol typing developments discussed in this chapter is presented in Tab. 2.1. These key aspects can be categorized in the applicability of the typing scheme (i.e., ground-based or spaceborne lidars), on the type of input required (i.e., height-resolved profiles of optical parameters or layer-mean values) and on whether or not a quantification of the aerosol types present in the mixture observed is provided. Tab. 2.1 reveals that no typing scheme is superior to others and that the application influences heavily the choice of the methodology. For instance, if one is interested in height-resolved aerosol typing based on ground-based lidars, POLIPHON is a suitable choice, as its output provides the height-resolved fraction of backscatter and extinction coefficient attributable to dust and non-dust particles. A minor drawback of this methodology is that the non-dust category is not automatically labeled/classified. On the other hand, if one is interested in automatic typing procedures, the Mahalanobis distance methodology (Sec. 2.3) or NATALI (Sec. 2.4) are to be preferred. Key difference between these two methodologies is that the first one

depends on a supervised learning technique, while the second one uses ANN that have been trained via a supervised learning technique, but perform the classification unsupervised.

Driven by the advantages and disadvantages of the methodologies discussed above and by considering the need of ground-based validation of the upcoming EarthCARE mission, a new aerosol typing methodology was developed and is presented in Ch. 4. The methodology is based on HETEAC, but utilizes the optimal estimation method to provide a quantification of the contribution, in terms of relative volume, of four different aerosol components to an aerosol mixture. The aerosol components retrieved are named based on their microphysical and optical properties. The methodology is flexible in terms of required input and thus suitable for most lidar systems (ground-based or spaceborne). A synopsis of the key research questions aimed to be addressed by this dissertation is provided below.

Research questions

This dissertation endeavours to provide answers to the following key questions:

1. Can an optimal-estimation-based methodology be applied effectively for aerosol typing purposes?

2. Is it possible to describe an aerosol mixture observed by means of basic aerosol components?

3. Can we retrieve the contribution of each aerosol component, in terms of relative volume, leading thus to a quantitative characterization of the aerosol mixture?

4. Is the proposed methodology sufficient for ground-based validation purposes and further support of EarthCARE's aerosol classification scheme (HETEAC)?

To address these questions, it is essential to begin with the construction of a basis of lidar-derived aerosol intensive optical parameters at different wavelengths. The basis will allow the identification of the different major aerosol types and the different mixtures that they form as aerosol components, as well as their corresponding optical properties (Ch. 3). An optimal-estimation-based methodology was chosen as it ensures to provide the statistically most likely conditions that produce a given measurement, i.e., intensive optical properties (Ch. 4). At the same time, the goodness of the solution relies on the predefined aerosol components and, therefore, the experimental basis constructed is a crucial element for the typing scheme proposed. The retrieval scheme proposed here is then applied to several case studies, as well as to a long-term dataset, to asses whether or not it provides results that allow for a quantitative characterization of the aerosol mixtures (Ch. 5, 6).

Chapter 3

Optical properties of different aerosol types

As discussed in the literature review (Ch. 2), intensive, i.e., concentration-independent, particle optical parameters can be used as "optical footprints", allowing thus the identification and classification of the different aerosol types. This chapter begins with a brief overview of lidar-derived optical parameters that can and have been used in previous studies for aerosol typing (Sec. 3.1).

The typing methodology that is described in detail in the following chapter (Ch. 4) relies on measured intensive optical properties and, similarly to HETEAC (Sec. 2.5) but for more wavelengths (355 and 532 nm), an experimental basis of ground-based lidar measurements at various locations and from various platforms was created. More specifically, the lidar ratio, the particle linear depolarization ratio and the Ångström exponent (backscatter- and extinction-related) at the aforementioned wavelengths were used to create the experimental basis. The experimental basis allows the exploration of the solution space for the proposed typing methodology, i.e., the identification of the different aerosol types (or clusters) and their optical properties. In addition to the combination of optical parameters shown in the following, all possible combinations available are presented in Appendix A.

3.1 Aerosol-typing-relevant optical properties

Multiwavelength polarization Raman lidars emit polarized light at different wavelengths (typically at 355, 532 and 1064 nm) into the atmosphere. The fraction of the light that gets backscattered by molecules and particles is detected by the instrument's receiver with a high temporal resolution. After the appropriate corrections (e.g., background subtraction, temporal averaging, overlap correction), the retrieved profiles of backscatter (β) and extinc-

tion (α) coefficients (typically at 355, 532 and 1064 nm for β and 355, 532 nm for α) can be used to calculate optical parameters that are usually called intensive, i.e., concentration-independent, particle optical properties. These parameters, along with the particle linear depolarization ratio (at 355 and 532 nm), are sensitive to the size, shape and refractive index of the atmospheric particles and therefore can be used effectively for aerosol characterization (Ansmann and Müller, 2005; Müller et al., 2007; Freudenthaler et al., 2009; Groß et al., 2011). These intensive parameters are discussed in more detail in the following sections.

3.1.1   Lidar ratio

The particle lidar ratio (S^{p}) or simply lidar ratio (S), defined as the extinction-to-backscatter ratio (p denotes particle),

$$S^{\mathrm{p}}(R) = \frac{\alpha^{\mathrm{p}}(R)}{\beta^{\mathrm{p}}(R)}, \tag{3.1}$$

is a valuable quantity for aerosol characterization. The lidar ratio reveals information regarding the size, shape and the absorption efficiency of the aerosol particles. The particle size and the lidar ratio are usually inversely related, while the particle absorption efficiency and the lidar ratio are directly related (e.g., S increases with particle absorption efficiency). When spherical and non-spherical particles have the same size range, the non-spherical ones tend to exhibit higher lidar ratio values (Müller et al., 2007). However, all these different effects could counterbalance each other, leading to similar lidar ratios for different aerosol types. For example, in the study of Müller et al. (2007), planetary boundary layer (PBL) observations from Saharan desert dust (observed during the SAMUM–1 campaign) and Central European urban haze (observed within the EARLINET framework) exhibited similar lidar ratio values of 55–60 sr at 355 nm. It becomes clear that, since large, non-spherical, moderately absorbing particles (such as desert dust particles) exhibit similar lidar ratios as small, spherical and absorbing particles (e.g., urban haze particles), the lidar ratio alone is not sufficient to adequately characterize the different aerosol types observed in the atmosphere and, therefore, additional information is needed.

3.1.2   Particle linear depolarization ratio

As mentioned earlier, lidars emit polarized light and during the backscatter process, parts of this light get unpolarized. The depolarization ratio (Schotland et al., 1971) usually refers to the ratio of signals measured in the perpendicular and parallel receiver channels (cross-polarized and co-polarized, respectively). Cross-polarized ($\perp$) refers to light with a state of polarization that is perpendicular to the one of the emitted laser light, and co-polarized ($\parallel$) refers to light that has the same state of polarization as the emitted laser light. The

linear volume depolarization ratio (δ_{vol}) comprises contributions of molecules and aerosol particles (subscripted as m and p, respectively) and is defined as:

$$\delta_{\text{vol}} = \frac{\beta^{\perp}}{\beta^{\parallel}} = \frac{\beta_{\text{m}}^{\perp} + \beta_{\text{p}}^{\perp}}{\beta_{\text{m}}^{\parallel} + \beta_{\text{p}}^{\parallel}}. \tag{3.2}$$

The molecules and particles can be considered independently, leading to the molecular (δ_{m}) and particle (δ_{p}) linear depolarization ratio, respectively:

$$\delta_{\text{m}} = \frac{\beta_{\text{m}}^{\perp}}{\beta_{\text{m}}^{\parallel}}. \tag{3.3}$$

$$\delta_{\text{p}} = \frac{\beta_{\text{p}}^{\perp}}{\beta_{\text{p}}^{\parallel}}. \tag{3.4}$$

The particle linear depolarization ratio (δ_{p} or often denoted as δ, Eq. (3.4)) is an intensive property of the ensemble of scattering particles and can be used for aerosol typing because spherical particles do not alter the state of polarization of the emitted light ($\delta_p \approx 0$), while non-spherical particles do ($\delta_p > 0$). For instance, for the same example that was discussed in Sec. 3.1.1, desert dust particles exhibit particle linear depolarization ratio values of 30%–35%, while for urban haze particles the respective values are less than 5%. While the lidar ratio does not provide enough information to allow distinguishing one type from another in this case, the particle linear depolarization ratio provides the information that is needed to distinguish spherical from non-spherical particles.

3.1.3   Ångström exponent

The Ångström exponent (Å) is a parameter that contains information regarding the spectral dependency of the aerosol optical properties (Ångström, 1964; Ansmann and Müller, 2005) and can be obtained from, e.g., multiwavelength lidars. An Ångström exponent of zero denotes wavelength independence of the investigated quantity. The Ångström exponent can be calculated for the backscatter coefficient (β), the extinction coefficient (α) and the lidar ratio (S) when they are available at least for two distinctive wavelengths λ_1 and λ_2, respectively (Ansmann et al., 2002).

The so-called backscatter-related Ångström exponent can be acquired from the particle backscatter coefficients at the wavelength pairs of 355 and 532 nm and 532 and 1064 nm, respectively:

$$\mathring{A}_{\beta_{355/532}} = -\frac{\ln\left(\frac{\beta_{532}}{\beta_{355}}\right)}{\ln\left(\frac{532}{355}\right)}, \tag{3.5}$$

$$\AA_{\beta_{532/1064}} = -\frac{\ln\left(\frac{\beta_{1064}}{\beta_{532}}\right)}{\ln\left(\frac{1064}{532}\right)}. \qquad (3.6)$$

Similarly, for the extinction-related Ångström exponent it is:

$$\AA_{\text{ext}} = -\frac{\ln\left(\frac{\alpha_{532}}{\alpha_{355}}\right)}{\ln\left(\frac{532}{355}\right)}. \qquad (3.7)$$

In general, the extinction-related Ångström exponent is around zero for large particles, while scattering by small particles exhibits a strong wavelength dependence and causes Å values greater than unity (Eck et al., 1999). The Ångström exponent of molecules (much smaller compared to small aerosol particles) is around 4 (Antoine and Morel, 1998). Continuing with the example from Müller et al. (2007), which has been discussed already in Sec. 3.1.1 and 3.1.2, Saharan desert dust particles have extinction-related Ångström exponent values of 0.2±0.2, while for urban haze particles these values are of the order of 1.4±0.5 and, thus, the Ångström exponent would be a good additional typing parameter in the specific example. The extinction- and backscatter-related Ångström exponents act as a proxy for the size of the particles and, thus, a separation between large and small particles can be achieved.

3.1.4   Backscatter color ratio

The backscatter color ratio is an intensive property similar to the Ångström exponent. It is defined as the ratio of the backscatter coefficients acquired at two different wavelengths λ_1 and λ_2, respectively:

$$C_{\beta_{\lambda_1/\lambda_2}} = \frac{\beta_{\lambda_1}}{\beta\lambda_2}. \qquad (3.8)$$

Typically, backscatter color ratios are inversely related to the aerosol particle size and, therefore, have been used for aerosol classification purposes (Sasano and Browell, 1989; Sugimoto et al., 2002; Burton et al., 2012). The backscatter color ratio does not contain any additional information compared to the backscatter-related Ångström exponent.

3.2   Intensive optical properties at 355 and 532 nm

3.2.1   Data collection

Figure 3.1 shows a collection of ground-based observations of lidar ratio and particle linear depolarization ratio (at 355 nm) acquired during different field campaigns at different locations using mainly (but not only) Polly lidar systems, which are part of the TROPOS-led PollyNET (Baars et al., 2016). In addition to a basic experimental data set published in

Illingworth et al. (2015), new findings and new aerosol types identified during recent measurement campaigns have been added to consolidate and extend both the basis of HETEAC and the typing approach proposed in this thesis (described in Ch. 4).

All these different data points have been determined and identified using the same standard methodology. The first step is the identification of the aerosol layers followed by calculations of the layer-mean aerosol properties. Then these layer-mean aerosol properties can be attributed to one aerosol type or to a mixture. To further support the aerosol type characterization, an air mass source attribution scheme is usually applied to the selected case.
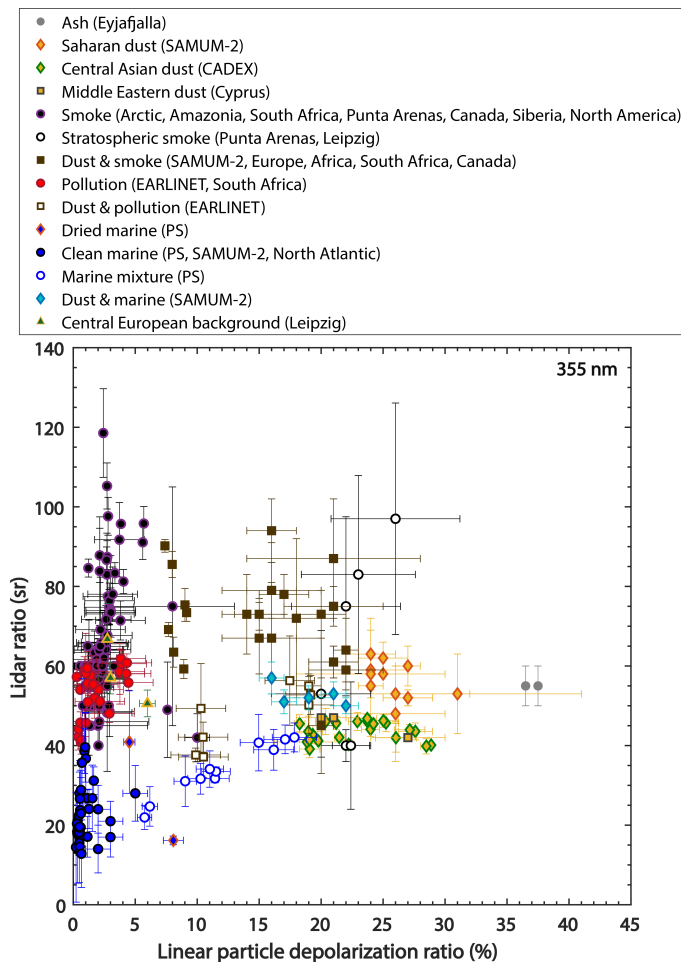
Figure 3.1: Intensive optical properties of different aerosol types, measured at 355 nm mainly within the PollyNET framework. Measurements from SAMUM–2 and Eyjafjallajökull (volcanic ash) were performed also with the Raman-polarization lidars POLIS (University of Munich).

The smoke category (depicted by black/purple circles in Fig. 3.1), which so far included data from the EUCAARI campaign in the Amazon Basin in 2008, was significantly ex-

tended (with respect to the HETEAC basis) to include smoke observations from other locations and from other fire types (e.g., smoldering or flaming combustion). Tropospheric and stratospheric smoke observations from the most recent field campaign DACAPO-PESO at Punta Arenas, Chile have been included in the basis (Floutsi et al., 2021; Ohneiser et al., 2020). For the stratospheric observations, the smoke particles originated solely from the record-breaking Australian wildfires of January 2020, while for the tropospheric observations, smoke particles originated also from fires that occurred in South America. Data for smoke from boreal forest fires in western Canada that reached and was observed above Central Europe (Leipzig, Germany) in 2017 were also included (Haarig et al., 2018). The collection of tropospheric smoke observations (originating from various locations and observed above Europe) by Haarig et al. (2018) was also incorporated. These observations gave new insights in the highly variable nature of the optical properties of smoke particles. Especially the different properties of smoke in the troposphere and stratosphere (resulting from Pyrocumulonimbus convection) were not realized before the 2017 smoke event and were first published by Haarig et al. (2018) and Ansmann et al. (2018). Smoke observations acquired during the MOSAiC campaign in the Arctic were also added to the basis (Ohneiser et al., 2021). The smoke particles originated mainly from wildfires in Siberia that occurred in July and August 2019 and reached stratospheric heights due to self-lifting effects. Smoke observations conducted at Elandsfontein, South Africa were considered as well (Giannakaki et al., 2016). These smoke observations exhibit a wide range of optical properties (due to differences in their chemical composition) and broaden even further the regional scales at which measurements were conducted.

The dust category considered for the development of HETEAC is based on Saharan dust observations that were conducted at Cape Verde during the SAMUM–2 campaign in 2008 (Groß et al., 2011). In the present study, this category (depicted by orange/red rhombuses in Fig. 3.1) has been extended to include Saharan dust observations over various locations (e.g., over Portugal (Preißler et al., 2011), over Leipzig (Baars et al., 2016) and all the Saharan dust observations provided in Baars et al. (2016)). Saharan dust measurements from the first SAMUM campaign were also added (adapted from Tesche (2011) and Müller et al. (2007)). Since the Eastern Mediterranean has been characterized as a hotspot for climate change, measurements and monitoring of the aerosols over that region have been intensified. Therefore, dust data from Cyprus obtained during the CyCARE campaign (Ansmann et al., 2019) were added to the basis (brown/yellow squares in Fig. 3.1). The dust particles observed above Cyprus originate either from the Saharan desert or the Middle Eastern deserts (e.g., the Arabian desert) or are a mixture from both sources. Therefore, their optical properties can be differentiated from the Saharan dust category. The CADEX campaign (already mentioned in Ch. 2) advanced the knowledge of the optical properties of

Central Asian dust. In particular, it was revealed that desert dust particles originating from the Central Asian region exhibit lidar ratio values that are lower than those of dust particles originating from the Saharan desert dust and comparable to the ones of dust particles originating from the Middle Eastern deserts (Hofer et al., 2017). Therefore, Central Asian dust measurements conducted in the framework of CADEX at Dushanbe, Tajikistan have been added to the basis as a separate dust aerosol type (green/yellow rhombuses in Fig. 3.1, Hofer et al. (2017)).

In HETEAC, the measurements considered for marine particles (i.e., sea salt) were obtained during the SAMUM–2 campaign in Cape Verde. This category (blue/black circles in Fig. 3.1) has been extended with additional measurements from the North Atlantic (reported in Groß et al. (2013)), to cover a broader geographical region. Recent marine boundary-layer measurements from Polly$^{\text{XT}}$ on board of the German research vessel Polarstern (PS) have been included to the experimental basis (blue/black circles in Fig. 3.1, Bohlmann et al. (2018)). Marine aerosol appears to have always similar optical properties regardless of its origin. However, in dry atmospheric conditions marine particles show significantly higher depolarization ratio values. Dry marine particles were observed (and measured) in the same PS cruise (Bohlmann et al., 2018) and have been introduced in the basis as a separate aerosol type (blue/red rhombuses in Fig. 3.1). In addition, the marine mixtures in the boundary layer that were observed during this cruise are now part of the basis as an additional aerosol category (blue/white circles in Fig. 3.1).

The "Central European background" aerosol type has been introduced for the first time. It includes measurements of the background aerosol load in a typical Central European region (performed at Leipzig, Germany). The optical values of this type (yellow/green triangles in Fig. 3.1) are similar to HETEAC's aerosol type "Pollution" (Illingworth et al., 2015) and depict exactly the close relation of those two aerosol types.

The categories of aerosol mixtures have been also extended and refined with respect to those presented in Illingworth et al. (2015) (namely "Dust and Smoke" and "Dusty mixtures"). For instance, the "Dusty mixtures" category in Illingworth et al. (2015) has been renamed into "Dust and Marine" to better reflect the aerosol types present in the mixture. New mixtures, such as Saharan dust and African biomass-burning aerosol (adapted from Baars et al. (2016)), were also included in the broader "Dust and Smoke" category (brown squares in Fig. 3.1).

As mentioned already, the typing methodology proposed in this work is applicable to multiwavelength ground-based lidars and, therefore, the experimental basis has been extended to include optical properties of different aerosol types at 532 nm. Fig. 3.2 shows the same collection of ground-based observations of lidar ratio and particle linear depolarization ratio as Fig. 3.1, but for 532 nm. In some cases, optical information was missing for

532 nm, but also some additional measurements were added: Saharan dust measurements over Portugal (Preißler et al., 2011), urban aerosol observations over the Pearl River Delta in Guangzhou, China (Heese et al., 2017), pollution observed above Leipzig and dust and pollution mixtures from India and Southeast Asia (Clarke et al., 2002).
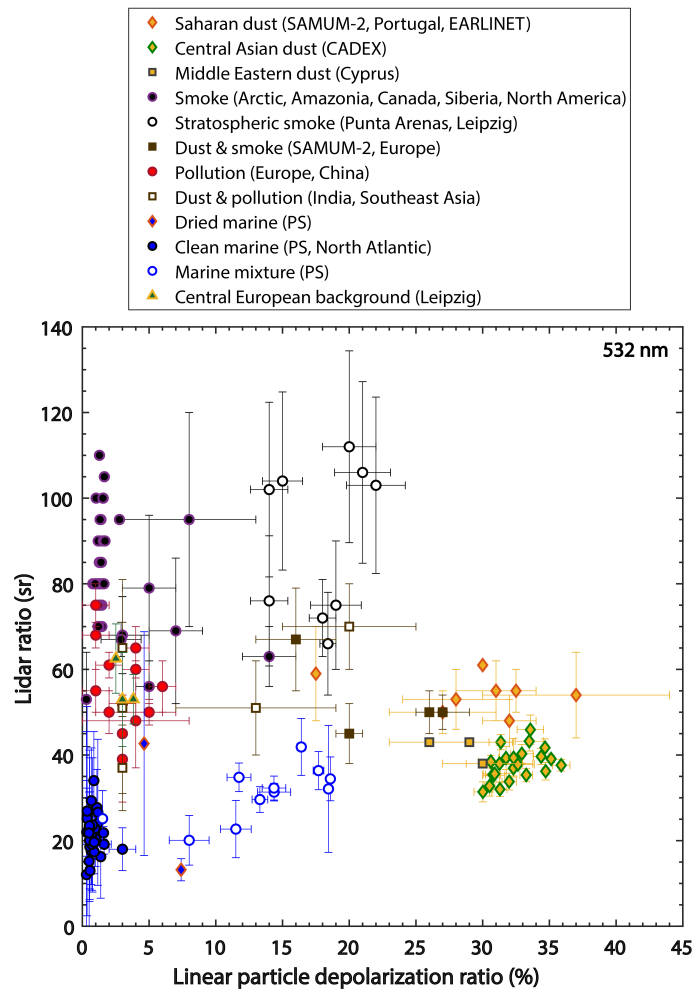
Figure 3.2: Intensive optical properties of different aerosol types, measured at 532 nm mainly within the PollyNET framework. Measurements from SAMUM–2 were performed with the Raman-polarization lidar POLIS (University of Munich).

The basis presented here corroborated previous findings and provided new insights regarding aerosol types. In addition, the basis is regularly updated aiming to provide a comprehensive climatology of the optical properties of the different aerosol types at different wavelengths. Currently, measurements (especially particle linear depolarization ratios) at the wavelength of 1064 nm are limited. However, we wish to extend the basis for that wavelength too, since spectrally resolved aerosol-type-linked intensive aerosol parameters are of great importance for aerosol typing. In a recent study, Haarig et al. provided the

first-ever lidar measurements of the lidar ratio and particle linear depolarization ratio for desert dust particles at all three lidar wavelengths (355, 532 and 1064 nm). The results are very promising and particularly helpful towards the extension of the basis. The full list of the aerosol types considered in the creation of the experimental basis presented in this work as well as the data sources are summarized in Tab. 3.1.

Table 3.1: Reference list for the optical properties of the aerosol types used in the creation of the experimental basis.

Aerosol type	Reference
Volcanic Ash	Groß et al. (2012), Baars et al. (2016) and references therein
Smoke	Müller et al. (2007), Baars (2011)[1], Tesche (2011)[2], Groß et al. (2013), Baars et al. (2016) and references therein, Giannakaki et al. (2016), Floutsi et al. (2021)[3], Haarig et al. (2018) and references therein, Ohneiser et al. (2021)[4]
Stratospheric Smoke	Haarig et al. (2018) and references therein, Ohneiser et al. (2020)[3]
Saharan Dust	Müller et al. (2007), Groß et al. (2011), Tesche (2011)[2], Kanitz et al. (2013), Baars et al. (2016) and references therein
Central Asian Dust	Hofer et al. (2020)[5]
Middle Eastern Dust	Müller et al. (2007)
Dust and Smoke	Tesche (2011)[2], Groß et al. (2013), Kanitz et al. (2013), Giannakaki et al. (2016), Baars et al. (2016) and references therein
Pollution	Groß et al. (2013), Kanitz et al. (2013), Giannakaki et al. (2016), Baars et al. (2016) and references therein, measurements at Leipzig (2018-2020)
Dust and Pollution	Müller et al. (2007), Baars et al. (2016) and references therein
Dried Marine	Bohlmann et al. (2018)
Clean Marine	Müller et al. (2007), Groß et al. (2011), Groß et al. (2013), Bohlmann et al. (2018)
Marine Mixture	Bohlmann et al. (2018)
Dust and Marine	Groß et al. (2011)
Central European Background	Müller et al. (2007), measurements at Leipzig (2018)

[1]EUCAARI, [2]SAMUM [3]DACAPO-PESO, [4]MOSAiC, [5]CADEX

As both Fig. 3.1 and Fig. 3.2 indicate, the different pure aerosol types and their mixtures spread over the 2D domain of the intensive particle parameters lidar ratio and linear depolarization ratio. On the bottom left corner, we observe the so-called marine aerosols, which

consist mainly of coarse, water-soluble spherical sea-salt particles. Due to their large size and low absorptivity, they produce low lidar ratios (of the order of 20–30 sr) and due to their sphericity, they show low depolarization ratios (lower than 5%). Ångström exponents are also low for this aerosol type (Cattrall et al., 2005; Müller et al., 2007; Burton et al., 2012). However, marine particles adopt a cubic-like shape when exposed to dry atmospheric conditions (with relative humidity lower than 45%) and can then produce higher depolarization ratios (Haarig et al., 2017; Bohlmann et al., 2018). This fact was not realized earlier, e.g., in Illingworth et al. (2015) or in typing schemes such as the one of CALIPSO (Sec. 2.2) and may have led to misclassification of this specific type (falsely classified as dust mixture). On the top left corner, we observe a cluster of particles identified as smoke/biomass-burning aerosol (BBA) particles. Smoke/BBA particles consist mainly of fine, nearly spherical, partly water-soluble and partly insoluble particles and are usually much more absorbing than other particle types. Typically, due to their small size and sphericity BBA particles produce low depolatization ratios (lower than 5%). The kind of burning (flaming or smoldering fires) influences the size of the smoke particles and the released amount of soot (insoluble material) and, thus, the optical properties of this aerosol type. The age of the observed smoke particles also is a crucial parameter determining their optical properties (Wandinger et al., 2002; Müller et al., 2005; Amiridis et al., 2009). Lidar ratios for aged smoke are higher than the ones for fresh smoke (Burton et al., 2012; Haarig et al., 2018), a fact that is well captured in the experimental basis (Fig. 3.1) given the wide range of lidar ratios for smoke (ranging from 50 up to 120 sr). Particles defined as pollution can be found near to the smoke/BBA cluster, exhibiting similar optical properties. The main difference in those two clusters is the absoptivity of the particles present, which is determined by the contribution of soot particles (Cattrall et al., 2005; Müller et al., 2007). Pollution aerosol consists of fine, nearly spherical and weakly to moderately absorbing particles. Typical lidar ratio values range from 50 to 70 sr depending on the soot content of the particles, while due to their sphericity they do not cause significant depolarization of the backscattered light. Another pure aerosol type comprises the dust particles. Their optical properties are very different compared to those of the other aforementioned aerosol types, which makes them, easily identifiable. Dust aerosol consists of coarse, non-spherical particles leading to moderate lidar ratios (40–60 sr, Mattis et al. (2002)), but significant depolarization of the backscattered light (20–35%, Freudenthaler et al. (2009)). Lidar ratios for desert dust aerosols vary regionally, e.g., Central Asian dust particles tend to have lower lidar ratios than Saharan dust particles, allowing thus even the identification of the source origin.

Mixtures of the aforementioned pure aerosol types appear to be scattered "in between" those pure clusters, exhibiting intermediate values of lidar ratio and particle linear depolarization ratio (Fig. 3.1, Fig. 3.2) and making thus the classification process much more

challenging. In such cases, the Ångström exponent can be used as an additional source of information, which serves as a proxy for aerosol particle size. Larger values of the extinction-related Ångström exponent indicate smaller particles and vice versa. This fact is illustrated in Fig. 3.3, which depicts two aerosol mixtures (isolated from Fig. 3.1) in a 3D domain of intensive optical properties (lidar ratio, particle linear depolarization ratio and extinction-related Ångström exponent). Both mixtures contain dust and one more aerosol component, which in one case is marine aerosol and in the other smoke aerosol (blue circle and brown square in Fig. 3.3, respectively). The two mixtures appear to have similar optical properties when considering lidar ratio and particle linear depolarization ratio, but the extinction-related Ångström exponent differs significantly. In the case of dust/marine mixture the extinction-related Ångström exponent has a value of 0.55 ± 0.22, indicating that the mixture consists of coarse particles. On the other hand, for the dust/smoke mixture the extinction-related Ångström exponent has a value of 1.10 ± 0.20, indicating the presence of fine particles in the mixture. Thus, the Ångström exponent should be considered for aerosol typing and classification purposes, if available, as a qualitative indicator of aerosol particle size.
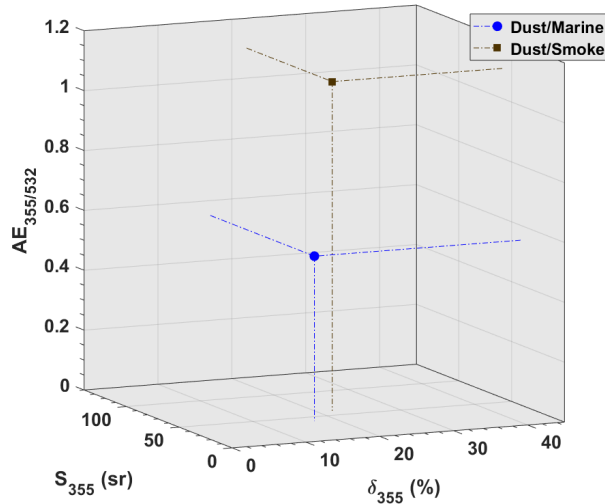
Figure 3.3: Lidar ratio (S), particle linear depolarization ratio (δ) at 355 nm and extinction-related Ångström exponent (AE) at the wavelength pair 355/532 nm for two aerosol mixtures. Errors have been omitted for readability.

3.2.2 Statistical analysis of intensive optical properties

Based on the datasets presented in Fig. 3.1 and Fig. 3.2, statistics of the intensive optical properties can be calculated. The statistics are an efficient way to summarize what has been

already discussed in Sec. 3.2.1, regarding the optical properties, and provide useful insights on the distribution of the data. For the pure aerosol clusters, the statistics are presented in the form of boxplots, in Fig. 3.4 and Fig. 3.5 for the 355 nm and 532 nm wavelengths, respectively. In the boxplots presented here, minimum and maximum values are indicated by the lower and upper whisker, median and mean values by the red lines and rhombuses, respectively. The lower part of each box indicates the 25% percentile and the upper part the 75% percentile. Red crosses represent the outliers.

As Fig. 3.4 indicates, the smoke cluster exhibits a positively skewed distribution for the lidar ratio and for the particle linear depolarization ratio. This means that more than half of the smoke data points have a lidar ratio lower than the mean lidar ratio (64 sr) and similarly a particle linear depolarization ratio lower than the mean depolarization ratio (3%). The data distribution for the marine and the dust lidar ratio as well as the marine and the pollution particle linear depolarization ratio is positively skewed. The only cases with a negatively skewed distribution are for the pollution lidar ratio and the dust particle linear depolarization ratio.

On the other hand, in Fig. 3.5, the distribution of the smoke and the dust lidar ratio is negatively skewed, while for the same categories the distribution of particle linear depolarization ratio is positively and negatively skewed, respectively. The data distribution for the pollution category is positively skewed for both lidar ratio and particle linear depolarization ratio. A symmetric data distribution (mean=median) is observed for the lidar ratio of marine particles, while for the same aerosol type the particle linear depolarization ratio distribution is positively skewed. The mean values of the intensive properties that were presented in Fig. 3.1 and Fig. 3.2, are listed in Table 3.2.

It should be noted that, since the data for each aerosol type are collected from various sources (see Tab. 3.1), naturally they exhibit a variation in the number of data points per aerosol category (i.e., not equal distribution of the data points for the different aerosol types). In addition, measurements were performed by different lidars and, therefore, not all intensive optical parameters are always available for all the observations (e.g., lidar ratio available only at one wavelength, or information on depolarization completely missing). Nevertheless, all available measurements were considered for the creation of the experimental basis and the statistical analysis presented here. To be more specific, the number of samples for the lidar ratio at 355 nm were 112, 38, 32 and 43 for smoke, marine, pollution and dust category, respectively. Similarly, 53, 37, 21, and 43 samples (same aerosol type order as above) were considered for the lidar ratio at 532 nm. The number of samples for the particle linear depolarization ratio at 355 and 532 nm were 203, 38, 28, 38 and 161, 34, 14, 34, respectively.
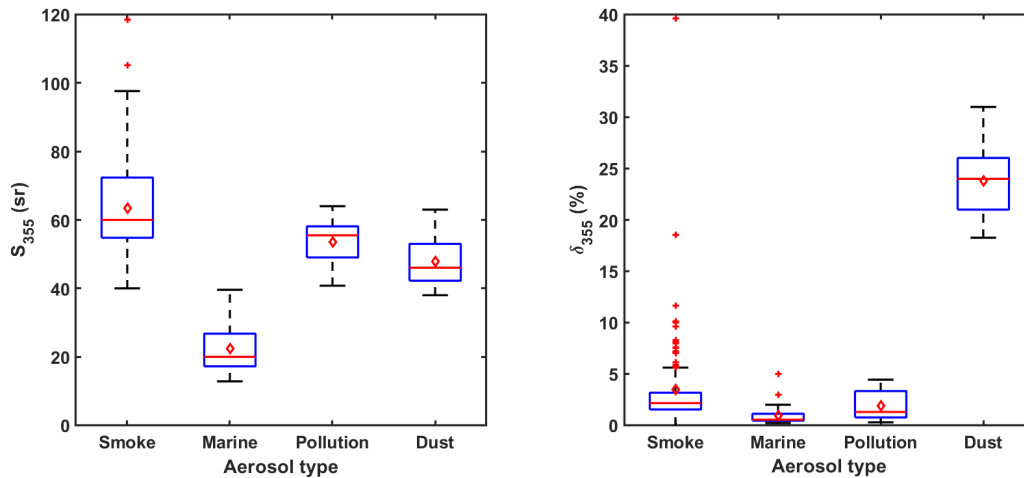
Figure 3.4: Statistics of the lidar ratio (left) and particle linear depolarization ratio (right) at 355 nm for the four most commonly found pure aerosol types. It should be noted that for the dust category all dust measurements, regardless of their source origin, have been considered.
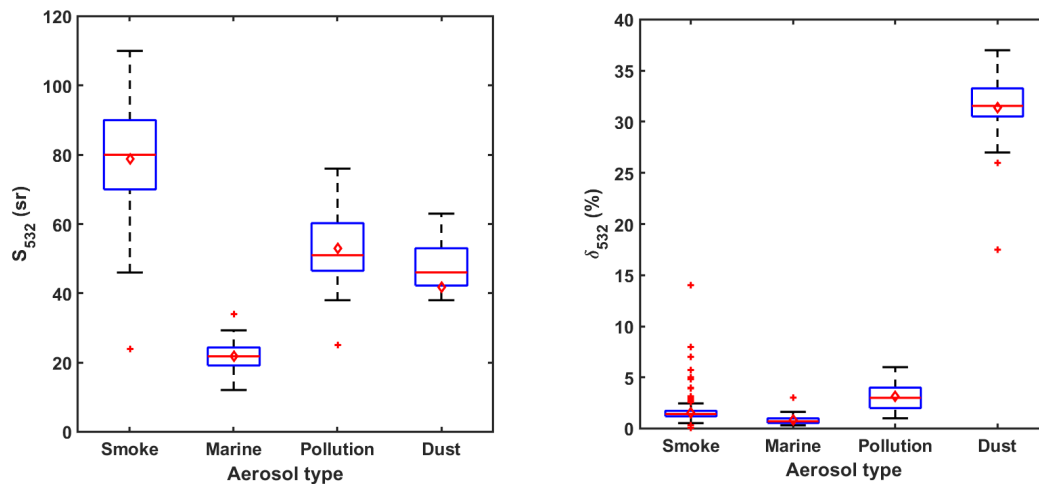
Figure 3.5: Statistics of the lidar ratio (left) and particle linear depolarization ratio (right) at 532 nm for the four most commonly found pure aerosol types. It should be noted that for the dust category all dust measurements, regardless of their source origin, have been considered.

3.3   Aerosol components

From the experimental basis presented here (Fig. 3.1 and 3.2), four pure and most commonly occurring aerosol types are distinguished and the rest are mixtures containing at least two of these pure aerosol types. The pure aerosol types are smoke, marine, (aerosol of) anthropogenic (origin) (e.g., pollution, urban aerosol) and desert dust aerosol. These

Table 3.2: Mean values of intensive properties at 355 and 532 nm, as derived from the respective experimental basis.

Aerosol type	S_{355} (sr)	δ_{355} (%)	S_{532} (sr)	δ_{532} (%)
Smoke	64	3	79	2
Marine	22	1	22	1
Pollution	54	1.8	53	3
Dust	48	24	42	31

aerosol types can be "translated" into aerosol components, with predefined microphysical and optical properties that reflect the aerosol types observed in nature. Even though there are other aerosol types that could also be considered as aerosol components, (e.g., volcanic ash or stratospheric smoke), we only consider these four to ensure consistency with HET-EAC and the experimental findings. Two components are fine-mode particles, namely the fine spherical absorbing (FSA) and the fine spherical non-absorbing (FSNA) component, while the other two are coarse-mode particles, namely the coarse spherical (CS) and the coarse non-spherical (CNS) component.

The FSA aerosol component, as the name indicates, describes aerosol of fine mode, spherical in shape and highly absorbing. It can therefore be associated with aerosol from direct combustion processes (e.g., smoke). When this component is present in an aerosol mixture, it can be used as a qualitative proxy for the amount of insoluble soot contained in that mixture. The FSNA component is from a microphysical perspective the same as FSA, but less absorbing (smaller imaginary part of the refractive index). This component can be associated with aerosol of anthropogenic origin (e.g., industrial pollution mixtures). CS describes coarse mode, water-soluble particles, usually associated with sea salt particles, under humid conditions. The spherical assumption holds with the exception of very dry conditions (as described in Haarig et al. (2017)). This component has a weak absorption efficiency. Finally, the CNS can be associated with dust particles. The shape of this component can be approximated by the spheroidal representation (Dubovik et al., 2006; Gasteiger et al., 2011). CNS is a component with relatively high absorption efficiency, although a great variety of values can be found in literature.

These four aerosol components are considered in HETEAC and in the retrieval scheme proposed below, because they can represent most of the aerosol conditions found in the cloud-free atmosphere. The main idea behind this is that any aerosol type or aerosol mixture can be assumed to be an external mixture of these components. In this case, external mixture means that there is no physical or chemical interaction between the particles of the different aerosol components (Hess et al., 1998). Having an estimation of the composition of an aerosol mixture, in terms of pre-defined aerosol components, allows for radiative-transfer calculations. In principle, the aerosol components could be extended to include

more aerosol types, but due to the lack of independent information that is currently leading to a much more ill-posed problem, here only the four aforementioned components are considered. Further details regarding the microphysical and optical properties of these four aerosol components are provided in Sec. 4.1.3.

Chapter 4

Methodology

In this chapter, a detailed description of the aerosol typing algorithm, along with its main constituents, output and additional products, is provided. In addition, the aerosol typing scheme is applied to four individual measurements from the experimental basis presented in Ch. 3. Each measurement corresponds to one of the basic aerosol categories, represented by one of the basic aerosol components considered in this typing scheme. The fifth measurement discussed here is a complex mixture of smoke, dust and marine particles. In all the aforementioned cases, the aerosol load has been well characterized beforehand and, therefore, the application of the typing scheme in these cases serves mainly as a proof of concept.

4.1 Retrieval methodology

The focus of this work is the retrieval of the aerosol components present in an aerosol mixture and their corresponding contributions, from ground-based remote-sensing lidar systems. This novel aerosol typing scheme applies the optimal estimation technique (see below) to a combination of lidar-derived intensive aerosol properties to estimate the contribution of the different aerosol components in terms of relative volume. The retrieval methodology along with its key features is presented in this chapter.

4.1.1 Optimal estimation method

In general, the optimal estimation method (OEM) is a nonlinear regression scheme applied to determine the statistically most-likely conditions to produce a given measurement, weighted against *a priori* knowledge of the system under investigation. Figure 4.1 gives a brief overview of the retrieval framework. In basic terms, the process starts with an initial guess of the state vector ($\boldsymbol{x}_\alpha$, i.e., the relative volume contribution per aerosol component),

which is iteratively modified such that the modelled intensive optical properties match those observed by the lidar ($\boldsymbol{y}$), resulting in the most probable estimated state ($\widehat{\boldsymbol{x}}$). The problem is ill-posed, since several different states may produce the same measurements and, therefore, *a priori* information is needed to constrain the state space (or solution space).
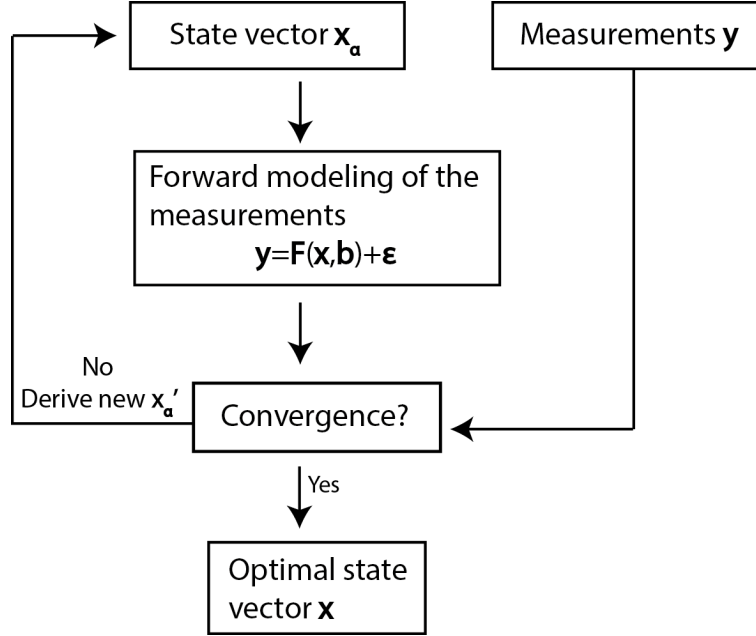
Figure 4.1: Generalized concept of the optimal estimation method.

As outlined in Rodgers (2000), optimal estimation solves the inverse problem:

$$\boldsymbol{y} = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{b}) + \boldsymbol{\epsilon}, \tag{4.1}$$

where $\boldsymbol{y}$ is a column vector describing the lidar measurements, $\boldsymbol{\epsilon}$ is the corresponding noise (error) of these measurements and $\mathbf{F(x,b)}$ the forward model that translates a state of the instrument and atmosphere, summarised by unknown parameters ($\boldsymbol{x}$) and known parameters ($\boldsymbol{b}$), into a simulated measurement. The probability that the system has a state ($\boldsymbol{x}$), given the measurements ($\boldsymbol{y}$), can be found by approximating the probability density function (PDF; P) for all quantities as Gaussian, and using Bayes' theorem:

$$- 2\ln P(\boldsymbol{x}|\boldsymbol{y}) = [\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{b})]^T \boldsymbol{S}_\epsilon^{-1} [\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{b})] + [\boldsymbol{x} - \boldsymbol{x}_\alpha]^T \boldsymbol{S}_\alpha^{-1} [\boldsymbol{x} - \boldsymbol{x}_\alpha], \tag{4.2}$$

where the covariance matrix $\boldsymbol{S}_\epsilon$ describes the measurement errors and $\boldsymbol{x}_\alpha$ the initial guess of the state (also referred to as the *a priori* state). The uncertainty in that expectation for the initial guess of the state is described by the *a priori* covariance $\boldsymbol{S}_\alpha$. The quantity $-2\ln P(\boldsymbol{x}|\boldsymbol{y})$ is hereafter referred to as the cost, as it measures the goodness of fit for a solution. The

iterational process converges where the costs are minimum, and that is the most probable state or the so-called optimal solution $\hat{\boldsymbol{x}}$. Evaluation ceases after the cost function decreases by much less than the number of measurements (i.e., 1/10 of the measurement's freedom degrees) or automatically after 30 iterations, which is considered a failure to converge. For clarity, the notation of Rodgers (2000) is followed throughout this dissertation.

4.1.2 State and measurement vectors

The quantities to be retrieved are represented by the state vector $\boldsymbol{x}$. This vector contains information on the relative contribution of the aerosol components (to the total aerosol volume) present in an aerosol mixture. As mentioned earlier (Sec. 3.3), every aerosol mixture is assumed to be consisting of a maximum of four basic aerosol components, two fine modes (spherical absorbing and spherical non-absorbing; FSA and FSNA, respectively) and two coarse modes (spherical and non-spherical; CS and CNS, respectively). Therefore, the state vector consists of four dependent and continuous (in the interval [0, 1]) variables. The basic aerosol components considered here enable the estimation of several extensive optical properties (see Sec. 4.1.5). More details on the aerosol components and their optical and microphysical properties are given in Sec. 4.1.3.

The initial guess of the state vector $\boldsymbol{x}_\alpha$, along with its covariance matrix $\boldsymbol{S}_\alpha$, is needed to start the iterational process. In the OEM, *a priori* information is used to regulate the ill-posed retrieval problem, making use of existing knowledge of the atmosphere and making the solution a physically meaningful result. The retrieval is thus sensitive to the choice of $\boldsymbol{x}_\alpha$, which for this retrieval scheme is the output of a decision tree (see Sec. 4.1.4), while the setting of $\boldsymbol{S}_\alpha$, which describes the estimated uncertainty of the initial-guess state vector elements as well as the correlation between the state vector elements, is a very controversial part of the OEM, since it might constrain the solution space inappropriately (by forming a subspace in which the solution must lie, with the correlations rejecting unrealistic solutions). In this study and for the given nature of $\boldsymbol{x}$, there is no manner in which an exact covariance $\boldsymbol{S}_\alpha$ can be derived (e.g., based on existing observations as in Foth and Pospichal (2017)) and hence, it is set in a parametric way. Since $\boldsymbol{S}_\alpha$ reveals the variability of the state vector, it is also used partly as a tuning parameter, meaning that the variances of the aerosol components are given with some margin with respect to the true values expected (Sec. 3.2), to maintain as much as possible the measurement information. The covariances (non-diagonal elements) are set to zero, since no relationship between the elements of the state vector is observed in nature, i.e., every basic aerosol component can coexist independently from the others and there is no correlation between them (e.g., the existence of absorbing particles does not prohibit or favor the existence of non-absorbing particles etc.). However, setting $\boldsymbol{S}_\alpha$ to a purely diagonal matrix translates to a stronger regularisation (Rodgers, 2000).

The quantities actually measured to retrieve the state vector $\boldsymbol{x}$ are represented by the measurement vector $\boldsymbol{y}$. This vector includes the intensive optical properties of the aerosol layer of interest, measured by lidar. To be more specific, the properties considered in this retrieval scheme are (in the order of appearance in the vector): the particle linear depolarization ratio at 355 nm, the lidar ratio at 355 nm, the extinction-related Ångström exponent (for the wavelength pair of 355/532 nm), the particle linear depolarization ratio at 532 nm, the lidar ratio at 532 nm and the backscatter-related color ratio (for the wavelength pair of 532/1064 nm). These intensive optical properties were chosen because they have high typing discrimination power (see Sec. 3.1). It should be noted that the measurement vector $\boldsymbol{y}$ can contain different combinations of the aforementioned properties, depending on the lidar capabilities and measurement availability (e.g., in the case of EarthCARE only the 355 nm particle linear depolarization ratio and lidar ratio will be present in the vector). Since measurements are made to a finite accuracy, the corresponding measurement errors are included in the diagonal matrix $\boldsymbol{S}_\epsilon$.

4.1.3 *A priori* aerosol components

The microphysical and optical properties of the four basic aerosol components considered in this retrieval scheme (i.e., FSA, CS, FSNA and CNS) are utilized as *a priori* information, which facilitates the construction of the forward model (discussed in Sec. 4.1.5). As mentioned already, the aerosol components along with their microphysical properties are the same as the ones defined in HETEAC (Sec. 2.5). The microphysical properties for the four aerosol components are summarized in Tab. 4.1. The effective radius (r_{eff}) is set to 0.14 µm for the fine-mode particles and to 1.94 µm for the coarse-mode particles. Together with the mode radius for the number and volume size distributions ($r_{0,\text{N}}$ and $r_{0,\text{V}}$, respectively), these quantities reflect the differences in the size of the particles. The real part of the refractive index at 355 nm (m_{R}) is the highest for the CNS component, slightly lower and comparable for the two fine-mode aerosol components (FSA and FSNA) and the lowest for the CS component. On the other hand, the imaginary part of the refractive index at 355 nm (m_{I}) is the highest for the FSA component, followed by the CNS and FSNA components, and finally the CS component. The refractive index reflects the chemical composition of the aerosol particles.

The scattering properties per unit particle volume (e.g., 1 µm^3/cm^3) of the four different aerosol components that are used directly in the OEM scheme (Eq. (4.3–4.6)) as *a priori* information are summarized in Tab. 4.2. For HETEAC, these properties have been calculated with a Mie scattering model for spherical particles (FSA, FSNA, CS) and with Dubovik's spheroid model (Dubovik et al., 2006) by assuming the spheroidal distribution of Gasteiger et al. (2011) for non-spherical particles (CNS). However, since HETEAC is optimized for

the wavelength of EarthCARE's ATLID (355 nm) and the calculated properties hold for idealized spheres and spheroids, the scattering properties used in the OEM-based aerosol typing scheme have been slightly adjusted to meet the experimental findings (Sec. 3.2). In particular, adjustments were made in the 532 nm backscatter coefficient per unit particle volume for the CNS aerosol component (a decrease of 0.013 in the case of CNS based on Saharan dust observations). In addition, the particle linear depolarization ratio (per unit particle volume) for the FSA, FSNA and CS aerosol components is zero in HETEAC, but adjusted for the OEM-based typing scheme to better reflect the observations (Sec. 3.2). The depolarization ratio (per unit particle volume) for the CNS aerosol component was also slightly decreased for 355 nm and increased for 532 nm in the OEM typing scheme. At the same time, the microphysical properties used in the OEM scheme (Tab. 4.1) were deliberately kept the same as in HETEAC, since the adjustments were small and consistency between these two approaches is important in terms of cross-validation of the algorithms and further support for the EarthCARE mission.

Table 4.1: Microphysical properties of the four basic aerosol components used to simulate multimodal particle distributions in HETEAC and OEM. r_{eff} stands for the effective radius, $r_{0,\text{N}}$ and $r_{0,\text{V}}$ for the mode radius of the number and volume size distributions, respectively, σ for the mode width (variance) and m_{R} and m_{I} represent the real and imaginary part of the refractive index, respectively, at two distinctive wavelengths 355 and 532 nm, respectively.

	Fine mode absorbing	Fine mode non-absorbing	Coarse mode spherical	Coarse mode non-spherical
r_{eff} (µm)	0.14	0.14	1.94	1.94
$r_{0,\text{N}}$ (µm)	0.07	0.07	0.788	0.788
$r_{0,\text{V}}$ (µm)	0.1626	0.1626	2.32	2.32
σ	0.53	0.53	0.6	0.6
m_{R}	1.50, 1.50	1.45, 1.44	1.36, 1.36	1.54, 1.53
m_{I}	4.3e−2, 4.3e−2	1.0e−5, 1.0e−3	1.0e−8, 1.0e−9	6.0e−3, 3.0e−3
Shape representation	Spherical	Spherical	Spherical	Spheroid distribution (Gasteiger et al., 2011)

Figure 4.2(a) shows two of the resulting intensive optical properties (at 355 nm) for the four basic aerosol types and their multimodal external aerosol mixtures obtained by using the optical properties of Tab. 4.2. The different aerosol mixtures have been calculated with a step of 5% in terms of relative volume (per aerosol component), while the sum of all the aerosol components is always equal to 100%. The external mixing assumption leads to straight connecting lines in the optical parameter space. As it can be seen, most of the multimodal mixtures produce lidar ratios between 40 and 80 sr and depolarization ratios

Table 4.2: Optical properties of the four basic aerosol components at two wavelengths (355 and 532 nm). α^* and β^* are calculated per unit particle volume (1 µm^3/cm^3).

Aerosol component	α^* (Mm^{-1})		β^* (Mm^{-1}sr^{-1})		δ^*		S (sr)	
	355	532	355	532	355	532	355	532
FSA	10.7	6.45	0.09	0.07	0.024	0.024	117.3	93.8
CS	0.88	0.94	0.051	0.049	0.035	0.035	17.4	19.2
FSNA	9.61	5.03	0.16	0.08	0.033	0.033	60.9	59.3
CNS (Saharan)	0.93	0.97	0.016	0.018	0.24	0.33	57.9	55.0
CNS (Asian)	0.93	0.97	0.022	0.024	0.25	0.28	43.3	40.0

between 2.4% and 5%. Such values are indeed most often observed in nature. It should be noted that a considerable change of the optical parameters is observable only when one of the components starts dominating the mixture. In addition, a large contribution of dust is needed to cause a considerable particle linear depolarization ratio. Similarly, very large or very small lidar ratios are produced only when the small, spherical, strongly absorbing component or the coarse spherical component dominates, respectively. Similar behavior is observed also for the intensive optical properties at 532 nm (see Fig. 4.2(b)).
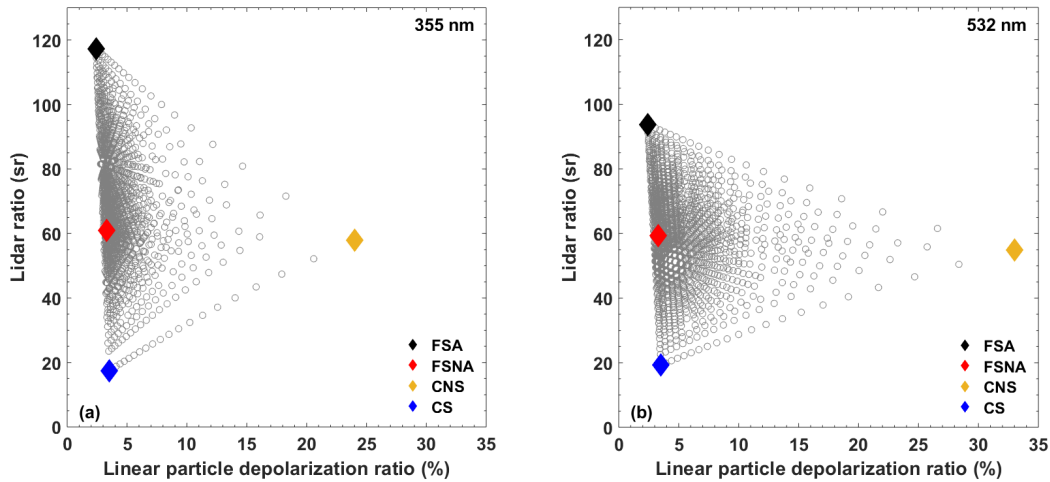
Figure 4.2: Simulated values of lidar ratio versus particle linear depolarization ratio at (a) 355 nm and (b) 532 nm, for multimodal mixtures (grey open circles) of four basic aerosol components (rhombuses) based on Tab. 4.2.

Similarly to Fig. 4.2, Fig. 4.3 shows the extinction-related Ångström exponent versus the particle linear depolarization ratio (a) and lidar ratio (b), both at 355 nm for the four basic aerosol types and their multimodal external aerosol mixtures. Again, the different aerosol mixtures have been calculated with a step of 5% in terms of relative volume (per

aerosol component). The 2D spaces created by the different optical parameters are different compared to the respective ones in Fig. 4.2. The coarse-mode aerosol components (i.e., CS and CNS) exhibit extinction-related Ångström exponent values of around zero, while the fine-mode aerosol components (i.e., FSA and FSNA) show extinction-related Ångström exponent values above unity. It can be seen that most multimodal external aerosol mixtures have extinction-related Ångström exponent values ranging between 0.8 and 1.4. A small Ångström exponent requires high relative volume contributions of coarse-mode aerosol particles and therefore not so many aerosol mixtures (simulated) have low Ångström exponent values.
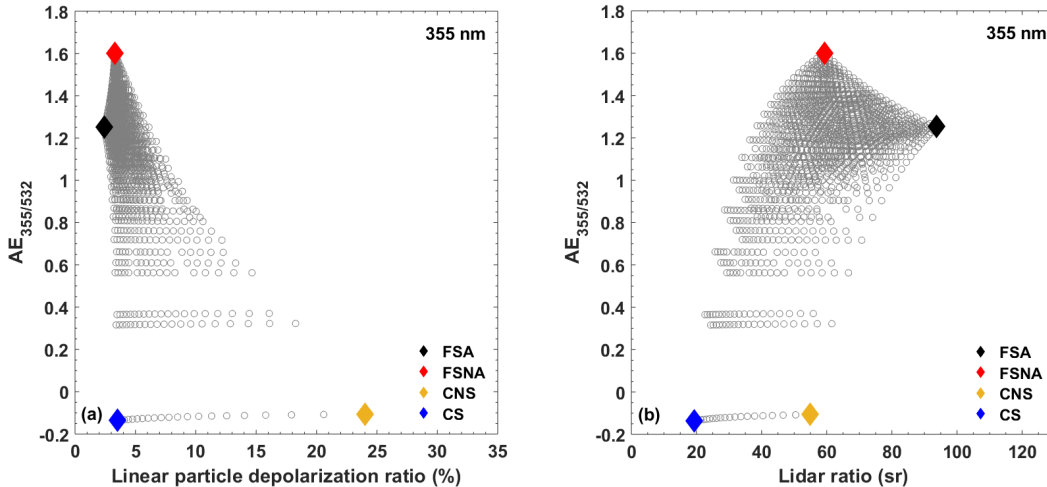
Figure 4.3: Same as Fig. 4.2, but for the extinction-related Ångström exponent versus the particle linear depolarization ratio (a) and lidar ratio (b) at 355 nm.

4.1.4 Decision tree

As mentioned in Sec. 4.1.2, the initial guess $\boldsymbol{x}_\alpha$ that is used to kick off the OEM is an output of a decision tree (Fig. 4.4). The decision tree is created based on aerosol optical parameters with high discriminatory power from the experimental basis presented in Sec. 3.2. To be more specific, the root node (topmost node) contains the mean aerosol layer values of S and δ (either at 355 or 532 nm). The first splitting parameter is the particle linear depolarization ratio (δ), which has been already found to have the highest discriminatory power (Sec. 2.3, Burton et al. (2012)). As Fig. 4.4 indicates, the highest value of particle linear depolarization ratio considered is 0.35. Aerosol particles exhibiting particle linear depolarization ratios higher that 0.35, such as volcanic-ash, are not considered in the OEM-based aerosol typing scheme and, therefore, are excluded from the decision tree. Further splitting is then done by the lidar ratio S. In the end, the terminal nodes contain the different labels along with

the corresponding values of the initial guess of the state vector $\boldsymbol{x}_\alpha$. While the terminal node labels contain the names used for the aerosol components, they should not be confused with the components themselves and therefore are accompanied with an asterisk (*). The asterisk indicates which aerosol component (or components in the case of mixtures) should be considered as dominant in the initial guess of the state vector $\boldsymbol{x}_\alpha$. For instance, the label CS* indicates that based on the lidar observations, an initial guess of the state vector where CS particles are dominant should be considered. The different sets of the initial guess of the state vector are provided in Tab. 4.3.
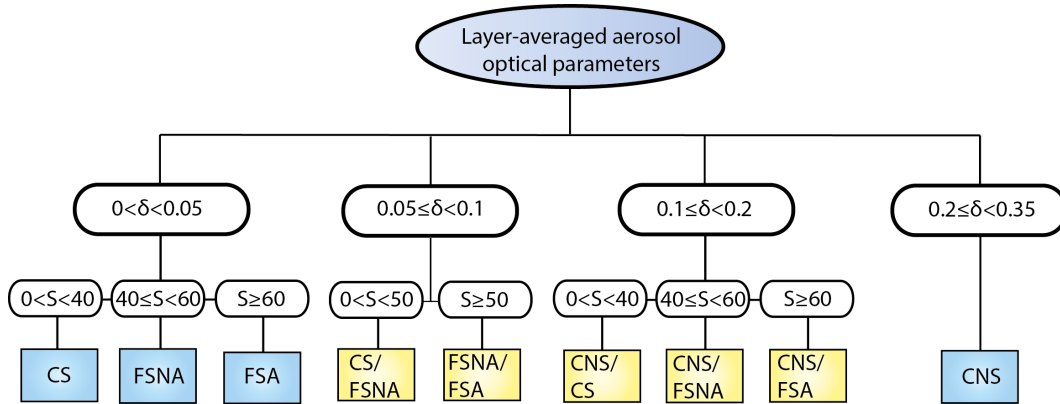
Figure 4.4: Decision tree used for the determination of the initial guess of the state vector $(\boldsymbol{x}_\alpha)$ in the OEM typing scheme. Unit for S is sr.

Table 4.3: Sets of initial guesses for the state vector $(\boldsymbol{x}_\alpha)$, based on the terminal node label of the decision tree (Fig. 4.4).

Terminal node label	FSA	CS	FSNA	CNS
CS*	0.05	0.85	0.05	0.05
FSNA*	0.05	0.05	0.85	0.05
FSA*	0.85	0.05	0.05	0.05
CS*/FSNA*	0.0	0.5	0.5	0.0
FSNA*/FSA*	0.5	0.0	0.5	0.0
CNS*/CS*	0.0	0.7	0.0	0.3
CNS*/FSNA*	0.0	0.0	0.7	0.3
CNS*/FSA*	0.7	0.0	0.0	0.3
CNS*	0.0	0.0	0.0	1.0

The relative volume contributions of each aerosol component used as the initial guess of the state vector $\boldsymbol{x}$ (Tab. 4.3) are chosen to some extent arbitrarily (i.e., there is no way to determine the exact relative volume contributions), but at the same time they clearly reflect the dominant components in the state vector. As a general rule, when the decision

tree returns a terminal node label with only one aerosol component as dominant, then its relative volume contribution is set to 0.85 and all the other components are set to 0.05 (to reach a total relative volume of 1), with the exception of the CNS component. When the CNS aerosol component is dominant, then its relative volume contribution is set to 1. Mixtures with two components are treated in a way that takes into account the particle linear depolarization ratio values. Mixtures that do not include the CNS component (i.e., associated with depolarizing particles) have equal contributions of the two dominant components (relative volume contributions of 0.5 for each) and the other two components are set to 0. In mixtures where the CNS component appears as one of dominant two components, the relative volume contributions of these components is distributed as 0.7 and 0.3, with the lowest contribution being assigned to the CNS component. That is because the CNS component is the only component with such high value of particle linear depolarization ratio and depending on the case, the OEM might need to adjust (through the corresponding Jacobian) quite a lot the initial guess before convergence is met.

4.1.5 Forward model

As mentioned in the OEM introduction (Sec. 4.1.1), a forward model $\mathbf{F}(\mathbf{x},\mathbf{b})$ is needed for the transformation of the state space to the observation space via utilizing the *a priori* information ($\mathbf{b}$). In this retrieval scheme, aerosol mixtures comprise four different aerosol components and therefore a multimodal representation is needed. The multimodal representation takes into account different aerosol size modes, refractive indexes and particle shapes for each mode (e.g., mixtures of dust and smoke). Currently, appropriate scattering models are missing. For instance, the Dubovik model (Dubovik et al., 2006) can handle bimodal size distributions, but only with constant refractive index and shape distribution for all sizes. Thus, to obtain the intensive optical parameters of multimodal aerosol compositions, mixing rules have to be applied. For each component, the scattering properties per unit particle volume are predefined (Tab. 4.2, Sec. 4.1.3). The optical parameters of interest are then derived from the extinction and backscatter coefficients per unit volume (α_λ^*, β_λ^*, respectively), the particle linear depolarization ratio (δ_λ^*) of the aerosol components (indexed with j) and the relative volume contribution (x) of all modes as follows (below, λ has been omitted from the first two equations for the sake of clarity, as the scattering properties are at the same wavelength):

$$\delta = \frac{\sum x_j \beta_j^* \frac{\delta_j^*}{1+\delta_j^*}}{\sum x_j \beta_j^* \frac{1}{1+\delta_j^*}}, \tag{4.3}$$

$$S = \frac{\sum x_j \alpha_j^*}{\sum x_j \beta_j^*}, \tag{4.4}$$

$$\mathring{A} = \frac{\ln\left(\frac{\sum x_j \alpha_{j,\lambda_1}^*}{\sum x_j \alpha_{j,\lambda_2}^*}\right)}{\ln\left(\frac{\lambda_1}{\lambda_2}\right)}, \tag{4.5}$$

$$C_{\beta_{\lambda_1/\lambda_2}} = \frac{\sum x_j \beta_{j,\lambda_1}^*}{\sum x_j \beta_{j,\lambda_2}^*}. \tag{4.6}$$

All four equations presented above (Eq. (4.3–4.6)) assemble the forward model. The first forward model equation (Eq. (4.3)) simulates the particle linear depolarization ratio, the second one the lidar ratio (Eq. (4.4)), the third one the extinction-related Ångström exponent (Eq. (4.5)) and the last one the backscatter-related color ratio (Eq. (4.6)). Depending on the available measurements in the measurement vector (**y**), the forward model is adjusted accordingly (to simulate only the available measurements). To be more specific, there are six different predefined forward model configurations or retrieval modes that are summarized in Tab. 4.4. The different retrieval modes provide great flexibility in terms of available input and ensure a retrieval with a minimum amount of two parameters per measurement (both at the same wavelength). It should be noted that the nature of the algorithm is such that the forward operator can be easily modified and extended according to the user needs or the application considered. This feature is of great importance, especially for ground-based lidars, where the channel configuration might differ from instrument to instrument.

Table 4.4: Available forward model configurations.

Retrieval mode	Parameters
1	δ_{355}, S_{355}
2	δ_{532}, S_{532}
3	$\delta_{355}, S_{355}, \mathring{A}_{355/532}$
4	$\delta_{532}, S_{532}, C_{\beta_{532/1064}}$
5	$\delta_{355}, S_{355}, \delta_{532}, S_{532}$
6	$\delta_{355}, S_{355}, \mathring{A}_{355/532}, \delta_{532}, S_{532}, C_{\beta_{532/1064}}$

4.1.6   Optimum solution and convergence

The variational approach for obtaining an optimal estimate of the atmospheric state $\widehat{\boldsymbol{x}}$, given a measurement vector **y** and a forward model $\mathbf{F}(\mathbf{x},\mathbf{b})$, is performed by minimizing a

cost function of the form (Rodgers, 2000):

$$J(\widehat{\boldsymbol{x}}) = J_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{x}}) + J_{\boldsymbol{y}}(\widehat{\boldsymbol{x}}) + J_{\mathrm{con}}(\widehat{\boldsymbol{x}}), \qquad (4.7)$$

where $J_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{x}})$ represents the initial guess costs (or *a priori* costs), $J_{\boldsymbol{y}}(\widehat{\boldsymbol{x}})$ the observation costs and $J_{\mathrm{con}}(\widehat{\boldsymbol{x}})$ penalty terms to ensure physically meaningful retrievals of relative volume per aerosol component. While from a mathematical point of view relative volume contributions below 0% or above 100% are feasible, this is not the case from a physical point of view. Therefore, $J_{\mathrm{con}}(\widehat{\boldsymbol{x}})$ adds a penalty, if the retrieval produces a relative volume per aerosol component that exceeds the interval 0 to 1, where the variables are continuous. The function is defined as:

$$J_{\mathrm{con}}(\widehat{\boldsymbol{x}}) = \begin{cases} 0 & \text{for } 0 \leq \boldsymbol{x}_{\mathrm{j}} \leq 1 \\[2ex] \zeta |(\boldsymbol{x}_{\mathrm{j}})|^3 & \text{else,} \end{cases} \qquad (4.8)$$

where $\boldsymbol{x}_{\mathrm{j}}$ are the elements of the state vector (relative volume contribution of the different aerosol components) and ζ is a constant that is proportional to the strictness of the constraint. Here it is set to a value large enough (10^6) to avoid relative volumes exceeding the interval boundaries. In addition to the penalty terms above, if the retrieved relative volume of an aerosol type is smaller than 0, then it is automatically set to 0, and if the total relative volume contribution (sum of the relative volume contribution per aerosol component) is greater than 1, then the state vector is normalized, i.e., each value of the state vector $\boldsymbol{x}_{\mathrm{j}}$ is divided by the sum of the relative volume of the four different aerosol components. There is no constraint in place in the case of a total relative volume contribution that is less than 1. In such cases, the remaining contribution is characterized as uncategorized/unknown aerosol. Most usually, when a total relative volume contribution less than 1 appears in the optimal-solution state, it is a result of the normalization of the state vector (in the previous step).

Expanding Eq. (4.7) we get:

$$J(\widehat{\boldsymbol{x}}) = [\boldsymbol{x} - \boldsymbol{x}_{\alpha}]^T \boldsymbol{S}_{\alpha}^{-1} [\boldsymbol{x} - \boldsymbol{x}_{\alpha}] + [\boldsymbol{y} - \boldsymbol{F}(\widehat{\boldsymbol{x}})]^T \boldsymbol{S}_{\epsilon}^{-1} [\boldsymbol{y} - \boldsymbol{F}(\widehat{\boldsymbol{x}})] + J_{\mathrm{con}}(\widehat{\boldsymbol{x}}). \qquad (4.9)$$

The optimum solution can be found iteratively using the Levenberg-Marquardt method (LM), which is a combination of the gradient/steepest descent and Gauss-Newton methods:

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + [(1 + \gamma_i) \boldsymbol{S}_{\alpha}^{-1} + \boldsymbol{K}_i^T \boldsymbol{S}_{\epsilon} \boldsymbol{K}_i + \ddot{\boldsymbol{J}}_{\mathrm{con}}]^{-1} \times \{ \boldsymbol{K}_i^T \boldsymbol{S}_{\epsilon} [\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_i)] - \boldsymbol{S}_{\alpha}^{-1} [\boldsymbol{x}_i - \boldsymbol{x}_{\alpha}] + \dot{\boldsymbol{J}}_{\mathrm{con}} \}, \ (4.10)$$

with i being the iteration step and the dots over $J_{\mathrm{con}}(\widehat{\boldsymbol{x}})$ denoting the first and second derivative with respect to the state vector. $\boldsymbol{K}_{\mathrm{i}}$ is the weighting function matrix, or Kernel

or Jacobian (from now on Jacobian), defined as $\boldsymbol{K}_i = \frac{\partial \boldsymbol{F}(\hat{\boldsymbol{x}})}{\partial \hat{\boldsymbol{x}}}$ and calculated analytically for the lidar ratio and the linear particle depolarization ratio and numerically for the remaining quantities of the forward model, by perturbing the corresponding variable of the state vector by 10^{-3}. The LM parameter (γ) is a factor that minimizes the cost function (Eq. (4.9)). When $\gamma \to 0$, the solution tends to the Gauss-Newton solution ($\gamma = 0$), while when $\gamma \to \infty$, the solution tends to the steepest-descent solution, allowing thus the solution to leave a local minimum towards a global minimum. In this study, the initial value for the γ parameter is 2 and it is increased by a factor of 10, if the cost function in the current iteration step is greater than the one in the previous step ($\boldsymbol{J}(\boldsymbol{x}_{i+1}) \geq \boldsymbol{J}(\boldsymbol{x}_i)$), and it is reduced by a factor of 2, if the cost function is smaller ($\boldsymbol{J}(\boldsymbol{x}_{i+1}) < \boldsymbol{J}(\boldsymbol{x}_i)$). In retrospect, the LM method was found to converge not faster but more reliably than a Gauss-Newton iteration and, hence, it is preferred in this study.

The iteration procedure of Eq. (4.10) starts with the initial guess of the state vector ($\boldsymbol{x}_i = \boldsymbol{x}_\alpha$) and is repeated until the following criterion is fulfilled:

$$[\boldsymbol{F}(\boldsymbol{x}_{i+1}) - \boldsymbol{F}(\boldsymbol{x}_i)]^T \boldsymbol{S}_{\delta\hat{y}}^{-1} [\boldsymbol{F}(\boldsymbol{x}_{i+1}) - \boldsymbol{F}(\boldsymbol{x}_i)] \ll d_f, \tag{4.11}$$

where $\boldsymbol{S}_{\delta\hat{y}} = \boldsymbol{S}_\epsilon (\boldsymbol{K}\boldsymbol{S}_\alpha\boldsymbol{K}^T + \boldsymbol{S}_\epsilon)^{-1}\boldsymbol{S}_\epsilon$ is the covariance matrix between the measurement ($\boldsymbol{y}$) and $\boldsymbol{F}(\hat{\boldsymbol{x}})$ and d_f describes the degrees of freedom of the measurement, i.e., the number of independent observables (see. Tab. 4.4). In the algorithm, the "much smaller" mathematical operator ($\ll$) translates into $1/10$ of the degrees of freedom of the measurement. The degrees of freedom of the measurement are defined as $d_f = d_s + d_n$, where the first term is attributable to the state vector and the second term corresponds to the noise, and they can range between two and six degrees (depending on the input parameters).

Finally, the covariance matrix of the optimally estimated state vector (*a posteriori*) is calculated as follows:

$$\hat{\boldsymbol{S}} = (\boldsymbol{K}^T\boldsymbol{S}_\epsilon\boldsymbol{K} + \boldsymbol{S}_\alpha)^{-1}. \tag{4.12}$$

The diagonal elements of $\hat{\boldsymbol{S}}$ are the retrieval errors of the final optimal state vector $\hat{\boldsymbol{x}}$. A pseudo-code summarizing the whole OEM procedure as described above is provided in Sec. 4.3. In addition, a detailed diagram of the OEM-based typing algorithm is shown there (Fig. 4.5).

4.1.7   Statistical significance of the retrieved state

Once the iteration has converged, a Pearson's chi-squared (χ^2) test must be carried out to ensure correct convergence (i.e., avoid convergence at a false minimum). This is done by comparing the forward-modelled measurements at the optimal state $\boldsymbol{F}(\hat{\boldsymbol{x}})$ with the mea-

surement vector $\boldsymbol{y}$, along with the corresponding error covariance matrix $\boldsymbol{S}_{\delta\widehat{y}}$:

$$\chi^2 = [\boldsymbol{F}(\widehat{\boldsymbol{x}}) - \boldsymbol{y}]^T \boldsymbol{S}_{\delta\widehat{y}}^{-1} [\boldsymbol{F}(\widehat{\boldsymbol{x}}) - \boldsymbol{y}]. \tag{4.13}$$

This statistical test is usually used for outlier identification (i.e., a retrieved state that does not belong to a Gaussian distribution). All retrievals for which the computed χ^2 is greater than a threshold value χ_{thr} are discarded and all the rest, for which $\chi^2 \leq \chi_{\text{thr}}$, are accepted and further analysed. χ_{thr} is calculated for a 5% probability that χ^2 is greater than the threshold for a theoretical χ^2 distribution with d_{f} degrees of freedom (Chi-Square Table, 2008). A significance level of 95% is used in this study. In other words, if the estimated retrieved state is found to be statistically significant within the 95% significance level, there is a 5% chance of not being true.

4.2 Additional products

Once the OEM aerosol typing scheme has successfully converged to a statistically significant solution (retrieved state vector), the final product, which is the relative volume contribution of each aerosol component, can be used in combination with the microphysical and optical properties of the aerosol components (*a priori*) to obtain further products. At the moment, additional products included in this algorithm and discussed below are the extinction and backscatter coefficient per component, the volume and number concentration per component, the effective radius of the mixture and the refractive index of the mixture.

Extinction and backscatter per component

With the extinction and backscatter coefficient per unit particle volume of the different aerosol components known as *a priori* (α^* and β^* in Tab. 4.2, respectively), once the relative volume contribution x_{j} of each aerosol component j is retrieved with the OEM typing scheme, the fraction of the extinction and backscatter coefficient (α_{j}' and β_{j}', respectively) attributable to a specific aerosol component can be calculated:

$$\alpha_{\text{j}}' = \alpha_{\text{j}}^* x_{\text{j}} \tag{4.14}$$

and

$$\beta_{\text{j}}' = \beta_{\text{j}}^* x_{\text{j}}. \tag{4.15}$$

The total extinction and backscatter coefficients per unit volume then follow as the sum of the respective fractions of all the aerosol components:

$$\alpha_{\text{tot}}^* = \sum \alpha_j'$$ (4.16)

and

$$\beta_{\text{tot}}^* = \sum \beta_j'.$$ (4.17)

When the the total extinction and backscatter coefficients measured by the lidar (α_{tot} and β_{tot}, respectively) are known, the absolute extinction and backscatter fraction per aerosol component can be calculated:

$$\alpha_j = \alpha_j' \frac{\alpha_{\text{tot}}}{\alpha_{\text{tot}}^*}$$ (4.18)

and

$$\beta_j = \beta_j' \frac{\beta_{\text{tot}}}{\beta_{\text{tot}}^*}.$$ (4.19)

Volume and number concentration per component

When the total extinction coefficient measured by the lidar (α_{tot}) and the total extinction coefficient per unit volume (α_{tot}^*) are known, we can determine the volume concentration per component (V_j):

$$V_j = x_j \frac{\alpha_{\text{tot}}}{\alpha_{\text{tot}}^*}.$$ (4.20)

The number concentration per component (N_j) can then be calculated with a simple conversion from (V_j), using the corresponding microphysical properties from Tab. 4.1:

$$N_j = \frac{V_j}{\frac{4}{3}\pi r_{0,V_j}^3 \exp\left(-\frac{9}{2}\sigma_j^2\right)}.$$ (4.21)

The total volume and number concentration are calculated from the sum of the volume and number concentration per component, respectively:

$$V_{\text{tot}} = \sum V_j$$ (4.22)

and

$$N_{\text{tot}} = \sum N_j.$$ (4.23)

Surface area per component

The particle surface area of the different aerosol components is calculated as follows:

$$A_{\rm j} = 4\pi r_{A_{\rm j}}^2 \exp(-2\sigma_{\rm j}^2) N_{\rm j}. \tag{4.24}$$

The term $r_{A_{\rm j}} = r_{0,N_{\rm j}} \exp(2\sigma_{\rm j}^2)$ is directly linked to the fine- or coarse-mode width and radius (see Tab. 4.1). The total surface area can be calculated as the sum of the respective surface areas of all the aerosol components:

$$A_{\rm tot} = \sum A_{\rm j}. \tag{4.25}$$

Effective radius of the mixture

With the total surface area of the aerosol particles known, the effective radius can then be calculated as follows:

$$R_{\rm eff} = \frac{3V_{\rm tot}}{A_{\rm tot}}. \tag{4.26}$$

Refractive index of the mixture

The refractive index $m = n - ki$ for an aerosol mixture can be calculated by applying a volume mixing rule and using the predefined values of the complex refractive index of each aerosol component (given in Tab. 4.1) and the relative volume contributions of the components as derived from the OEM:

$$m = \frac{\sum n_{\rm j} x_{\rm j}}{\sum x_{\rm j}} - \frac{\sum k_{\rm j} x_{\rm j}}{\sum x_{\rm j}} i. \tag{4.27}$$

Error propagation in additional products

The error propagation from the OEM output to the additional products is carried out by Monte Carlo simulations (MCS). The retrieval errors are randomly distributed to the corresponding OEM-derived aerosol components, creating in total 50000 different and random MCS. Before moving to the next step, two filters are applied to these 50000 MCS. The first filter checks for negative relative volume contributions in the aerosol components. Negative volume contributions can occur, when the retrieval error is larger than the component's relative volume. When that occurs, even if only one of the four aerosol components has a negative contribution, then the whole group of components is discarded. The second filter checks for the total relative volume of the four components, which should not exceed 1. If it does, then similarly to the first filter, the whole group of components is discarded. After filtering out the unrealistic MCS, each additional product is then calculated for the remaining MCS, along with the respective mean and standard deviation. As expected, the mean

is found to be close to the value calculated using the original OEM output and, therefore, the standard deviation is used as the corresponding error.

4.3   OEM pseudocode

A step-by-step description of the algorithm follows. The program code is written in the proprietary programming language MATLAB (MATLAB, 2018).

1. Define the input measurements along with the corresponding error ($\boldsymbol{y}$ and $\boldsymbol{\epsilon}$, respectively).

2. The measurements are used as an input for a decision tree that returns a first guess for the state vector ($\boldsymbol{x}_\alpha$).

3. For the given first guess of the state vector calculate the modelled measurements $\boldsymbol{F}(\boldsymbol{x}_\alpha)$, the Jacobian $\boldsymbol{K}$ and the covariance matrix between the measurement and the modelled measurement $\boldsymbol{S}_{\delta\widehat{y}}$.

4. For a maximum of 30 iteration steps (start of the iterative process, default: Levenberg-Marquardt method)

    (a) Initialize γ parameter. Calculate the next state vector $\boldsymbol{x}_i$, new Jacobian $\boldsymbol{K}_i$, new modelled measurements $\boldsymbol{F}_i$ for the given $\boldsymbol{x}_i$, new covariance matrix between the measurement and the modelled measurement.

    (b) Calculate the cost function $\boldsymbol{J}(\boldsymbol{x}_i)$.

    (c) If the cost function is greater than or equal to the one calculated in the previous step, increase the γ parameter by a factor of 10. If it is smaller, decrease the γ parameter by a factor of 2.

    (d) Calculate the convergence criterion. Once its value is much lower than the $\mathrm{d_f}$ of the measurement (lower than $\mathrm{d_f}/10$), stop the iteration.

5. Return the optimal state vector $\widehat{\boldsymbol{x}}$ along with the retrieval errors $\widehat{\boldsymbol{S}}$.

6. Perform χ^2 test. Discard the results, if the retrieved state is not statistically significant within the desired significance level. Consider starting the iterative process again with a new $\boldsymbol{x}_\alpha$ or with different *a priori* settings.

Fig. 4.5 provides a more detailed overview of the OEM (compared to Fig. 4.1). Details regarding the different symbols and terms used in the figure are provided in the text above.
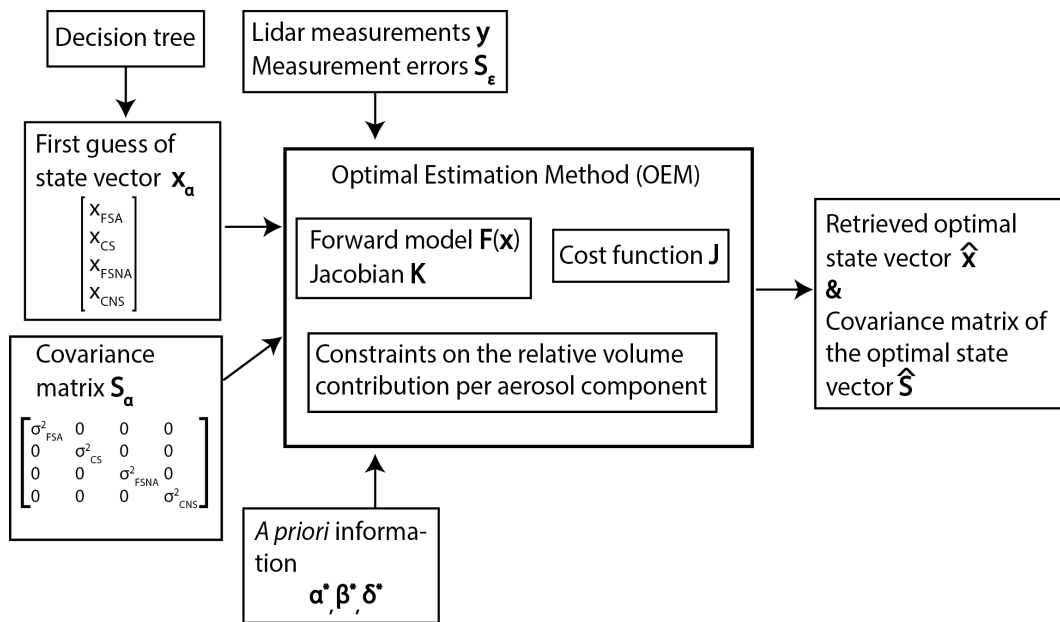
Figure 4.5: Detailed illustration of the optimal estimation method.

4.4   Evaluation of the methodology

In this section, four measurements from the experimental basis (presented in Ch. 3) and one adapted from Bohlmann et al. (2018) are analysed with the OEM-based typing scheme to evaluate the proposed methodology. The measurements that were selected from the experimental basis represent the four basic pure aerosol types, meaning that in every case there is a clear aerosol component predominance. The first case (Sec. 4.4.1) is used to discuss the typing algorithm in detail and serves as a walk-through example, while the other three cases (Sec. 4.4.2, 4.4.3, 4.4.4) show the applicability of the algorithm only briefly. The fifth case (Sec. 4.4.5) is an example for a complex aerosol mixture and demonstrates the ability of the algorithm to resolve correctly the different aerosol components present in that mixture. For every case, the observed aerosol type/mixture was known beforehand and, therefore, these cases serve as a proof of concept.

Table 4.5 summarizes the cases discussed in the following sections in terms of date of the measurement, aerosol types (known from the experimental basis and literature), OEM retrieval mode and output. It can be seen already that the typing scheme is very flexible and adjustable to different lidar systems and configurations (e.g., temporal coverage from 2008 until 2021, different retrieval modes based on the available lidar capabilities). In addition, all retrievals (the solutions of which were statistically significant within the 95% confidence interval) succeeded in the correct identification of the dominant aerosol component.

Table 4.5: Overview of the five case studies used to evaluate the aerosol typing scheme, along with the aerosol type (as in the experimental basis and literature), the OEM output (i.e., relative volume contribution) and the retrieval mode (see Tab. 4.4).

Date	Aerosol type	OEM output (%)				Retrieval mode
		FSA	CS	FSNA	CNS	
14/09/2008	Smoke	**50** ± 20	21 ± 19	21 ± 18	8 ± 21	3
15/04/2016	Marine	1 ± 10	**99** ± 22	0 ± 6	0 ± 7	5
18/04/2021	Pollution	1 ± 15	13 ± 19	**80** ± 22	0 ± 17	2
05/02/2008	Dust	5 ± 9	2 ± 18	4 ± 11	**88** ± 22	1
29/04/2016	Mixture	1 ± 6	21 ± 11	7 ± 11	**71** ± 21	5

4.4.1   Smoke

The first example (from the experimental basis) discussed here is a measurement of smoke from Fig. 3.1. The measurement was conducted in Brazil during the dry season of 2008 (Baars, 2011) on 14/09/2008. The vertical profiles of the aerosol properties (not shown here) were derived with the Raman method (between 01:00 and 02:00 UTC). A mean lidar

ratio of 78 ± 7 sr, a mean particle linear depolarization ratio of $3.2 \pm 2\%$ (both at 355 nm), and a mean extinction-related Ångström exponent of 0.7 ± 0.5 were used as input to the OEM scheme (retrieval mode 3, see Tab. 4.4). The values of the optical parameters already indicate that the particles that were present in the atmosphere were rather absorbing (high S), mostly spherical (low δ) and mostly of fine mode (relatively high $\mathring{A}$).

Based on the aforementioned values, the decision tree (Sec. 4.1.4) assigned a set of four values as the initial guess of the state vector, which allowed the iterational process to begin. In this case, the initial guess of the state vector was determined as "FSA*" (see Tab. 4.3) with contributions of 85% of the relative volume from FSA particles and the remaining 15% distributed equally between the three other aerosol components (5% per aerosol component).
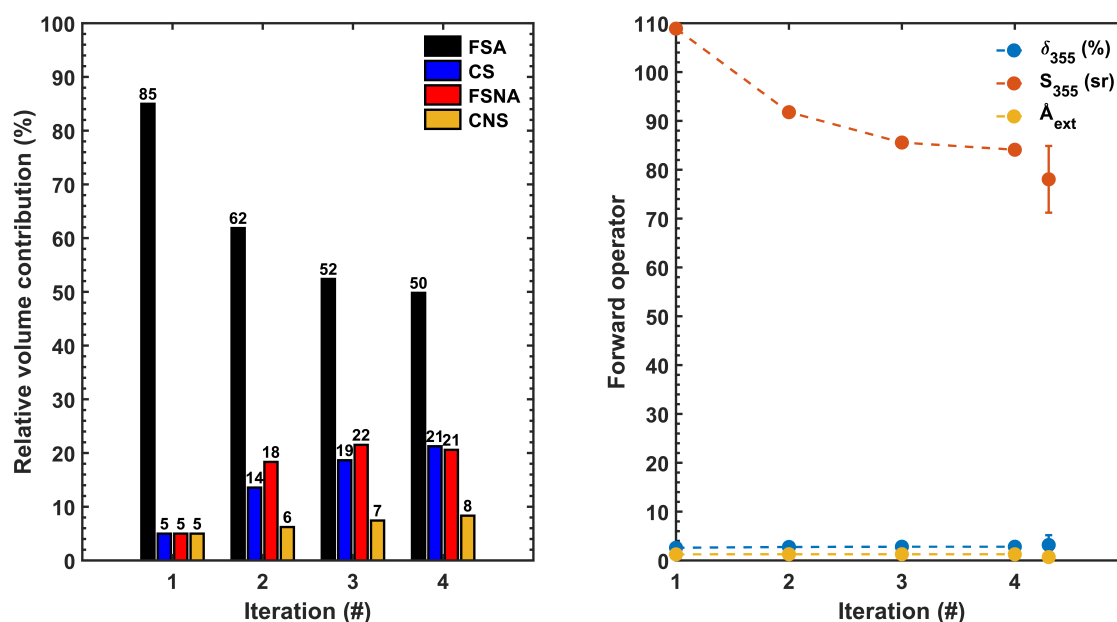
Figure 4.6: Evolution of the state vectors (left) and the optical parameters (right) as function of the iteration step for the smoke case of 14/09/2008.

The evolution of the state vector at each iteration step is visualized in the left panel of Fig. 4.6. In the right panel of this figure, the values of the different optical properties considered in this retrieval (as resulting from the evaluation of the forward model at each iterational step) are also presented. With the given initial guess (iteration step = 1), the modelled optical parameters were: $S = 109$ sr (355 nm), $\delta = 2.6\%$ (355 nm) and $\mathring{A} = 1.2$ (extinction-related). While the forward-modelled values of particle linear depolarization ratio and Ångström exponent are both within the measurement errors, the forward-modelled value of lidar ratio is overestimated. This overestimation translates into a cost function (Eq. (4.9)) of 22 (Fig. 4.7, left). As Fig. 4.7 (right panel) indicates, the costs at the first step (iteration step = 1) are solely observational costs (term $\boldsymbol{J_y}(\widehat{\boldsymbol{x}})$ in Eq. (4.7)), meaning

that they originate only from the differences between the measured and forward-modelled optical parameters. The initial-guess costs (term $J_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{x}})$ in Eq. (4.7), denoted as "a priori" in Fig. 4.6) are zero, since the evaluation is performed for the initial guess of the state vector ($\boldsymbol{x} = \boldsymbol{x}_\alpha$). Constraint costs ($J_{\mathrm{con}}(\widehat{\boldsymbol{x}})$ in Eq. (4.7)) are also zero, since there is no negative relative volume per aerosol component. In addition to the high costs, the convergence criterion (Eq. 4.11) is not met and, therefore, in the next step the algorithm needs to adjust the initial guess of the state vector to decrease these observational costs (and thus get "closer" to the measured optical parameters) and minimize this cost function.
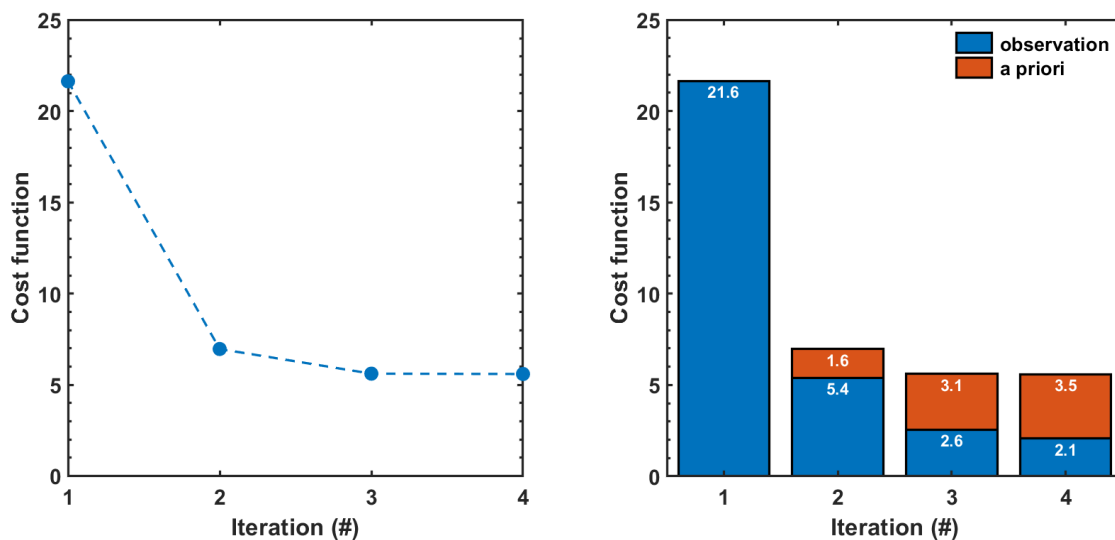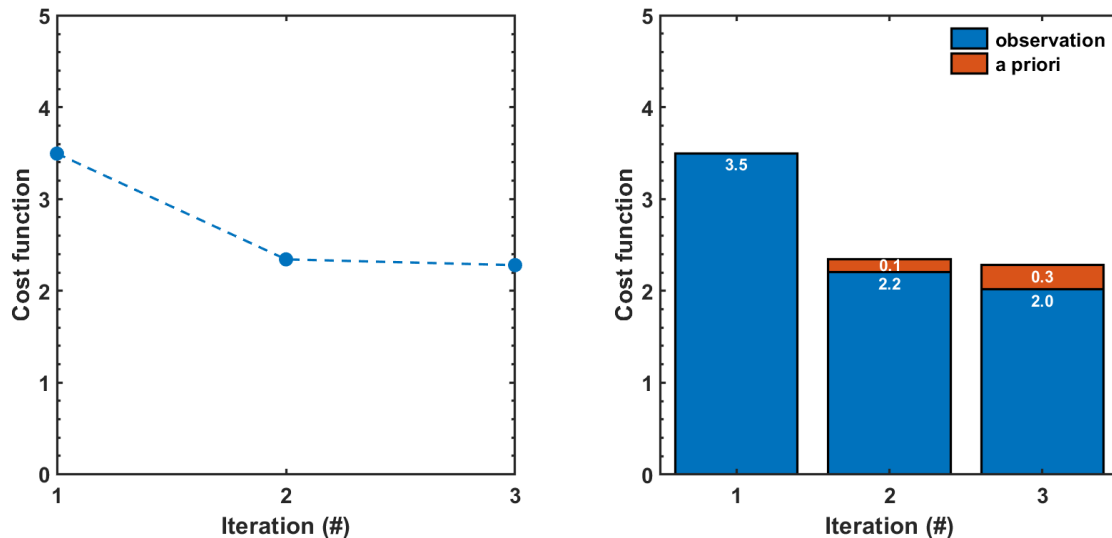

Figure 4.7: Evolution of the cost function (left), and breakdown of costs (right) as function of the iteration step for the smoke case of 14/09/2008.

Indeed, in the next iteration step (iteration step = 2) a new state vector is calculated (based on Eq. (4.10), Fig. 4.7, left), with lower contribution of FSA particles (62%) and higher contributions of the remaining three components (14%, 18% and 6% for CS, FSNA and CNS, respectively). These adjustments in the state vector resulted in a decrease of the lidar ratio (Fig. 4.7, right), however still not within the measurement error. The particle linear depolarization ratio and the Ångström exponent remained almost constant. The cost function (Fig. 4.7, left) decreased to 7, with a significant decrease of the observational costs (since the deviation from the measurements decreased) and an increase of the initial-guess costs, reflecting the changes in the state vector. A similar behavior of the state vector, forward operator and cost function is observed for the third iteration.

Finally, at iteration step four, the convergence criterion of Eq. (4.11) was fulfilled ($0.26 \ll 3$) and convergence was met. The state vector, which is the optimal solution (Fig. 4.6, left), is dominated by FSA particles ($50 \pm 20\%$ of the total volume). We also observe equal contri-

butions of CS and FSNA particles ($21 \pm 19\%$ and $21 \pm 18\%$ of the total volume, respectively) and only a small contribution of CNS particles ($8 \pm 21\%$). It should be noted that the error range should be treated in a physically meaningful way (e.g., $8 \pm 21\%$ CNS particles contribution means that the relative volume contribution of CNS particles can range between 29% and 0%). The aforementioned mixing ratio (relative volume contributions per component) produced a lidar ratio of 84 sr at 355 nm, a particle linear depolarization ratio of 2.8% at 355 nm and an extinction-related Ångström exponent of 1.2. At this step (iterational step = 4), the cost function reached a minimum (Fig. 4.7, left), with a value of 5.6 (difference of 0.02 to previous step). The respective observational costs decreased further (as the modelled values are within the measurement error range) and the initial-guess costs increased (as the optimal solution is further away from the initial guess of the state vector). The optimal solution must undergo a Pearson's chi-squared test to ensure that the convergence is correct (e.g., convergence at a global minimum instead of a local minimum). The chi-squared value was calculated as in Eq. (4.13), and the χ_{thr} was retrieved from the chi-square distribution table for 3 degrees of freedom and a p-value of 0.05. Since $\chi^2 < \chi_{\text{thr}}$ (5.5 < 7.8), the solution was found to be statistically significant within the 95% confidence interval.

The OEM-based typing scheme was able to identify the predominance of the FSA component ($50 \pm 20\%$ of the total volume) from the optical parameters of the observed aerosol mixture. This component is associated with highly absorbing soot-containing aerosol, although the exact amount of soot cannot be determined by this typing scheme. Relatively high contributions of the FSNA component ($21 \pm 18\%$) indicate the presence of less absorbing particles (with respect to the FSA component) that could be possibly attributed to aged smoke particles (mixture of sulphates, nitrates and soot). Fine-mode particles were clearly dominant in the aerosol mixture with a total relative volume of 71%. Coarse-mode particles were also present in the mixture, with a total relative volume of 29%. The CS component could be attributed to fine-mode particles that have undergone hygroscopic growth (Zhang et al., 2008), while the CNS component could be associated with a mixture of irregularly shaped soil dust particles with partly coated soot particles (Nisantzi et al., 2014).

4.4.2   Marine aerosol

Measurements in the marine boundary layer (MBL) from the Polarstern cruise PS98 on 15/04/2016, at around 04:40 UTC (Bohlmann et al., 2018), were extracted from the experimental basis and used as an input to the OEM typing scheme. Between 0.2 and 0.3 km, mean lidar ratios of 26.8 ± 9 sr and 19.1 ± 2 sr at 355 and 532 nm, respectively, and particle linear depolarization ratios of $1.5 \pm 0.2\%$ and $1.6 \pm 0.5\%$ at 355 and 532 nm, respectively, were observed. Since the optical parameters are available at both wavelengths, retrieval mode 5 was selected (Tab. 4.4).
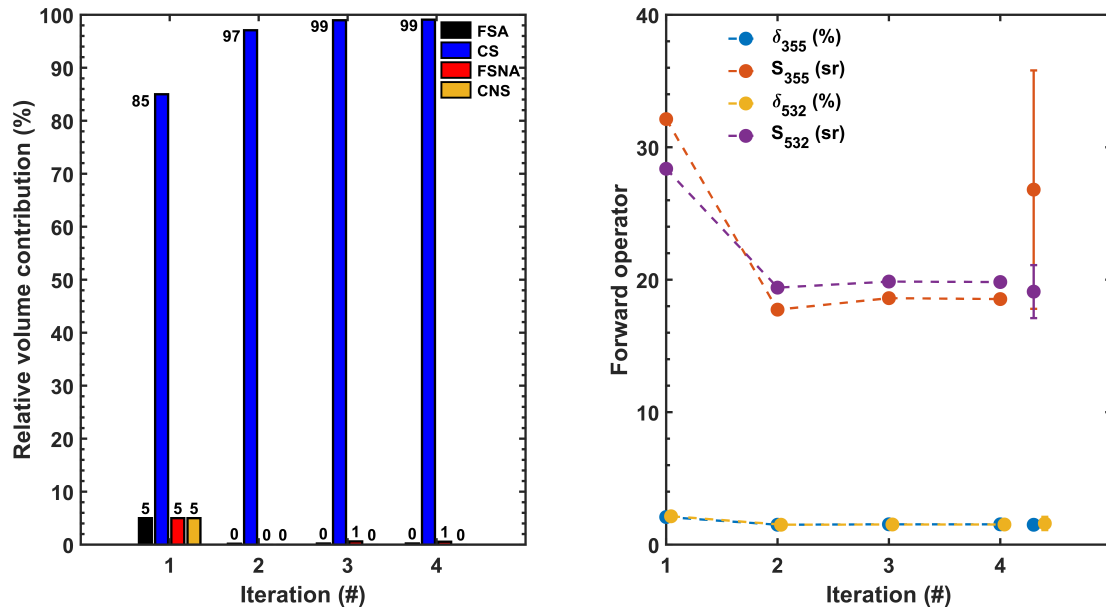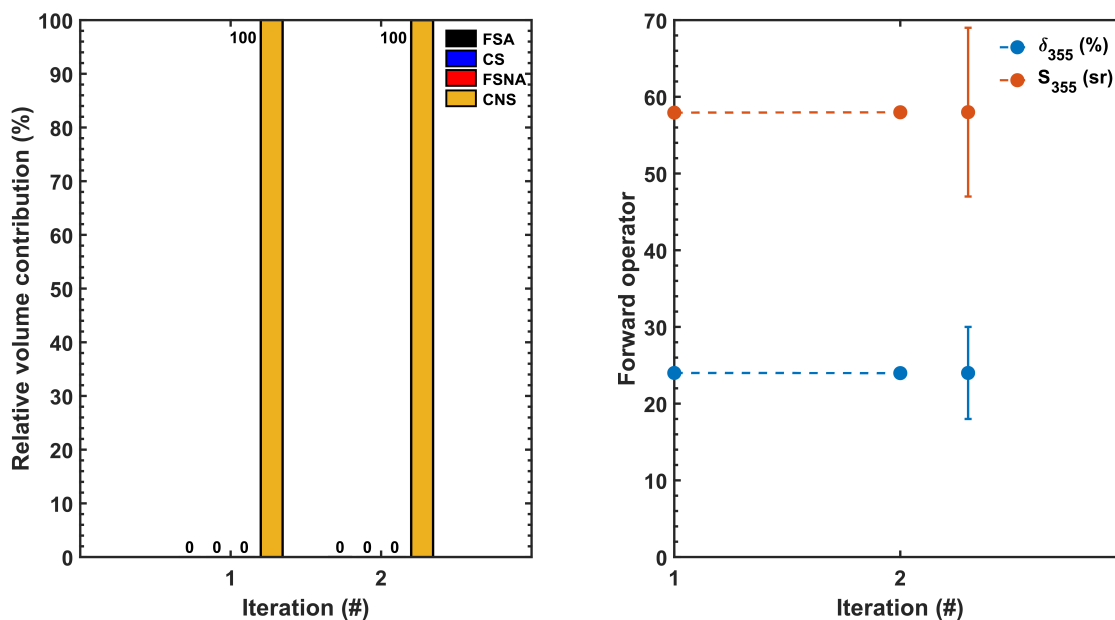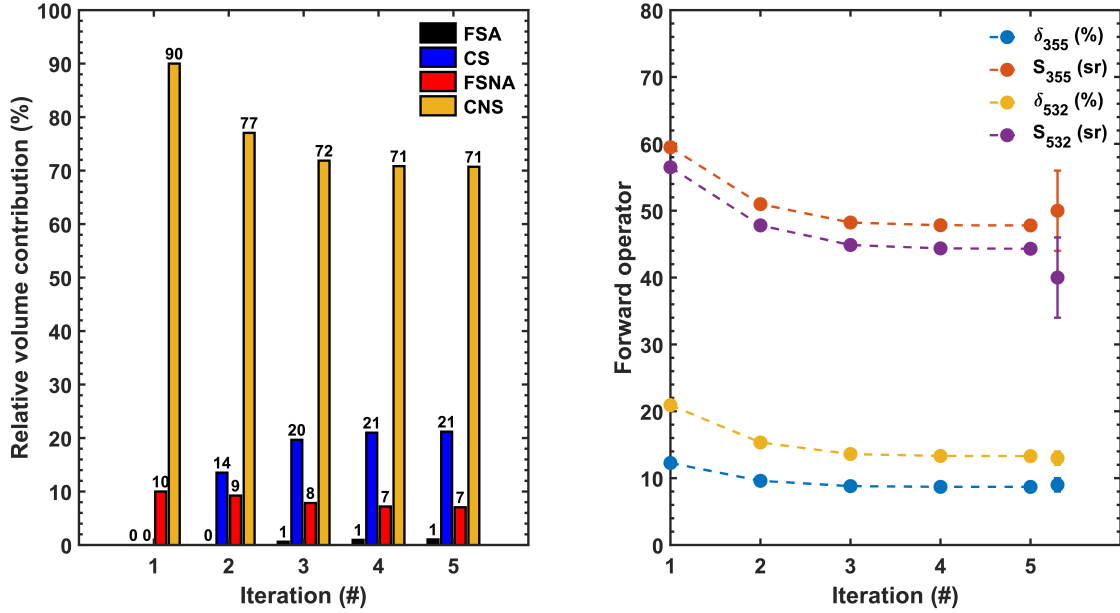
Figure 4.8: Evolution of the state vectors (left) and the optical parameters (right) as function of the iteration step for the marine case of 15/04/2016.

The low lidar ratios as well as the low particle linear depolarization ratios are characteristic for aerosol of marine sources and, therefore, already from the initial guess of the state vector (Fig. 4.8, left) the CS component, which is associated with marine aerosol, dominated. Convergence to a statistically significant (within the 95% confidence interval) solution was reached after four iterations and the optimal solution revealed that the CS component dominated the aerosol load with a relative volume contribution of $99 \pm 22\%$. The contributions of the FSNA, FSA and CNS components are negligible ($1 \pm 10\%$, $0 \pm 6\%$ and $0 \pm 7\%$, respectively). For the optimal solution, all the forward-modelled optical parameters at both wavelengths were within the measurement error (Fig. 4.8, right).

Even though the initial guess of the state vector was dominated by the CS component, the rather small contributions of the other aerosol components (5%) resulted in lidar ratios much higher than the ones observed (Fig. 4.8, right), which led to high observational costs and, thus, a cost function of 31.6 (Fig. 4.9). It is worth mentioning that in the second iteration step, while the observational costs decreased and the initial-guess costs increased, the minimum of the cost function (1.51) is observed. This feature occurred due to the normalization of the state vector in the second iteration step (all components except CS were 0, Fig. 4.8, left). However, the convergence criterion (Eq. (4.11)) was not met (due to the high differences between the forward-modelled optical parameters between iteration steps one and two) and, thus, the iterational procedure continued. At the optimal solution (iteration step = 4) the total costs were 1.56 (difference from the previous step of the order
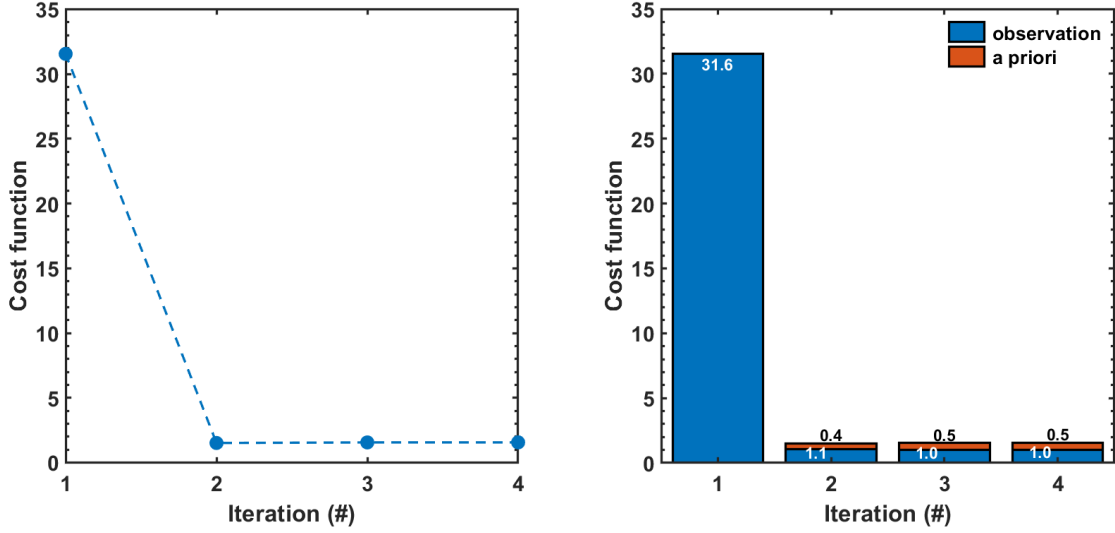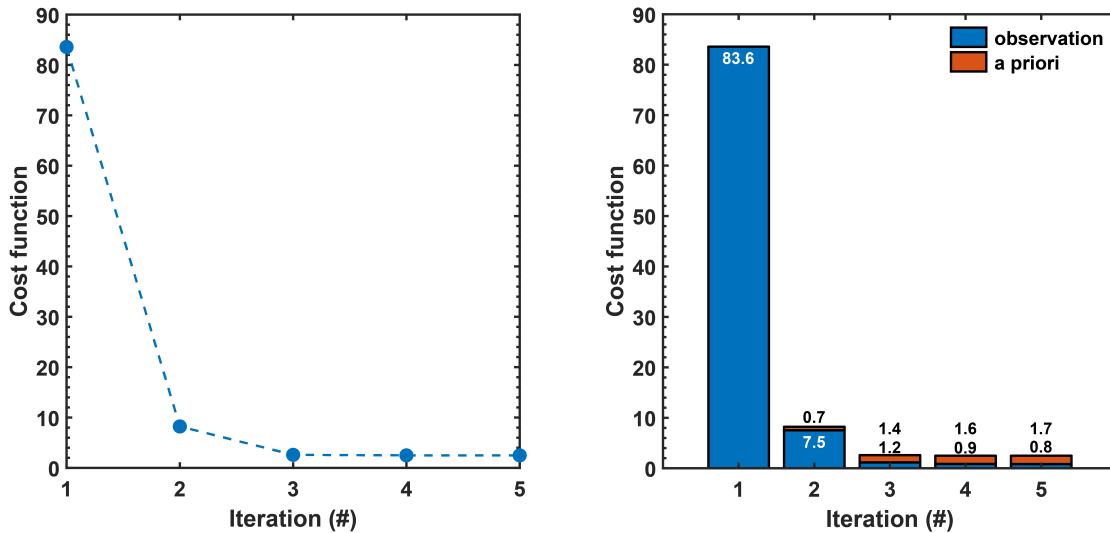
Figure 4.9: Evolution of the cost function (left), and breakdown of costs (right) as function of the iteration step for the marine case of 15/04/2016. Values above each bar have been rounded up to the first significant digit.

of 10^{-2}).

In this case (retrieval mode 5), the algorithm was able to retrieve successfully the mixing ratio of the different aerosol components from the MBL measurements. Since multiwavelength measurements were available, for the sake of demonstration, this case was examined using all the other retrieval modes as well. Retrieval mode 1 (355 nm only) resulted in a statistically significant solution (within the 95% confidence interval) with a clear predominance of CS particles in terms of relative volume ($91 \pm 22\%$). Contributions of FSA, FSNA and CNS particles are almost negligible ($7 \pm 10\%$, $2 \pm 16\%$ and $0 \pm 12\%$, respectively). Retrieval mode 2 (532 nm only) also resulted in a statistically significant solution with a clear dominance of the CS component ($97 \pm 22\%$). Understandably, the contribution of the remaining components was imperceptible ($1 \pm 11\%$, $2 \pm 18\%$ and $0 \pm 5\%$ for FSA, FSNA and CNS, respectively). Retrieval mode 3 (355 nm and extinction-related Ångström exponent) resulted in a solution, which was not statistically significant. Nevertheless, the dominant component was again the CS one, with a relative volume contribution of $97 \pm 22\%$. The FSA, FSNA and CNS components exhibit almost no contributions at all ($2 \pm 10\%$, $0 \pm 10\%$ and $1 \pm 7\%$, respectively). The reason why this particular retrieval was not statistically significant is that even though the forward-modelled lidar ratio at 355 nm was within the measurement error, it was still "too far" (in absolute terms) from the measured lidar ratio. The same reason also led to large values in the covariance matrix between the measurement ($\boldsymbol{y}$) and $\boldsymbol{F}(\widehat{\boldsymbol{x}})$ ($\boldsymbol{S}_{\delta\widehat{y}}$) and thus to large χ^2 values (see Eq. (4.13)). Retrieval modes 4 and 6 were not available for the evaluation of this case due to the lack of data for the backscatter

coefficient at 1064 nm.

In summary, all retrievals (regardless the statistical significance status) captured successfully the predominance of the CS component (relative volume contribution $> 90\%$) in the aerosol mixture observed. However, it is difficult to argue about which retrieval mode produced the best results, since for every retrieval performed with the OEM-based typing scheme the cost function evaluated (Eq. (4.9)) is unique.

4.4.3 Pollution

Measurements conducted in Leipzig, which is a typical urban Central European site, were used as an example case to demonstrate the ability of the algorithm to distinguish the pollution-related aerosol component. The measurements were carried out on 18/04/2021 and vertical aerosol profiles were derived with the Raman method between 04:30 and 05:20 UTC. Averaging for altitudes between 1 and 1.5 km yields to a mean lidar ratio of 55 ± 5 sr and a mean particle linear depolarization ratio of $2 \pm 1\%$, both at 532 nm.
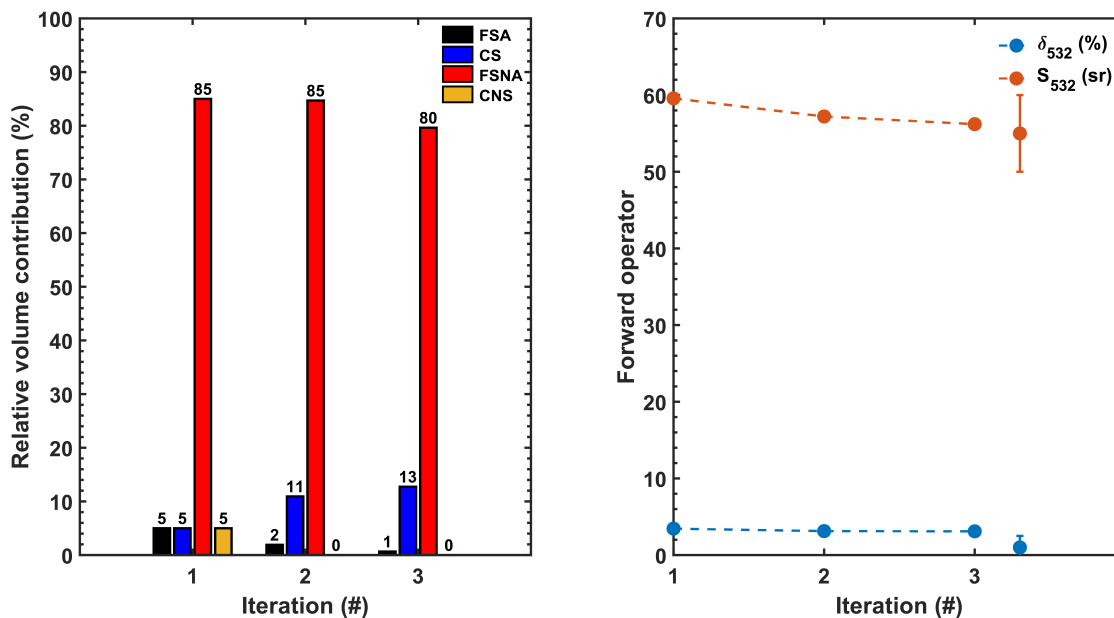
Figure 4.10: Evolution of the state vectors (left) and the optical parameters (right) as function of the iteration step for the pollution case of 18/04/2021.

Both measurements and errors were used as the algorithm's input and retrieval mode 2 (Tab. 4.4) was performed. The iterational process converged successfully at a statistically significant solution after three iterations. The initial guess of the state vector as well as the evolution of the state vector until convergence are presented in Fig. 4.10 (left). It can be seen that already the initial guess of the state vector is dominated by the FSNA

Figure 4.11: Evolution of the cost function (left), and breakdown of costs (right) as function of the iteration step for the pollution case of 18/04/2021. Values above each bar have been rounded up to the first significant digit.

component. The optimal state (iteration step = 3) has small deviations from the initial guess of the state vector. The FSNA component exhibits a clear predominance, with a relative volume contribution of $80 \pm 22\%$ (decrease of 5% from the initial guess), while the second most present aerosol component is the CS one, with a relative volume contribution of $13 \pm 19\%$ (increase of 8% from the initial guess). Since the retrieval error is higher than the contribution of the CS component itself, we cannot draw solid conclusions on the contribution of CS particles, only that values can range from 0 up to 33%. The FSA and CNS components have negligible contributions ($1 \pm 15\%$ and $0 \pm 17\%$, respectively). The remaining undistributed 6% of the relative volume can be characterized as "unidentified aerosol", and it is a result of the measures in action prohibiting physically meaningless volume contributions (i.e., constraints, see Sec. 4.1.6). For the optimal solution state, the forward operator was able to simulate the measurements within the measurement error (Fig. 4.10, right). Even though at the initial guess of the state vector (iteration step = 1) the costs are as low as 3.5 (Fig. 4.11, left), the convergence criterion is not met (criterion of Eq. (4.11)) due to the difference between the modelled and measured optical properties and their associated covariance. At convergence the costs are minimum (2.3) and the reduction of the cost function from the previous step is very small (of the order of 10^{-2}).

The clear predominance of the FSNA component ($80 \pm 22\%$) in the aerosol mixture above Leipzig on 18/04/2021 was well captured by the algorithm. Given the small relative volume contributions of the other components, along with their retrieval-associated errors, we can

conclude that most probably the aerosol observed consisted of solely FSNA particles.

4.4.4   Dust

The dust measurement examined in this section was conducted at the Praia airport, Cape Verde during the SAMUM–2 field experiment on 05/02/2008 (Groß et al., 2011). A lidar ratio of 58 ± 11 sr and a mean particle linear depolarization ratio of $24 \pm 6\%$ at 355 nm already indicate a strong presence of moderately absorbing, non-spherical (high values of δ) particles. These measurements were used as the algorithm's input and, therefore, retrieval mode 1 (Tab. 4.4) was applied.

Figure 4.12: Evolution of the state vectors (left) and the optical parameters (right) as function of the iteration step for the dust case (SAMUM–2).

Already from the initial guess of the state vector, a clear predominance of the CNS component is evident (iteration step $= 1$ in Fig. 4.12, left). The forward-modelled optical parameters are already within the measurement error (Fig. 4.12, right) and, thus, the observational costs are very small (Fig. 4.13). Convergence is therefore met after 2 iterations and the optimal solution indicates the presence of $99.9 \pm 22\%$ CNS particles, $0.1 \pm 8\%$ FSA particles, $0 \pm 11\%$ FSNA particles, and $0 \pm 18\%$ CS particles, all in terms of relative volume contribution. The solution is statistically significant within the 95% confidence interval. However, given the low relative volumes and high retrieval uncertainties of all the components except the CNS one, from a physical point of view, we can only conclude with confidence that the dominant component in the mixture observed is the CNS aerosol
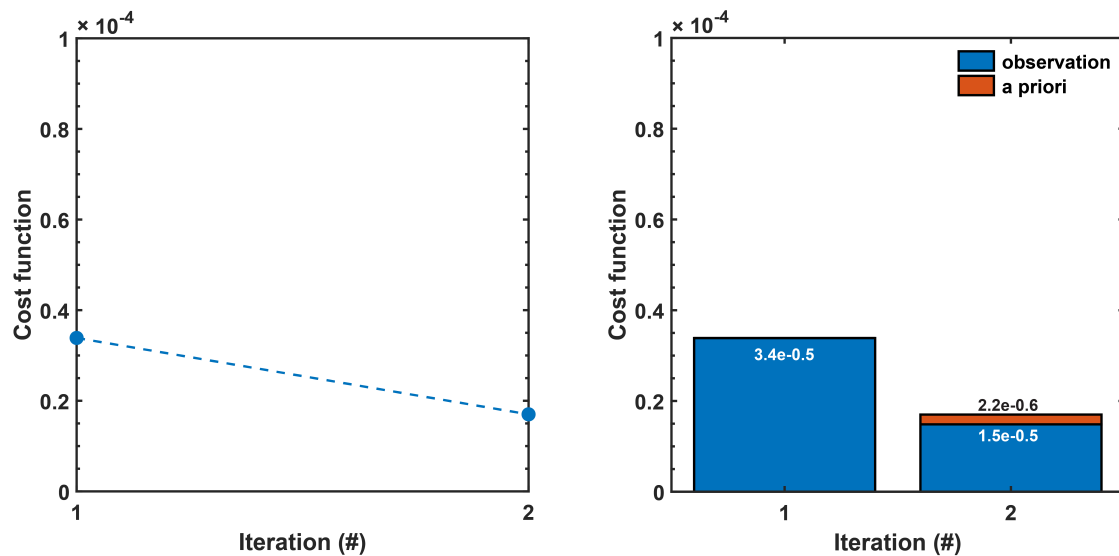
Figure 4.13: Evolution of the cost function (left) and breakdown of costs (right) as function of the iteration step for the dust case (SAMUM–2).

component.

4.4.5 Complex aerosol mixture of smoke, dust and marine particles

During the Polarstern cruise PS98 (Bohlmann et al., 2018), apart from pure marine conditions (such as in the case discussed in Sec. 4.4.2), complex aerosol mixtures containing smoke and dust mixed with marine particles were observed and characterized. On 29/04/2016, from 20:15–21:00 UTC, several aerosol layers were observed between 0.9 and 3 km. Here, the focus is put on only one layer, for a complete overview readers may refer to Bohlmann et al. (2018). The layer of interest was observed between 1.7 and 2.2 km and it has been characterized as a complex mixture of dust, smoke and marine particles. Mean lidar ratios of 50 ± 6 sr and 40 ± 6 sr and particle linear depolarization ratios of $9\pm1\%$ and $13\pm1\%$ at 355 and 532 nm, respectively, were observed. The low particle linear depolarization ratio values in combination with the moderate lidar ratios and a mean extinction-related Ångström exponent of 1 ± 0.4 indicate that the mixture contains small, non-depolarizing and absorbing particles.

A retrieval based on retrieval mode 5 was chosen (Tab. 4.4). The iterational process converged after 5 iterations to a statistically significant solution (within the 95% confidence interval). Based on the optical parameters used as an input to the OEM typing scheme, the initial guess of the state vector was determined as mixture of the CNS and FSNA components (iteration step = 1 in Fig. 4.14, left). At this point, it should be noted that the initial guess has been modified with respect to the one listed in Tab. 4.3, because using

Figure 4.14: Evolution of the state vectors (left) and the optical parameters (right) as function of the iteration step for the complex aerosol mixture of 29/04/2016.

that initial guess led to a statistically non-significant solution. The state vector at the first iteration step produces much higher lidar ratios and particle linear depolarization ratios than those observed (Fig. 4.14, right) and consequently high costs that are solely observational (iteration step = 1 in Fig. 4.15). To meet the observations, in the following iteration steps the state vector is increasing the contributions of the CS component, while decreasing the contributions of both CNS and FSNA components. At convergence (iteration step = 5), the optimal state vector shows contributions of $71 \pm 21\%$ CNS, $21 \pm 11\%$ CS, $7 \pm 11\%$ FSNA and $1 \pm 6\%$ FSA particles, all in terms of relative volume. For the optimal state vector all the forward-modelled optical properties meet the observations within the measurement error. The total costs are 2.5 (difference of the order of 10^{-4} from previous step) and mainly attributable to the initial-guess costs (Fig. 4.15).

   In this case, the OEM typing scheme was proven to be able to separate the different aerosol components present in the mixture observed. The component with the highest contribution in terms of relative volume is the CNS, usually associated with desert dust particles. The second highest component in abundance is the CS one, which is associated with particles of marine origin. Fine-mode particles contribute to the mixture with a total relative volume contribution of about 9%. Given the corresponding contributions of FSA and FSNA particles, the small, absorbing and non-depolarizing particles speculated to be present in the aerosol mixture already from the optical properties of the layer are more likely FSNA particles (typically pollution-associated) than FSA particles (usually associated with

Figure 4.15: Evolution of the cost function (left) and breakdown of costs (right) as function of the iteration step for the complex aerosol mixture of 29/04/2016.

smoke). However, FSA particles might also contribute to the mixture with relative volume contributions up to 7%.

For the same case, retrieval modes 1, 2 and 3 were additionally applied (results not shown here). All the retrieved optimal state vectors were not statistically significant within the 95% confidence interval. Nevertheless, they were all consisting of the same aerosol components as in the optimal state vector of retrieval mode 5, but with different mixing ratios. Due to these differences, the optimal solutions were in principle not able to reproduce the measurements within the experimental error, thus causing large differences in the corresponding terms of Eq. (4.13) (and, therefore, failing to be statistically significant). It is clear that in this case information at only one wavelength (e.g., as in retrieval mode 1 and 2) is not sufficient to resolve this complex aerosol mixture. The information provided by the extinction-related Ångström exponent (included in retrieval mode 3) did not improve the situation, as apparently there are many state vectors that can produce optical properties similar to those measured by the lidar. Therefore, it can be concluded that multispectral information on the lidar ratio and particle linear depolarization ratio is essential for the adequate characterisation of such complex aerosol mixtures.

Chapter 5

Application - Part I: Case studies

In this chapter, the OEM-based aerosol typing scheme is applied to three selected case studies to provide a more insightful overview of the algorithm capabilities. The first case examined and presented below (Sec. 5.1) concerns a geometrically thick dust layer from the A-LIFE (Absorbing aerosol layers in a changing climate: aging, LIFEtime and dynamics) field campaign, which took place at Cyprus in April 2017. This case study was selected because in addition to the lidar observations, airborne in-situ measurements from the DLR (Deutsches Zentrum für Luft– und Raumfahrt – German Aerospace Center) Falcon 20E research aircraft were also available. For the first time, a comparison between lidar-estimated and airborne-derived number and volume size distributions, as well as the effective radius has been attempted here.

The second case regards two aerosol layers that were observed in January 2008 over Praia, Cape Verde during the SAMUM–2 field campaign (Sec. 5.2). The aerosol layers were characterized as a mixture of smoke and desert dust particles (Tesche et al., 2011a,b) and the corresponding aerosol contributions to the backscatter and extinction coefficients have been determined with the POLIPHON method. Therefore, this case study facilitates a comparison between the POLIPHON and the OEM results to examine the consistency between the two methodologies.

In the third case analysed and discussed here (Sec. 5.3), the focus is given to a pronounced smoke plume observed above Leipzig, Germany in September 2020. The smoke plume originated from southwestern North America and was associated with the Californian wildfires that occurred during the same period (Baars et al., 2021). This case is ideal for the identification of any potential challenges with respect to aerosol typing. Due to the long-range transport, the smoke particles observed were aged and, therefore, exhibit different optical properties than those considered by the typing scheme.

5.1   Case 1: Desert dust from A-LIFE field campaign

5.1.1   Overview

The A-LIFE project, led by the University of Vienna, aimed to investigate the properties of absorbing aerosol and in particular those of mineral dust, black carbon and their mixtures. For that purpose, a field experiment took place in Cyprus, an island located in the Eastern Mediterranean region, in April 2017. The collection of the required aerosol data was performed by high-sophisticated in-situ and remote-sensing instruments integrated on the DLR research aircraft Falcon 20E as well as with ground-based in-situ and remote-sensing instruments from various institutes and universities. TROPOS contributed to this project with a ground-based Raman lidar of type Polly$^{\text{XT}}$ (Engelmann et al., 2016), as part of LACROS (Leipzig Aerosol and Cloud Remote Observations System), which was located in Limassol (34.71° N, 33.02° E) and operated on the premises of the Cyprus University of Technology (CyCARE campaign, Ansmann et al. (2019); Radenz et al. (2021a)).

A-LIFE was a great opportunity for further support of EarthCARE's preparation and, thus, the joint project A-CARE emerged (A- from A-LIFE and CARE from EarthCARE). A-CARE aims at further development and refinement of EarthCARE's geophysical product algorithms. The consistency, in terms of microphysical assumptions, between the OEM-based aerosol typing scheme and EarthCARE's aerosol classification algorithm (HETEAC), allows the usage of the OEM-based scheme for testing and validation purposes. To be more specific, lidar-derived optical parameters were used as input to the OEM typing scheme and then the output products were compared directly with the airborne in-situ data. By doing so, first estimates for the capabilities of EarthCARE's classification schemes can be produced. In addition, a full assessment of the microphysical assumptions used in the classification scheme can be performed, allowing thus fine re-adjustments (if needed), for calibration/validation purposes. Below, a short description and information regarding the data used is provided.

Airborne in-situ data

The DLR research aircraft Falcon 20E was equipped with in-situ and remote-sensing aerosol instruments (in the cabin, on the wings, and on the nose of the aircraft). Some key parameters measured were the total number concentration of fine-mode particles measured by three Condensation Particle Counters (CPC) with different lower cut-off diameters D_{p} ($0.005 < D_{\text{p}} < 2.5$ µm), the aerosol particle size distributions (0.5 µm $< D_{\text{p}} < 50$ µm), particle optical properties and particle shapes measured by a Cloud and Aerosol Spectrometer with Detector for Polarization Detection (CAS-DPOL), the total and non-volatile size distribution ($0.25 < D_{\text{p}} < 2.5$ µm) measured by two Optical Particle Counters (OPC, type Grimm

Sky 1.129), as well as the chemical composition and shape of particles (D_{p} < 2.5 µm) collected by filters by an impactor sampler. It should be noted that all the airborne in-situ data presented and used in this dissertation have been analysed and provided to us by the University of Vienna. The full list of the instrumentation and parameters measured during A-LIFE can be found in the following web page: `https://www.a-life.at/instrumentation/` (last access: 21/03/2022).
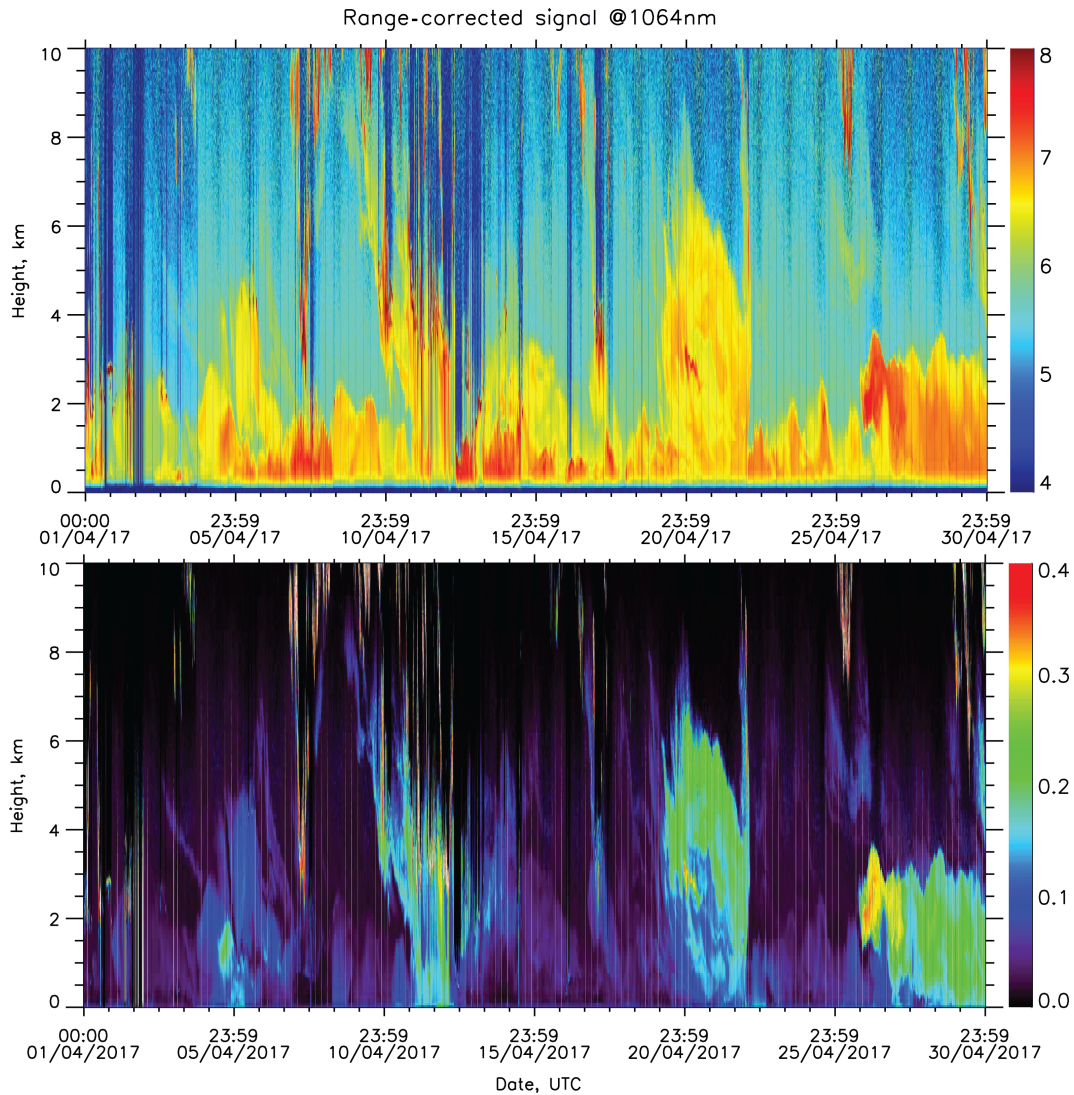
Figure 5.1: Overview of the atmospheric conditions, in terms of range-corrected signal at 1064 nm (top panel) and particle linear depolarization ratio at 532 nm (bottom panel) at Limassol, Cyprus during the whole A-LIFE measurement period (April 2017).

Table 5.1: A-CARE selected case studies, along with the layer-mean optical properties and their uncertainties (as derived by Polly^{XT} lidar). The first value in each optical parameter field corresponds to 355 nm, and the second one to 532 nm. The air-mass sources have been identified by means of backwards trajectories (performed by University of Vienna).

Case #	Date	Layer altitude (km)	Source	δ (%)	S (sr)
1	06/04/2017	2.0–4.0	Sahara	15.6 ± 0.6 19.3 ± 0.6	38 ± 6 33 ± 6
2	06/04/2017	0.8–1.2	Arabia	17.2 ± 0.6 19.8 ± 0.9	42 ± 7 38 ± 6
3	11/04/2017	1.0–1.5	Cyprus	7.3 ± 0.8 2.9 ± 0.9	75 ± 14 74 ± 28
4	14/04/2017	0.9–1.3	Cyprus	5.1 ± 0.3 2.6 ± 0.3	36 ± 6 33 ± 8
5	20/04/2017	3.0–5.0	Sahara	20.6 ± 2.0 26.4 ± 0.8	49 ± 8 44 ± 7
6	21/04/2017	4.0–6.0	Sahara	21.3 ± 2.0 27.7 ± 1.0	53 ± 8 46 ± 7
7	25/04/2017	1.0–2.0	Turkey	7.4 ± 1.0 6.7 ± 1.1	53 ± 8 46 ± 9
8	27/04/2017	1.5–3.8	Arabia	28.5 ± 1.5 32.3 ± 0.9	38 ± 6 34 ± 5

Ground-based lidar data

An overview of the atmospheric conditions during the A-LIFE campaign is given in Fig. 5.1. During a one-month measurement period, several aerosol types were observed as indicated by the range-corrected signal at 1064 nm (Fig. 5.1, top panel) and the wide range of particle linear depolarization ratio at 1064 nm (Fig. 5.1, bottom panel). A selection of case studies was considered and analysed in the framework of A-CARE (presented in Tab. 5.1). The case studies were carefully selected (after coordination with the responsible scientists for the airborne in-situ measurements) to ensure coverage of various aerosol scenarios and, thus, facilitate A-CARE's activities and project goals. The diverse optical properties presented in Tab. 5.1 reflect the different aerosol conditions that were under investigation, as well as the different source regions. The key lidar parameters for the OEM-based typing scheme (namely the lidar ratio and the particle linear depolarization ratio) were derived from measurements temporally co-located with the Falcon 20E flights. Since the flights were conducted mostly during daytime, the return signals from the Raman scattering could not always be used for the retrievals due to the too strong background light. In these cases, the lidar ratio was derived from the closest night-time observation in the same aerosol layer

and then used to analyse further the daytime lidar data. While all cases of Tab. 5.1 are of great importance and interest, only case 5 will be discussed in detail here. The results of the OEM-based aerosol typing scheme for all the A-CARE cases as well as some relevant additional products are available in Appendix B.

5.1.2 Lidar observations on 20 April 2017, Cyprus
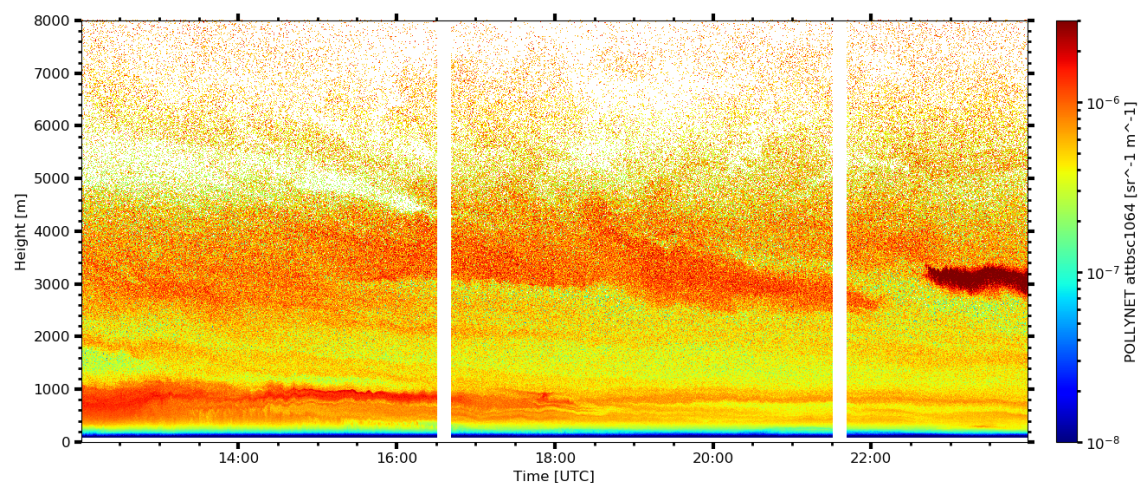
Figure 5.2: Overview of the atmospheric conditions, in terms of range-corrected signal at 1064 nm at Limassol, Cyprus on 20/04/2017, between 12:00 and 00:00 UTC. No data are available during regular depolarization calibration periods (white bars).

Figure 5.2 shows the temporal development of the range-corrected signal (at 1064 nm) on 20 April 2017, between 12:00 and 00:00 UTC. During the whole measurement period, a temporally stable, approximately 8 km geometrically thick aerosol layer was present with increased backscattering evident in parts of the aerosol layer.

On the same day, between 17:40 and 19:30 UTC, DLR Falcon 20E conducted measurements at altitudes between 0.5 and 9 km. Therefore, the vertical profiles of optical parameters from lidar between 17:00 and 19:00 UTC were averaged from 3 to 6 km altitude, aiming to capture the center of the aerosol layer. The temporal co-location of the lidar and Falcon's in-situ measurements facilitates meaningful comparisons (see Sec. 5.1.4).

The aerosol layer observed above the lidar site originated from the Sahara region as indicated by the HYSPLIT (Stein et al., 2015) backward trajectories (Fig. 5.3). Air masses arriving at 4 km originated directly from Algeria and crossed Tunisia and the Mediterranean Sea on the way to Limassol. These air masses carried mainly desert dust particles. Air masses arriving at higher altitudes (6 km) originated from France. However, four days prior to their observation at Limassol, the same air masses were located above Algeria at very low altitudes, thus, could pick up desert dust particles (aerodynamic lifting). The

source appointment performed based on Falcon 20E data also confirms high concentrations of desert dust particles.
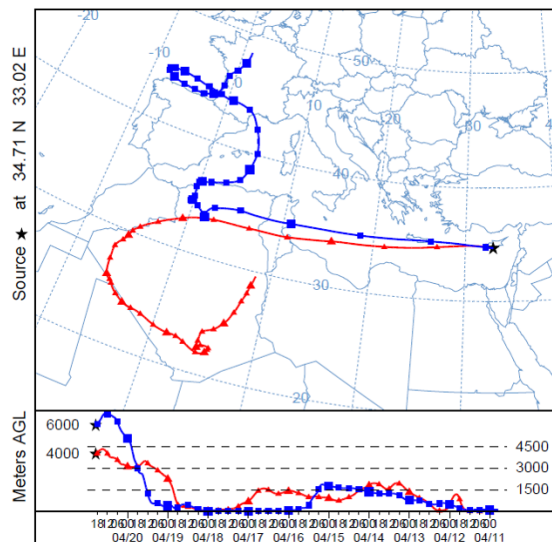
Figure 5.3: 10–day HYSPLIT backward trajectories ending at Limassol, Cyprus, on 20 April 2017, 19:00 UTC.

The lidar-derived optical parameters (between 17:00 and 19:00 UTC) are presented in Fig. 5.4. Maximum extinction coefficient values of approx. 120 Mm^{-1} occurred between 3.5 and 4 km. Extinction and backscatter-related Ångström exponents of around 0.1 and 0, respectively, were observed for this aerosol layer between 3 and 5 km. The lidar ratio was stable between the same altitudes with values of approx. 49 and 44 sr for 355 and 532 nm, respectively. The particle linear depolarization ratio was high, exceeding 20% above 3 km, indicating thus a strong presence of coarse-mode dust particles.

5.1.3   Aerosol characterization by the OEM-based typing scheme

Layer-averaged values (from 3–5 km) of lidar ratio (49 ± 8 sr) and particle linear depolarization ratio ($20.6 \pm 2\%$) at 355 nm were used as input to the OEM-based aerosol typing scheme, and the retrieval mode based only on UV information (mode 1, see Tab. 4.4) was carried out. The optimal solution (Fig. 5.5), which was also statistically significant within the 95% confidence interval, confirmed the dominance of the CNS aerosol component in the aerosol mixture: in terms of relative volume contribution, the CNS component occupied $86 \pm 22\%$ of the total aerosol mixture volume. The contributions of the remaining aerosol components were small and almost negligible: $10 \pm 11\%$ of FSNA, $4 \pm 18\%$ of CS, and $0 \pm 8\%$ FSA aerosol particles. Overall, coarse-mode particles dominated the aerosol mixture with a total relative volume contribution of 90%.
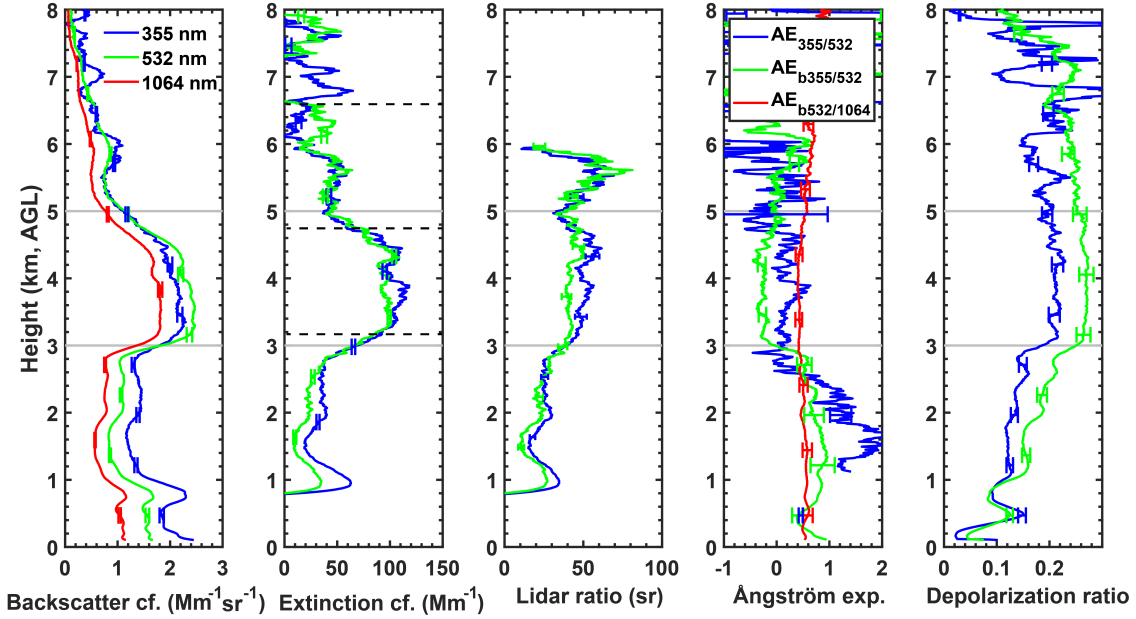
Figure 5.4: Vertical profiles of the particle backscatter and extinction coefficients, particle lidar ratio, Ångström exponents and particle depolarization ratio measured at Limassol, Cyprus, on 20 April 2017, from 17:00-19:00 UTC. A smoothing length of 750 m was used for the extinction and lidar ratio, and 200 m for the backscatter and depolarization ratio. Faint grey lines indicate the averaging layer boundaries. Dashed black lines in the extinction coefficient panel indicate the altitudes of the three Falcon 20E flight legs.

As more optical information was available than in the UV only (Tab. 5.1), the determination of the aerosol mixing ratio was also attempted based on retrieval modes 2 and 5 (see Tab. 4.4). The optimal solution for the retrieval mode 2 (information on 532 nm only) was statistically significant, while the one for retrieval mode 5 (simultaneous 355 and 532 nm retrieval mode) was not statistically significant. The statistically non-significant solution occurred due to the inability of the retrieved aerosol mixing ratio to reproduce the particle linear depolarization ratio at 532 nm within the measurement error range. Regardless of the statistical significance of the solution, both retrievals captured successfully the predominance of the CNS aerosol component ($87.3 \pm 22.3\%$ and $90.7 \pm 22.2\%$ for retrieval modes 2 and 5, respectively). While both optimal solutions of retrieval modes 1 and 2 were statistically significant, in the following sections the optimal solution of retrieval mode 1 is preferred for further product derivations since it corresponds to EarthCARE's operating wavelength.
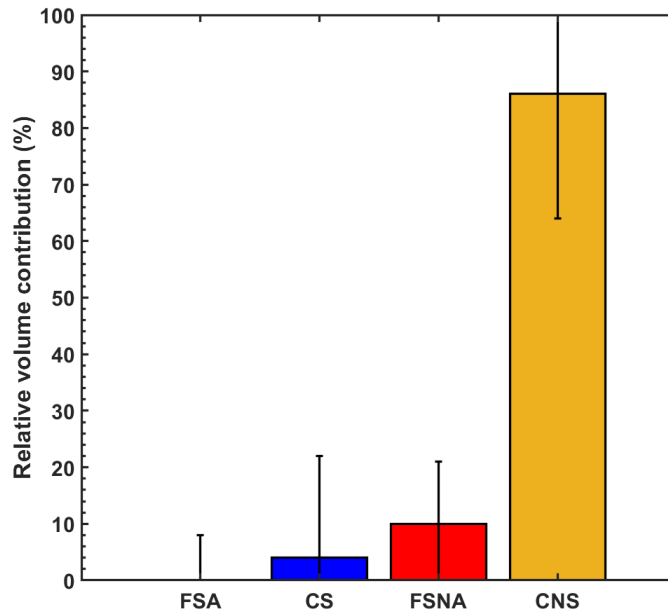
Figure 5.5: Mixing ratio of the four aerosol components as determined by the OEM typing scheme for the layer observed between 3 and 5 km at Limassol, Cyprus, on 20 April 2017, from 17:00-19:00 UTC. It should be noted that the error bars have been confined between the constraint-accepted limits.

5.1.4   Comparison with airborne in-situ data

The output of the OEM-based aerosol typing scheme (Fig. 5.5) can be transformed into aerosol size distributions, using the information of the four predefined aerosol components. In the first step, the relative volume per aerosol component is converted into the volume and number concentration per component, as described in Sec. 4.2 (Eq. (4.20) and Eq. (4.21), respectively). Then, the volume and number concentration can be converted into the respective size distributions by using each mode's width and radius (listed in Tab. 4.1) in the log-normal definition of the aerosol size distribution.

The volume and number concentration per aerosol component for the optimal solution and retrieval errors (discussed in Sec. 5.1.3) is presented in Tab. 5.2. The aerosol mixture observed is dominated by the CNS aerosol component in terms of volume concentration and by the FSNA aerosol component in terms of number concentration. The volume concentration predominance of the CNS aerosol component $(37.9 \pm 2.8 \ \mu m^3/cm^3)$ is expected since the optimal solution was also dominated by the same component. The second largest volume concentration was for the FSNA aerosol component $(4.4 \pm 2.3 \ \mu m^3/cm^3)$, followed by the CS and FSA aerosol components with very small volume concentrations $(1.7 \pm 2.4$ and $0 \pm 1 \ \mu m^3/cm^3$, respectively). The predominance of the FSNA component in the number concentration reflects the aerosol particles' size; the predefined mode radius of the fine-mode

Table 5.2: Volume and number concentration, along with the respective uncertainties, as derived from the OEM-retrieved optimal solution for the case of 20 April 2017.

	Volume Concentration ($\mu m^3/cm^3$)	Number Concentration ($\#/cm^3$)
FSA	0 ± 1	0 ± 195.6
CS	1.7 ± 2.4	0.2 ± 0.2
FSNA	4.4 ± 2.3	865.7 ± 464
CNS	37.9 ± 2.8	3.7 ± 0.3

particles is much smaller than the one of the coarse mode particles (see Tab. 4.1 for the exact values). The fact that many more FSNA particles are needed, in comparison to the CNS particles, to fill a certain volume explains the low volume and high number concentration of the FSNA aerosol component.

Aerosol size distributions were available from Falcon 20E measurements that were carried out on the same day (20 April 2017), between 17:40 and 19:30 UTC, at various altitudes ranging from 0.5 up to 9 km. The aerosol size distributions presented here are the outcome of data combined from two instruments: the in-cabin SkyOPC and the wing-mounted CAS-DPOL. The volume and number size distributions derived from the Falcon 20E measurements are presented in Fig. 5.6(a) and (b), respectively. Only flight legs that were performed at altitudes approximately within the lidar altitude range (3–6 km) are considered here and indicated with different colored lines. In the same figure, the volume and number size distributions that were derived with the lidar approach (3–5 km), using the OEM-based aerosol typing scheme (hereafter referred to as lidar/OEM approach), are presented (black line with black dashed lines indicating the uncertainties). In addition, the effective diameter calculated from the lidar/OEM approach (vertical black dashed line with grey shaded area for the uncertainty) and the measurements from Falcon 20E are presented (vertical colored dashed lines, estimates of the uncertainty were not available at the time of writing).

Before discussing the comparison results, it should be noted that a comparison between the extinction coefficient from the in-situ data and the lidar was performed. For the Falcon 20E measurements, the number size distribution was first converted to volume size distribution under the simplified assumption of spherical particles. Then the volume size distribution was used as an input to the spheroid model introduced by Dubovik et al. (2006) to obtain the scattering properties. The extinction coefficient values (at 355 nm) that were retrieved for the flight legs at 3.2, 4.8 and 6.6 km were 301.2, 205.9 and 62.3 Mm^{-1}, respectively. The extinction coefficient measured by the lidar at altitudes between 3 and 5 km was 79.2 Mm^{-1}, with maximum values around 120 Mm^{-1} between 3.5 and 4.5 km. The
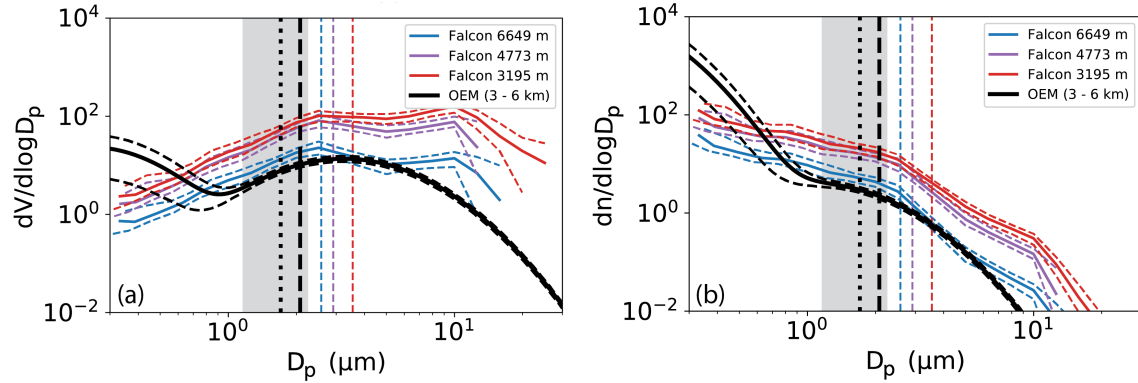
Figure 5.6: Volume (a) and number (b) size distributions derived from lidar measurements (3–6 km height) following the lidar/OEM approach (black line, dashed lines indicate uncertainties) compared to the corresponding size distributions measured on board of Falcon 20E at various altitudes (colored lines and dashed lines). The vertical lines indicate the effective diameter (black for the lidar/OEM approach and colored for Falcon 20E). The lidar/OEM-derived effective diameter has been calculated for the whole size distribution (dotted) and for a lower particle-diameter limit of 0.3 µm (dashed).

extinction coefficient obtained from the in-situ measurements appears to be vastly overestimated for the two lowest (in altitude) flight legs (3.2 and 4.8 km). For the highest (in altitude) flight leg the extinction coefficient retrieved is lower but comparable with the one measured by the lidar. The observed overestimation of the in-situ retrieved extinction coefficient could be attributed to horizontal effects. The horizontal distance between the the aircraft and the lidar site was 72 and 127 km (in the south direction) for the flight legs at 3.2 and 4.8 km, respectively. The flight leg at 6.6 km had the lowest horizontal distance from the lidar site: 57 km (in the south direction).

Figure 5.6 indicates an overall good agreement between the size distributions measured by DLR Falcon 20E and those obtained with the lidar/OEM approach, especially for the altitude for which the extinction values were consistent (around 6 km, blue lines). In particular, very good agreement was found for particles with diameters ranging between 1 and 8 µm. For particles with a diameter less than 1 µm (fine-mode aerosol particles) the agreement is less satisfactory, with the lidar/OEM approach overestimating both the volume and number size distributions. The overestimation in this size range reaches up to almost two and one orders of magnitude for the volume and number size distribution, respectively. The pronounced fine mode can be explained by two reasons. Firstly, the lidar/OEM approach is considering fine-mode particles with a fixed mode diameter of 0.14 µm (mode width = 0.53, see Tab. 4.1), which is a very rough generalization and, thus, is unable to capture the whole natural variability of the fine-mode particles. In addition, the in-situ observations used for this study have a lower size limit of 0.3 µm in particle diameter. At the time of

writing, Falcon 20E measurements for particles with a diameter smaller than 0.3 µm were still under analysis. Therefore, it would be advised to revisit this comparison when the data are available, to draw more solid conclusions with respect to the microphysical assumptions used in the OEM typing scheme and in extension to the HETEAC model. Similarly, the lidar/OEM approach produces lower values for the volume and number size distribution for particles with a diameter larger than 8 µm. The aforementioned effect is much more prominent for the volume size distribution (Fig. 5.6(a)), compared to the number size distribution (Fig. 5.6(b)), and its range reaches up to approx. three orders of magnitude. Again, the observed discrepancies can be attributed to the microphysical properties of the coarse-mode particles considered by the lidar/OEM approach (see Tab. 4.1).

While estimates of the uncertainty of the effective diameter from the in-situ measurements were not available at the time of writing, a first preliminary comparison between the in-situ effective diameter and the one calculated using the lidar/OEM approach is attempted here. It should be noted that for the effective diameter that was derived from the in-situ measurements, a particle size range of 0.3–50 µm in diameter was used. An effective diameter of 1.7 ± 0.5 µm was obtained from the lidar/OEM approach, when the whole size distribution is considered (dotted black line in Fig. 5.6). When the same calculation (lidar/OEM approach) is carried out with a lower particle-diameter limit of 0.3 µm (same as the in-situ data) the effective diameter is 2.1 µm. Considering the same size range as the in-situ data improves the comparison, especially for the effective diameter obtained from the in-situ Falcon 20E measurements at approx. 6 km altitude (even though in-situ uncertainties are not available). The lower limit of 0.3 µm in diameter is causing an overestimation of the effective diameter derived from the in-situ measurements. Once the particles with smaller size are also considered in the analysis, the in-situ derived effective diameter is expected to be lower.

In conclusion, the OEM-based aerosol typing scheme correctly identified the pronounced lofted aerosol layer (3–6 km) as an aerosol layer primarily consisting of CNS particles. The mixing ratio of the four different aerosol components (determined with the OEM scheme), together with the *a priori* microphysics of the aerosol components, allowed the comparison of the lidar/OEM approach products (i.e., volume and number size distributions and effective radius) with the Falcon 20E in-situ data. The comparison was overall satisfactory, especially when one considers that the OEM-based typing scheme is not designed to retrieve detailed aerosol microphysical properties since it is based on parameterizations.

5.2   Case 2: Smoke and desert dust above Cape Verde

5.2.1   Overview

During the SAMUM–2 campaign, a persistent and complex aerosol layering was observed above Praia, Cape Verde for almost one month (18/01 to 14/02/2008), with the exception of a few days, where clean marine conditions dominated. The vast majority of the aerosol mixtures were observed in lofted aerosol layers and were mixtures of smoke and dust particles, while pure layers were mostly confined in the lowest 1 km and they were mostly pure dust layers. Here, the focus is put on a lofted layer, which was observed at altitudes between 1–3 km, on 22/01/2008, from 20:05–21:00 UTC. The same lofted layer has been analysed in detail by Tesche et al. (2011a).
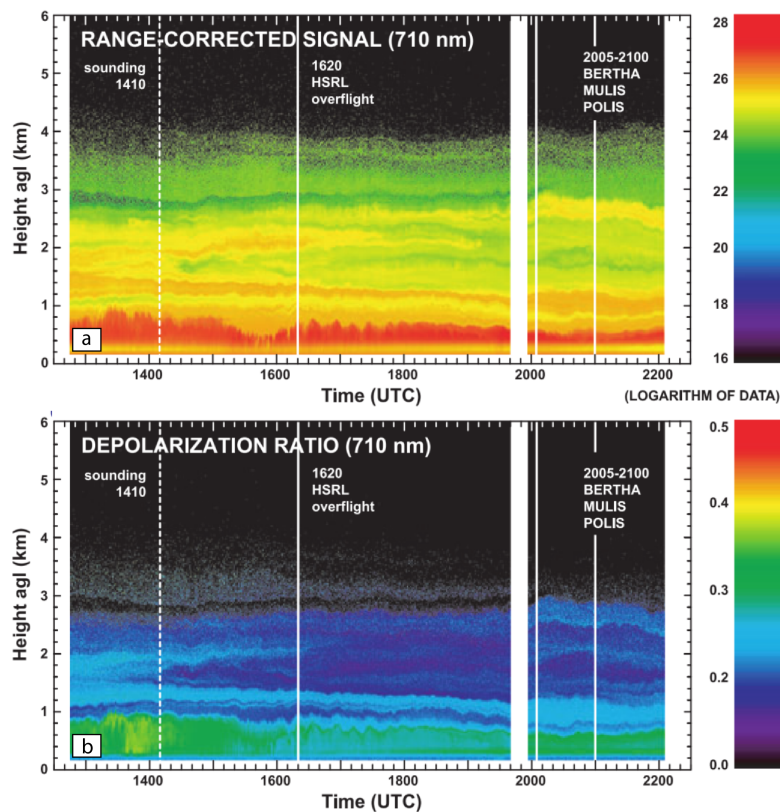
Figure 5.7: Aerosol layer evolution in terms of range-corrected signal (a) and linear volume depolarization ratio (b) at 710 nm at Praia, Cape Verde, on 22 January 2008, from 12:45–22:05 UTC (adapted from Tesche et al. (2011b).

An overview of the atmospheric conditions based on aerosol lidar observations is provided in Fig. 5.7. As the figure indicates, on that day several aerosol layers were observed at altitudes up to 4 km. High linear volume depolarization ratio values in the lowest 0.8 km indicate the presence of desert dust particles (originating from northern Africa). Between 0.8

and 1 km, a decrease of the depolarization ratio is observed. Above 1 km altitude, the lofted layer has lower depolarization ratio values (compared to the the lowest layer in altitude) and it has been characterized as a mixture of mineral dust and biomass-burning/smoke aerosol originating from southern West Africa (Tesche et al., 2009b, 2011a,b).
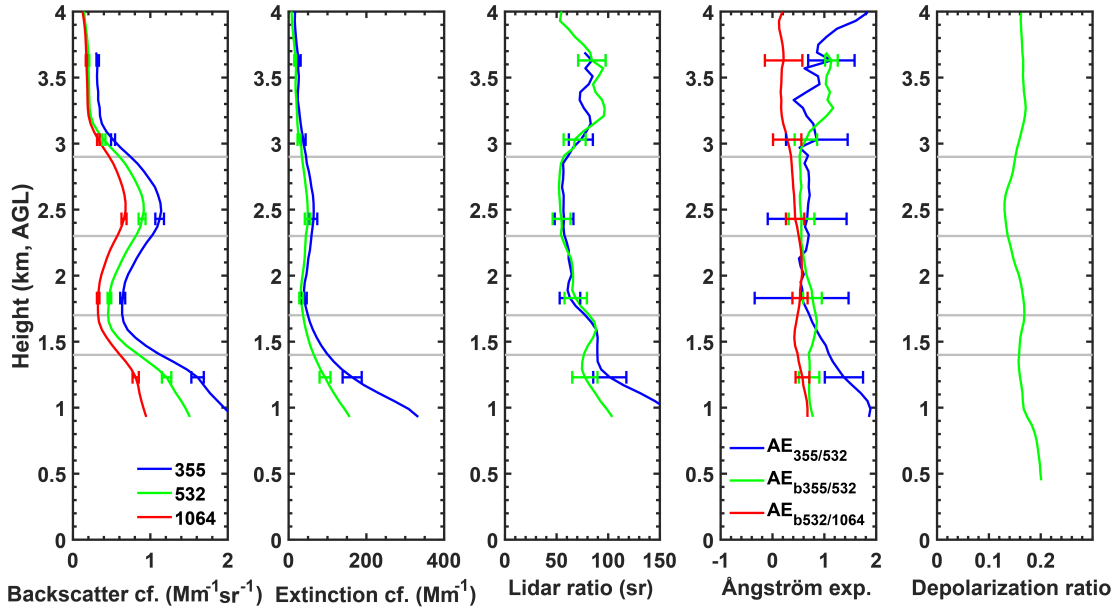
Figure 5.8: Vertical profiles of the particle backscatter and extinction coefficients, particle lidar ratio, Ångström exponents and linear particle depolarization ratio measured at Praia, Cape Verde, on 22 January 2008, between 20:20 and 22:30 UTC. A smoothing length of 660 m has been used (adapted profiles; analysed by Tesche et al. (2011a,b)). Faint grey lines indicate the averaging layer boundaries.

Profiles of the backscatter and extinction coefficients, the lidar ratio, the Ångström exponent and the linear particle depolarization ratio are presented in Fig. 5.8 (from left to right, respectively) for a 2-h interval (20:20–22:30 UTC). As the figure indicates, between 1–3 km altitude, two aerosol layers can be identified. The aerosol layers boundaries have been defined based on the retrieved optical properties, as 1.4–1.7 km for the lower aerosol layer and 2.3–2.9 km for the upper aerosol layer. The averaged lidar ratios for the lower aerosol layer are 85.6 ± 13.5 sr and 84.2 ± 13.3 sr at 355 and 532 nm, respectively. The mean particle linear depolarization ratio at 532 nm for that layer is 0.16. The backscatter-related Ångström exponents at the wavelength pairs of 355/532 and 532/1064 nm exhibit values that range between 0.5 and 1. The optical properties of this aerosol layer are indicative of an aerosol mixture that is rather absorbing and contains both spherical and non-spherical particles. The upper layer exhibits mean lidar ratios of 57.0 ± 9.0 sr and 53.9 ± 8.5 sr at 355 and 532 nm, respectively. The layer-averaged particle linear depolarization ratio (532 nm) is 0.14. Both the extinction- and backscatter-related Ångström exponents exhibit values

that range between 0.5 and 1, similar to the ones observed for the lower layer. The optical properties of the upper aerosol layer indicate again an aerosol mixture of spherical and non-spherical particles that is less absorbing compared to the lower aerosol layer. However, based solely on the aerosol measurements the mixing ratio of the different aerosols contributing to the aerosol mixture observed cannot be determined.

5.2.2   Aerosol characterization by the OEM-based typing scheme

To characterize the mixing ratio (in terms of relative volume) of the different aerosol components contributing to the observed aerosol mixtures, the OEM-based aerosol typing scheme was applied to the two aforementioned aerosol layers (Fig. 5.8). Due to the absence of particle linear depolarization ratio information at 355 nm the only possible retrieval modes were the modes 2 and 4 (see Tab. 4.4). In addition to the aforementioned optical properties at 532 nm, backscatter-related color ratios of 1.4 ± 0.5 and 1.3 ± 0.5 (for layer 1 and layer 2, respectively, both at the wavelength pair 532/1064 nm) were used as inputs to the OEM-based typing scheme. It should be noted that the errors of the particle linear depolarization ratio (0.05) and the backscatter-related color ratio (0.5) had to be assumed (due to the absence of data) to enable the OEM retrievals.

The outcome of the different retrieval modes for the two aerosol layers is presented in Tab. 5.3. Both retrieval modes produced results that were statistically significant within the 95% confidence interval for both aerosol layers and are therefore considered for further analysis. It can be seen that in both aerosol layers, regardless of the retrieval mode, the CNS aerosol component dominated in the aerosol mixture, in terms of relative volume, with values ranging from approx. 67% to 78%. The second most abundant aerosol component was the FSA one (≈ 26–29%) for the lower aerosol layer (layer 1) and the FSNA one (≈ 14–18%) for the upper aerosol layer (layer 2). The aerosol components least present in the lower aerosol layer were the CS and and FSNA ones, taking into account the retrieval uncertainties. For the upper aerosol layer, the least present aerosol components were the CS and FSA ones. The results (Tab. 5.3) for both aerosol layers are consistent for both retrieval modes (2 and 4).

For the lower aerosol layer (layer 1) both retrieval modes identified the aerosol mixture as a mixture of only two aerosol components; CNS and FSA. Retrieval mode 2, which takes into account only the lidar ratio and particle linear depolarization ratio at 532 nm, identified a CNS contribution of $67.3 \pm 21.4\%$ and a FSA contribution of $25.8 \pm 15.4\%$ (both in terms of relative volume). Overall, the relative volume contribution of coarse-mode particles was 2.5 times higher than the one of the fine-mode particles. While this solution is the statistically most likely solution to produce the given measurements, the aerosol components' contributions add up to a relative volume equal to 93.1%. Therefore, the remaining 6.9%

Table 5.3: Contribution and respective uncertainties of the four basic aerosol components to the aerosol mixtures observed at Praia, Cape Verde, on 22/01/2008, between 1.4–1.7 km (layer 1) and 2.3–2.9 km (layer 2), from 20:05–21:00 UTC. The retrieval mode is also indicated.

Aerosol layer	Retrieval mode	OEM output (%)			
		FSA	CS	FSNA	CNS
1	2	25.8 ± 15.4	0 ± 14.8	0 ± 17.6	67.3 ± 21.4
2	2	1.7 ± 11.7	6.3 ± 14.3	14.3 ± 17.7	77.7 ± 22
1	4	28.5 ± 16	0 ± 14.9	0 ± 17.2	67.8 ± 21.3
2	4	1.7 ± 12.1	6.6 ± 14.8	17.8 ± 17.6	73.9 ± 21.9

relative volume can be attributed as uncategorized/unknown aerosol. In addition to the optical parameters considered in retrieval mode 2, retrieval mode 4 takes also into account the backscatter-related color ratio at the wavelength pair 532/1064 nm. The relative volume contribution of the CNS aerosol component was $67.8 \pm 21.3\%$ for retrieval mode 4 (almost identical with retrieval mode 2) and $28.5 \pm 16\%$ for the FSA aerosol component (slightly increased compared to retrieval mode 4). The mixing ratio of coarse- and fine-mode particles slightly decreased compared to retrieval mode 2 (approx. 2.4). For this retrieval mode, the relative volume of the uncategorized/unknown aerosol was 3.7% (lower compared to retrieval mode 2). Given that both retrieval modes produced statistically significant results for this layer and that retrieval mode 4 characterized a lower percentage of aerosol as uncategorized/unknown compared to retrieval mode 2, we can conclude that retrieval mode 4, which includes more optical parameters compared to retrieval mode 2, is preferred. In general, when the quality of the input data is good, the retrieval mode that uses the most optical parameters is to be preferred.

The aerosol mixture observed in the upper layer (layer 2) was identified by retrieval mode 2 as a mixture of primarily CNS and FSNA aerosol with relative volume contributions of $77.7 \pm 22\%$ and $14.3 \pm 17.7\%$, respectively. CS and FSA particles also contributed to the aerosol mixture with much smaller relative volume contributions ($6.3 \pm 14.3\%$ and $1.7 \pm 11.7\%$, respectively). The relative volume contribution of the coarse-mode particles was 84%, and it was 16% for the fine-mode particles. For the same aerosol layer, the results of retrieval mode 4 are similar to those of retrieval mode 2. The dominant aerosol component was the CNS ($73.9 \pm 21.9\%$), followed by the FSNA ($17.8 \pm 17.6\%$). CS and FSA contributions to the aerosol mixture were small ($6.6 \pm 14.8\%$ and $1.7 \pm 12.1\%$, respectively). The mixing ratio of coarse- to fine-mode particles was approx. 4.

To summarize the OEM results, both aerosol layers were identified by both retrieval modes as aerosol mixtures, where coarse aerosol particles dominate in terms of relative

contribution. The lower aerosol layer appears to be a mixture of CNS and FSA aerosol particles with a mixing ratio of approx. 1.8. The upper layer appears to be a mixture of CNS and FSNA aerosol particles with a mixing ratio of approx. 4. For this specific case, the consistency between the results obtained with the two retrieval modes (2 and 4) implies that the additional optical information (backscatter-related color ratio in the case of retrieval mode 4) did not significantly change the outcome, but also did not obstruct the analysis. In the following section, the OEM results for the two aerosol layers are compared with the results of the POLIPHON method to investigate the consistency between the two methodologies.

5.2.3   Comparison with POLIPHON

Following the POLIPHON methodology (Sec. 2.1), the dust and smoke contributions to the measured total backscatter and extinction coefficient at 532 nm for the lofted aerosol layers observed on 22/01/2008 were calculated and presented by Tesche et al. (2011b). To compare the POLIPHON and OEM results, the OEM retrieval outputs had to be transformed into backscatter and extinction fractions. First, the extinction and backscatter coefficient per aerosol component was calculated as described in Sec. 4.2. Then, the fraction attributable to the different components was computed with respect to the total OEM-retrieved backscatter and extinction coefficients. Table 5.4 summarizes the results of the aforementioned transformation. In addition, to facilitate the POLIPHON-OEM comparison, the POLIPHON-derived backscatter and extinction fraction vertical profiles had to be averaged for the two examined aerosol layers.

Table 5.4: Contributing fractions of the different aerosol components to the total OEM-estimated backscatter and extinction coefficient at 532 nm for the two aerosol layers observed at Praia, Cape Verde, on 22 January 2008, from 20:20–22:30 UTC.

Layer	Ret. mode	Backscatter coef. 532 nm (%)				Extinction coef. 532 nm (%)			
		FSA	CS	FSNA	CNS	FSA	CS	FSNA	CNS
1	2	59.8	0	0	40.2	71.7	0	0	28.3
2	2	3.9	10.2	40.2	45.7	6.7	3.6	43.6	46.1
1	4	62.1	0	0	37.9	73.6	0	0	26.4
2	4	3.7	9.8	46.3	40.2	6.3	3.4	50.1	40.2

In the OEM-based typing scheme, the CNS aerosol component resembles desert dust particles and, therefore, can be directly compared with the dust fractions derived with the POLIPHON method. Both the FSA and FSNA aerosol components can resemble smoke particles with different absorption properties, such as in the case of smoke from different origins, e.g., from flaming or smoldering fires. Therefore, to allow for a fair POLIPHON-

OEM comparison, the contributions of the FSA and FSNA aerosol components have been summed up first (only applicable for aerosol layer 2). In addition, the contributions of the CS aerosol component (only applicable for aerosol layer 2) are not taken into account for this comparison, since POLIPHON has attributed the non-dust contributions to smoke particles entirely. The POLIPHON-OEM comparison results are summarized in Figure 5.9.
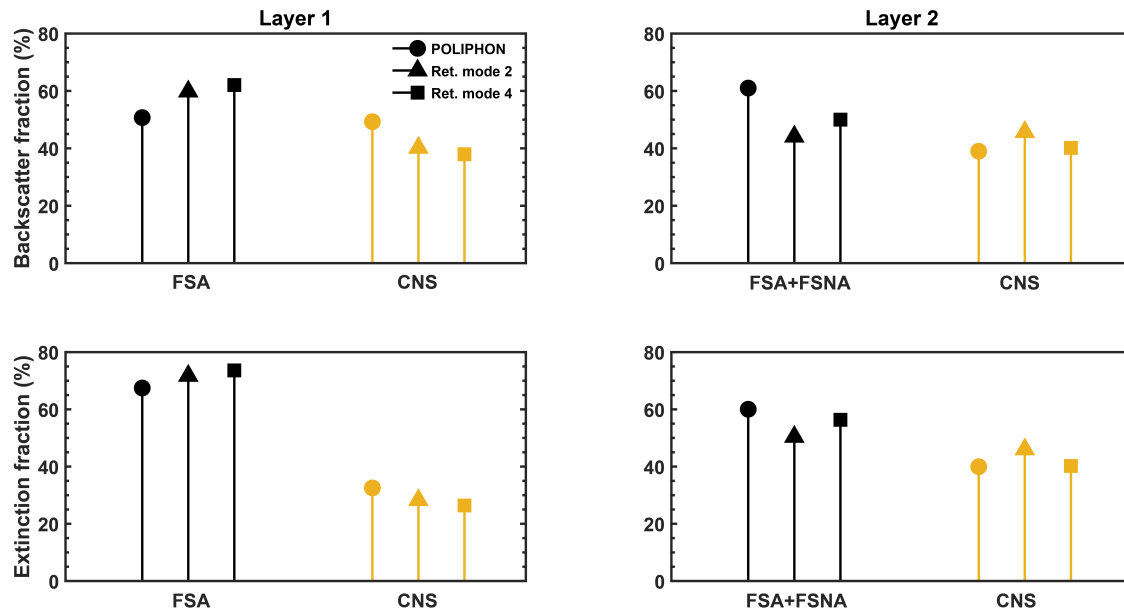
Figure 5.9: POLIPHON-OEM comparison of the fractions of backscatter (top panels) and extinction (bottom panels) coefficients for the lower (layer 1, left panels) and upper (layer 2, right panels) aerosol layers at Praia, Cape Verde on 22 January 2008, between 20:20 and 22:30 UTC.

The OEM-derived results for the lower aerosol layer (layer 1, left panels in Fig. 5.9) agree to a satisfactory level with the POLIPHON results. POLIPHON results (stem with circle) indicate that the dust (CNS) and smoke (FSA) aerosol particles contributed almost equally to the backscatter coefficient (50.7% and 49.3%, respectively), while the smoke particles dominated the extinction coefficient with contributions of 67.5%. The OEM results are rather similar for both retrieval modes (2 and 4 indicated by stems with triangle and square, respectively) and indicate that the FSA particles dominated both the backscatter and extinction coefficients with contributions of approx. 60% and 72%, respectively.

The POLIPHON results for the upper aerosol layer (layer 2, right panels in Fig. 5.9) indicate that the smoke particles dominated both the backscatter and extinction coefficients with contributions of 61% and 60%, respectively. On the other hand, for retrieval mode 2, the OEM results revealed that the sum of the FSA and FSNA aerosol components (smoke particles) and the CNS particles contributed almost equally to the backscatter (44.1% and

45.7%, respectively) and extinction coefficients (50.3% and 46.1%, respectively). For the same retrieval mode, small contributions of CS aerosol particles were also evident (see Tab. 5.4, not shown in Fig. 5.9). OEM results from retrieval mode 4 compared much better with the POLIPHON results; the sum of the FSA and FSNA aerosol components dominated both the backscatter and extinction coefficients with contributions of 50% and 56.4%, respectively. The contribution of the CNS particles was 40.2% for both the backscatter and extinction coefficients, which in addition compared very well with the respective dust fractions from POLIPHON. Similarly to retrieval mode 2, in retrieval mode 4, the CS aerosol component had small contributions to both the backscatter and extinction coefficients.
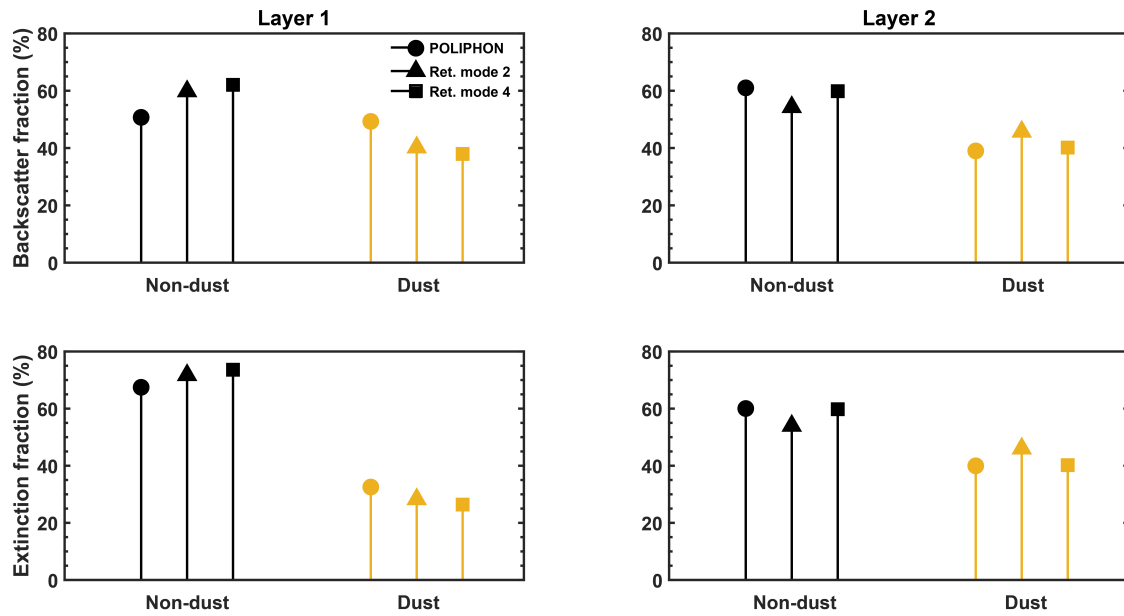
Figure 5.10: Same as Fig. 5.9, but including the fractions from the CS particles and separated as dust and non-dust components.

As mentioned earlier (see Sec. 2.1), POLIPHON is able to separate only dust and non-dust components (in this case smoke). The POLIPHON-OEM comparison presented in Fig. 5.10 follows the aforementioned notation. The non-dust fractions consist of contributions of the FSA, FSNA and CS aerosol particles. The lower aerosol layer (layer 1) consists only of FSA and CNS aerosol components (according to the OEM-typing scheme) and naturally has not been affected by this notation change. However, for the upper layer (layer 2) the OEM-retrieved non-dust fractions of backscatter and extinction coefficients fit very well with the POLIPHON results. The non-dust component clearly dominates now both the backscatter and extinction coefficients for both retrieval modes. That is especially true for the retrieval mode 4, where the results from both methodologies are almost identical. The non-dust particles dominate the backscatter (61% for POLIPHON and 59.8% for OEM)

and extinction (60% for POLIPHON and 59.8% for OEM) coefficients.

Overall, the comparison between the POLIPHON and the OEM results is satisfactory. The small discrepancies observed (e.g., in the upper aerosol layer, Fig. 5.9) originate from the differences in the two methodologies. POLIPHON is able to separate only dust and non-dust components (in this case smoke, see Sec. 2.1), while the OEM-based aerosol typing scheme can separate up to four different aerosol components, including two fine-mode and two coarse-mode components. When the notation between the two methodologies was harmonized, the results compared very well to each other (see upper aerosol layer, retrieval mode 4 in Fig. 5.10). Overall, the advantage of OEM compared to POLIPHON is that OEM provides information on the radiative properties of the aerosol particles (i.e., whether the particles are absorbing or not). For instance, regarding the upper aerosol layer, POLIPHON assigned the entire amount of non-dust particles to smoke, while OEM provided the most-likely contributions of the remaining three components. For the given optical properties, OEM suggested that the non-dust particles in the upper aerosol layer (see Tab. 5.3) are more likely to be less absorbing (FSNA) than strongly absorbing (FSA).

5.3   Case 3: Californian smoke above Central Europe

5.3.1   Overview

On 11 September 2020, around 03:00 UTC, lofted aerosol layers arrived above Leipzig, Germany and were observed with a Polly^{XT} at an altitude of approx. 12 km (Fig. 5.11). These aerosol layers contained smoke particles, which originated from wildfires that occurred 3–4 days prior to the observations at the US west coast (Baars et al., 2021). Later on the same day, the smoke layers were observed throughout the whole troposphere down to an altitude of approx. 4.5 km. Overall, smoke layers from this event were observed above Leipzig until 14 September 2020.
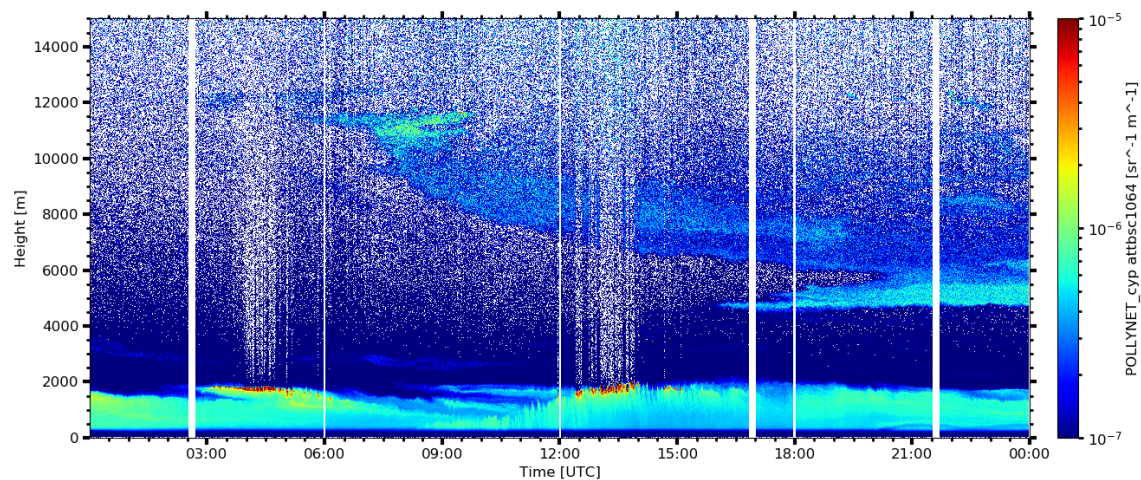
Figure 5.11: Overview of the atmospheric conditions, in terms of range-corrected signal at 1064 nm over Leipzig, Germany on 11/09/2020, between 00:00 and 22:00 UTC. No data are available during regular depolarization calibration periods (white bars).

Fig. 5.12 (vertical resolution: 500 m) shows the optical properties of the smoke layer, as observed on 11 September 2020, between 21:40–22:40 UTC, and retrieved using the Raman method (Ansmann et al., 1992). The backscatter coefficient shows a strong wavelength dependence (355–532 nm spectral range), while at the same time, the extinction coefficient exhibits a much less prominent wavelength dependency (at same spectral range). Therefore, the lidar ratio at 532 nm is larger than the respective one at 355 nm. This feature incorporates the most significant signature for aged smoke (Murayama et al., 2004; Müller et al., 2005; Tesche et al., 2011b; Baars et al., 2012; Giannakaki et al., 2015; Haarig et al., 2018). The extinction-related Ångström exponent exhibits values between zero and one, indicating that the aerosol particles are medium in size. The particle linear depolarization ratio exhibits low values (from 2–4% for both wavelengths) up to approx. 8 km altitude. Above that altitude, it increases to values up to 5.7% and 9.2% at 355 and 532 nm, respectively. The

low particle linear depolarization ratio values indicate the presence of spherically shaped particles. The slightly elevated particle linear depolarization ratio values might occur due to the presence of rather irregularly shaped soot or soil dust particles (Nisantzi et al., 2014) mixed with partly coated soot particles.
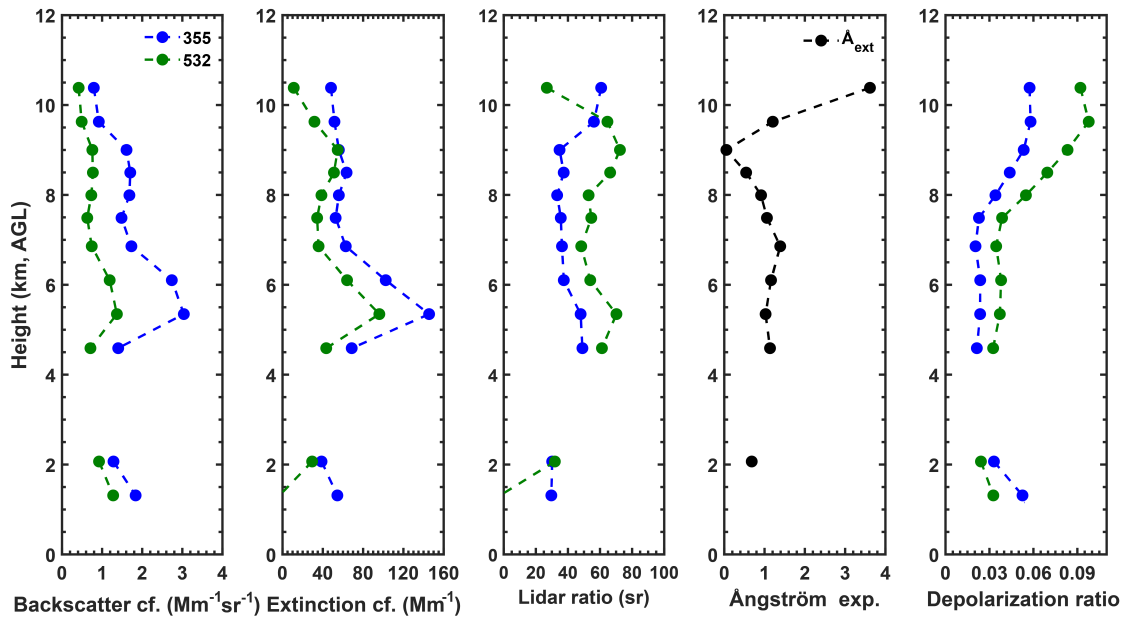
Figure 5.12: 1-h-mean profiles (21:40–22:40 UTC) of optical properties in the smoke layer observed on 11 September 2020 above Leipzig, Germany. From left to right: backscatter coefficient at 355 and 532 nm (blue and green, respectively), extinction coefficient (355 and 532 nm), lidar ratio (355 and 532 nm), extinction-related Ångström exponent and particle linear depolarization ratio (355 and 532 nm).

5.3.2   Aerosol characterization by the OEM scheme

The optical properties presented in Fig. 5.12 (filled circles) are layer-mean values (500 m averaged, 12 aerosol layers in total) and were used as input to the OEM-based typing scheme. The OEM scheme can, therefore, evaluate the whole profile. Tab. 5.5 presents the optical properties of the respective aerosol layers (OEM input) along with measurement errors and altitude information. Since measurement errors were not available, to be able to use the OEM typing scheme, the measurement errors had to be assumed. An error of 20% of each respective measurement was assumed (large enough to capture the variability of the optical parameters, e.g., errors for the backscatter and extinction coefficients are in the range of 5–10% and 10–20%, respectively (Althausen et al., 2009; Baars et al., 2012; Engelmann et al., 2016)).

For almost all layers (except the lowest one in altitude) all OEM input parameters were

Table 5.5: Optical properties of the different aerosol layers observed at Leipzig, Germany, on 11/09/2020. Unit for altitude range is in km, δ is expressed in % and S in sr.

Layer	Altitude	Optical properties					
		δ_{355}	S_{355}	$\mathring{A}$	δ_{532}	S_{532}	$C_{\beta_{532/1064}}$
1	1.1–1.6	5 ± 1	29.7 ± 5.9	$-$	3 ± 0.6	$-$	3.1 ± 0.6
2	1.8–2.3	3 ± 0.6	30.2 ± 6.04	0.7 ± 0.14	2.5 ± 0.5	31.8 ± 6.4	3.8 ± 0.76
3	4.3–4.8	2 ± 0.4	49 ± 9.8	1.1 ± 0.22	3 ± 0.6	61.2 ± 12.2	4.1 ± 0.82
4	5.1–5.6	2 ± 0.4	48 ± 9.6	1 ± 0.2	3.7 ± 0.7	70.2 ± 14	3.7 ± 0.74
5	5.8–6.3	2.4 ± 0.5	37.4 ± 7.5	1.2 ± 0.24	3.8 ± 0.76	54 ± 10.8	3.7 ± 0.74
6	6.6–7.1	2 ± 0.4	36.3 ± 7.3	1.4 ± 0.28	3.5 ± 0.7	48.3 ± 9.7	3.8 ± 0.76
7	7.2–7.7	2 ± 0.4	35.5 ± 7.1	1.1 ± 0.22	3.8 ± 0.76	54.5 ± 10.9	3.7 ± 0.74
8	7.7–8.2	3 ± 0.6	33.3 ± 6.66	0.9 ± 0.18	5.5 ± 1.1	53 ± 10.6	3.6 ± 0.72
9	8.2–8.7	4.4 ± 0.9	37.4 ± 7.5	0.54 ± 0.11	7 ± 1.4	66.3 ± 13.3	3.6 ± 0.72
10	8.7–9.2	5.3 ± 1.1	34.8 ± 7	0.05 ± 0.01	8 ± 1.6	72.4 ± 14.5	3.5 ± 0.7
11	9.3–9.8	5.8 ± 1.2	56.2 ± 11.2	1.2 ± 0.24	9.7 ± 1.9	64.6 ± 12.9	3.5 ± 0.7
12	10.1–10.6	5.7 ± 1.1	60.7 ± 12.1	3.6 ± 0.7	9.2 ± 1.8	26.8 ± 5.4	3.3 ± 0.66

available and, therefore, all retrieval modes (six in total, Tab. 4.4) were performed. For layer 1, only one retrieval mode was available (retrieval mode 1 in Tab. 4.4). In total, 64 retrievals were performed, out of which 34 were found to be statistically significant (approx. 53% of the cases). The remaining retrievals were not statistically significant and three of them failed to converge.

An overview of all the retrieval results, regardless of their statistical significance, is presented in Fig. 5.13. With a quick glance, one can see that in the majority of the layers the dominant aerosol component was the CS one. The CS predominance is especially evident for retrieval modes 1, 3, 5 and 6 (Fig. 5.13 a, c, e and g, respectively). The low values of the particle depolarization ratio along with the low values of the lidar ratio (both at 355 nm) resulted in a CS-dominated initial guess of the state vector (see Sec. 4.1.4) and in several cases a statistically significant solution was found. In the majority of the cases, the CS-dominated aerosol mixture has also high contributions of the CNS aerosol component. However, layers 3, 4, 11 and 12 were the exceptions (due to the much higher lidar ratio values), as the dominating aerosol component was the FSNA. In particular, the lowest (in altitude) two aerosol layers were identified as FSNA/CS mixture, a mixture that could be associated with aged smoke particles that have undergone some hygroscopic growth and coating, but the results were not statistically significant. On the other hand, the highest two (in altitude) aerosol layers were primarily composed of FSNA particles, mixed mainly with CNS and FSA particles. The results of retrieval modes 1 and 3 were statistically

significant for the aerosol layer 11. None of the retrievals performed for the aerosol layer 12 was found to be statistically significant.

The FSNA and FSA aerosol components appear to dominate the results retrieved using the retrieval modes 2 and 4 of the OEM-based aerosol typing scheme. Large contributions of CNS particles were also evident, especially for the higher (in altitude) aerosol layers. For these two retrieval modes, the higher lidar ratio (at 532 nm) resulted in a more FSNA-dominated initial guess of the state vector compared to the rest of the retrieval modes. For retrieval mode 2, all layer retrievals except for layer 12 were found to be statistically significant. While the results obtained with retrieval mode 4 are overall consistent with the respective ones of retrieval mode 2, only results for layers 5, 6, 7 and 8 were found to be statistically significant.
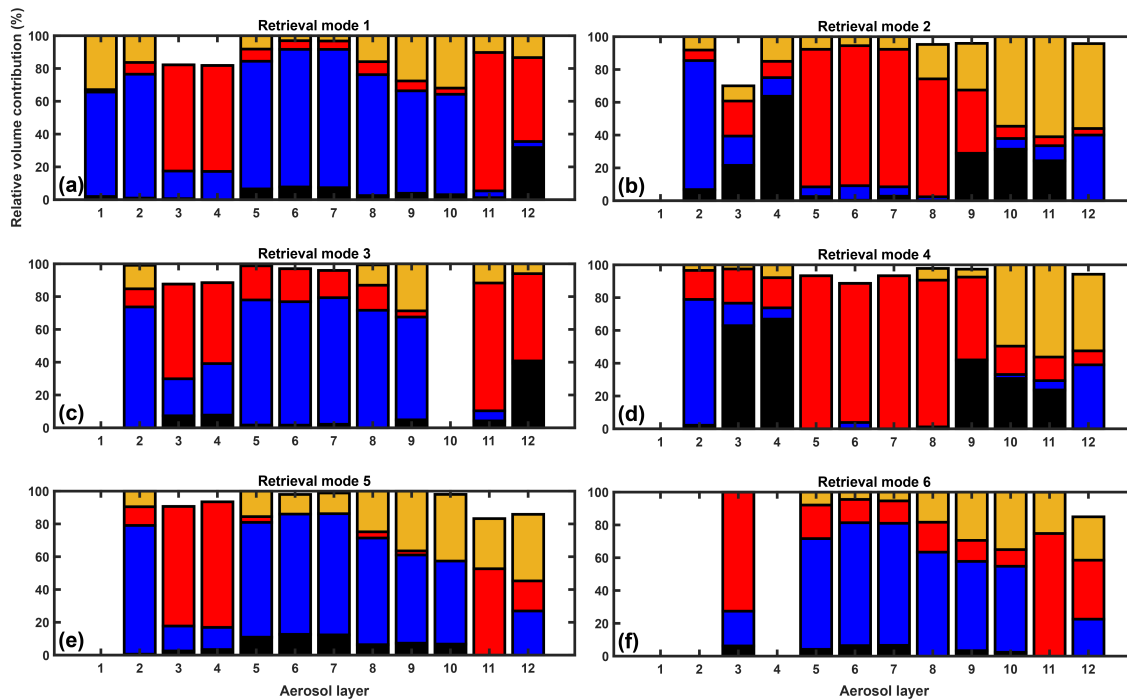
Figure 5.13: Relative volume contributions of the four different aerosol components contributing to the aerosol layers of Fig. 5.12 for the different retrieval modes. Black, blue, red and orange colors indicate FSA, CS, FSNA and CNS particles, respectively.

The initial guess of the state vector has clearly affected the optimal state (i.e., the optimal solution). In retrieval modes 1, 3, 5 and 6 the selection of the initial guess is based on the lidar ratio and particle linear depolarization ratio at 355 nm and, according to the decision tree (Sec. 4.1.4), a CS-dominated initial guess was assigned to the majority of the aerosol layers. On the other hand, in retrieval modes 2 and 4 the choice of the initial guess is made based on the same information but at 532 nm. That results in a FSNA- or
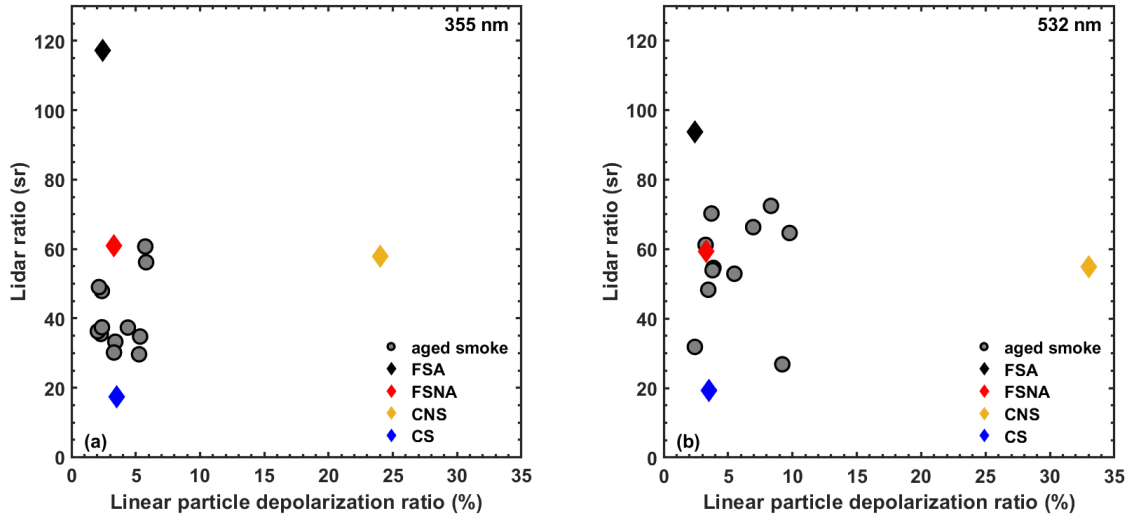
Figure 5.14: Lidar ratio versus linear particle depolarization ratio (at 355 and 532 nm) of the four basic aerosol components considered in the current version of the OEM-based typing scheme and of the aged smoke particles observed on 11/09/2020.

FSA-dominated initial guess for the majority of the aerosol layers.

Apart of the importance of the initial guess of the state vector, inconsistencies between the different retrieval modes also reveal that the nature of the aged smoke particles is not well captured by the OEM-based aerosol typing scheme as it is. For instance, layer 5, which was observed in an altitude of approx. 6 km, is clearly consisting of aged smoke particles, as the optical properties indicate (Tab. 5.5). According to retrieval modes 1, 3, 5 and 6 (all statistically significant except 6) the layer consisted mainly of CS particles, with contributions of at least 70% in terms of relative volume. On the other hand, retrieval mode 2 and 4 (statistically significant results) identified a FSNA dominance of at least 83% in terms of relative volume contributions. Clearly, neither the CS nor the FSNA aerosol component can represent optically and microphysically the aged smoke particles.

The correct identification and quantification of the relative volume contribution of the aged smoke particles by the OEM-based aerosol typing scheme requires the introduction of at least one new aerosol component to the typing scheme. The necessity of the introduction of a new aerosol component is pointed out in Fig. 5.14 and 5.15. As the figures indicate, the optical properties of the aged smoke particles do not match with either the FSA nor the FSNA aerosol components. The backscatter coefficient of the new aerosol component should exhibit a clear and strong wavelength dependence for the 355-532 nm range. In the same spectral range, the wavelength dependence of the extinction coefficient should be weaker. The particle linear depolarization ratio for aged smoke particles present in the troposphere is typically low and does not show a strong wavelength dependence. However,
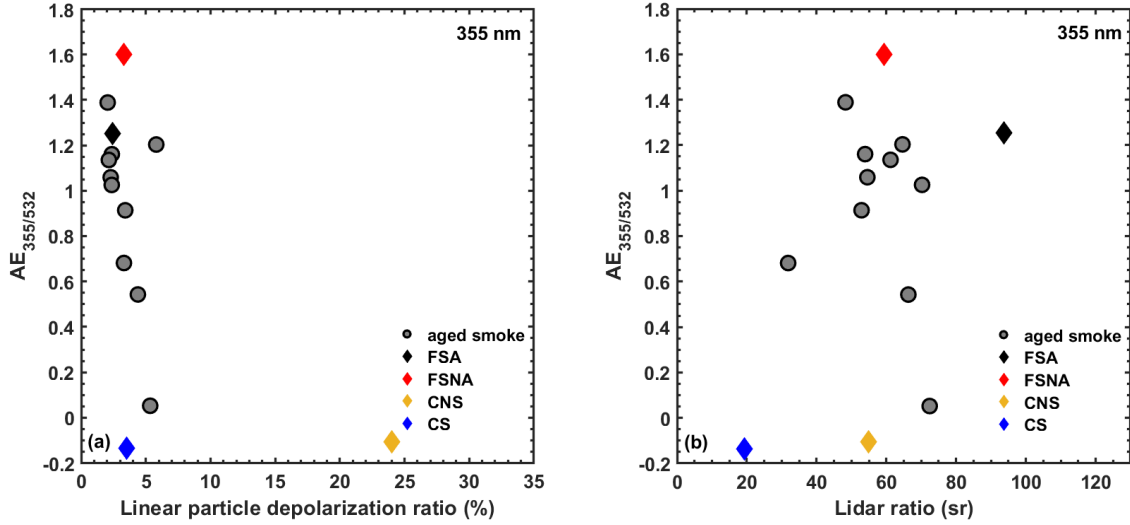
Figure 5.15: Same as Fig. 5.14, but for the extinction-related Ångström exponent versus the particle linear depolarization ratio (left) and lidar ratio (right) at 355 nm.

when the particles are present in the stratosphere, the particle linear depolarization ratio is considerably higher and exhibits a strong wavelength dependence (Burton et al., 2015; Haarig et al., 2018; Hu et al., 2019; Ohneiser et al., 2020). The reason for this behavior has been hypothesized by Haarig et al. (2018). The fast vertical transport of the fire smoke to the upper troposphere and lower stratosphere by pyroconvection has possibly prohibited the interaction of the particles with gases (coating) and/or the internal mixing with other aerosol particles and, therefore, only the pure and irregularly shaped soot particles entered the rather dry lower stratosphere. Previous studies (e.g., Müller et al. (2005); Wandinger et al. (2002)) have shown that the size of the aged biomass-burning particles tends to be larger than the particle size of anthropogenic pollution aerosol. Therefore, the new aerosol component should be of medium size (e.g., $R_{\text{eff}} > 0.2$ µm). The real and imaginary part of the refractive index of the component should be comparable (same order of magnitude) with the FSA aerosol component. Given the rather peculiar nature of those particles, a further splitting of the new aerosol component into two separate components might be necessary also from a microphysical point of view (i.e., tropospheric aged smoke particles with spherical and stratospheric aged smoke particles with irregular shape). The introduction of such new particle types requires a more detailed research, as the topic is quite new and it is beyond the scope of this thesis, although can be worked on in a follow-up project.

5.4   Conclusions

The application of the OEM-based aerosol typing scheme in the case studies presented in this chapter demonstrated not only the algorithm's strengths but also its limitations. It was shown that for an aerosol layer, the mixing ratio (in terms of relative volume) of up to four different aerosol components can be adequately determined. The determination of the mixing ratio was achieved by utilizing several optical properties in various combinations. The flexibility of the forward-model operator is a huge advantage of this methodology, as it allows the usage of OEM in a wide-ranging extent. The aerosol layers discussed in the case studies above covered various scenarios of atmospheric load and the algorithm was able (in most cases) to distinguish the aerosol components present in those aerosol layers. The determination of the mixing ratio was more challenging in the case of aerosol layers consisting of aerosol particles that are not directly resembled by an aerosol component of the OEM-based typing scheme (e.g., see Sec. 5.3). In such cases, while the resulting mixing ratios were statistically significant, they clearly did not reflect the atmospheric state, as there is no unique solution that can reproduce the measurements. To improve the results for such cases, the introduction of a new aerosol component is necessary, but beyond the scope of this thesis.

Chapter 6

Application - Part II: Long-term lidar observations

The structure of the OEM-based aerosol typing algorithm is such that it allows its application not only to case-specific aerosol layers (see Ch. 5) but also to long-term observations, making it thus ideal for aerosol monitoring and characterization purposes. Therefore, in this Chapter the application of the OEM-based typing scheme to long-term lidar observations that were conducted in Haifa, Israel (March 2017–May 2019) is presented and discussed. First, information regarding location, instrumentation and data analysis methods is provided. Afterwards, the aerosol mixtures observed above Haifa are identified and characterized in terms of mixing ratio based on the OEM-based aerosol typing scheme. Finally, the seasonal variability of the aerosol components is presented, along with height-resolved air-mass source attributions, which were performed using TRACE (Radenz et al., 2021b). Results presented in this Chapter have been published in Heese et al. (2022).

6.1 Haifa, Israel: location, instrumentation and data handling

Haifa is a coastal city located in Israel (Fig. 6.1) and situated between the northern slopes of Mount Carmel and the Haifa Bay (eastern shore of the Mediterranean Sea/Levantine Sea). The city itself is neighbouring to the east with big industries (e.g., large petrochemical plants, oil-fueled power station, cement factories) as well as smaller industries and workshops (Ganor et al., 1998).

Due to its location, Haifa receives air masses that firstly cross the Mediterranean basin and typically originate from Europe, Africa and the Atlantic Ocean (Alpert et al., 1990). In addition to locally produced pollution-associated aerosol particles (from the industries in the

vicinity of the city), the air masses carry long-range transported pollution-related particles (Europe), desert dust particles from the Saharan region (Africa) as well as marine particles (Atlantic Ocean and Mediterranean Sea) to Haifa. An additional source for desert dust particles is the Arabian peninsula (Kubilay et al., 2000). The different aerosol types and the mixtures that they create have been studied mainly by means of in-situ measurements (e.g., Ganor et al., 1998; Koçak et al., 2004) and thus, there was a knowledge gap regarding the vertical aerosol distribution.

For that reason and in collaboration with the Viterby Faculty of Electrical Engineering at the Technion in Haifa, continuous (24/7) and height-resolved lidar measurements of the vertical aerosol distribution have been performed for two years, from March 2017 until May 2019. The location of the lidar measurement site is shown in Fig. 6.1 (right panel, yellow star). A multi-wavelength Raman and polarization lidar Polly$^{\text{XT}}$ (Engelmann et al., 2016) was situated at 32.8° N and 35.0° E at an altitude of 230 m ASL. A near-range telescope allows the detection of scattered light at 355, 387, 532 and 607 nm from about 60–80 m AGL (above ground level). Vertical profiles can be obtained with an uppermost detection height of around 20 km AGL. This setup allows the determination of the backscatter coefficient at 355, 532 and 1064 nm and the extinction coefficient at 355 and 532 nm, as described by Ansmann and Müller (2005). Typically, uncertainties for the Raman-retrieved backscatter and extinction coefficients are in the range of 5–10% and 10–20%, respectively (Althausen et al., 2009; Baars et al., 2012; Engelmann et al., 2016). The lidar ratio, Ångström exponent, backscatter-related color ratio as well as the particle linear depolarization ratio can also be determined according to standard procedures applied in EARLINET (D'Amico et al., 2015; Mattis et al., 2016). Data are quality assured, since quality assurance procedures are a key aspect within the EARLINET network (e.g., Wandinger et al., 2016b; Freudenthaler, 2016; Belegante et al., 2018) and are followed even when Polly lidar systems are operated at non-stationary sites.

For the period March 2017 to May 2019, in total 397 lidar aerosol profiles were analysed (Mewes, 2018) and about 1013 aerosol layers were manually identified. However, not all aerosol layers were suitable for the OEM-based typing scheme. Input criteria require the availability of height-averaged lidar ratio and linear particle depolarization ratio at least at one wavelength (either 355 or 532 nm), as already mentioned in Sec. 4.1.5 (see Tab. 4.4). Even though the optical information available was very diverse, given OEM's flexibility in terms of optical parameters and their different combinations that can be used as input, a total of 659 aerosol layers were found suitable for further analysis. More precisely, 659 aerosol layers were found suitable for further analysis, out of which 474 led to statistically significant retrievals (71.9%). To be more specific, 57 statistically significant solutions were obtained by using OEM's retrieval mode 1, 359 by using retrieval mode 2 and 57 by
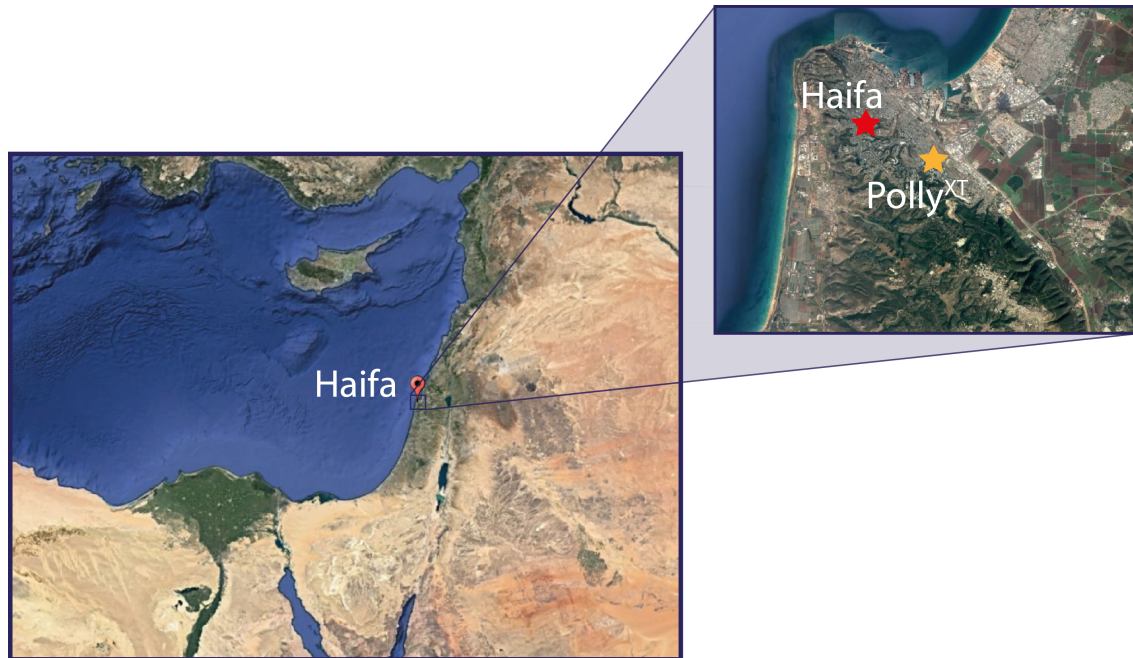
Figure 6.1: Map of the Levantine Sea, showing the location of Haifa, Israel (left). The zoomed map (right) shows Haifa (red star) and the location of the Polly$^{\text{XT}}$ lidar (yellow star). Source: Google Earth (22/03/2021).

using retrieval mode 5 (see Tab. 4.4). In this case, the fact that the vast majority of the statistically significant retrievals is an outcome of the retrieval mode 2 is only reflecting the available input parameters; most aerosol layers were identified and analysed based on optical information at 532 nm. In the following sections, the only layers further considered and discussed are the ones that led to a statistically significant retrieval.

6.2 Aerosol mixture identification

An overview of the altitude and extent of the examined layers is shown in Fig. 6.2. It should be noted again that only the layers that led to statistically significant retrieval results are shown here. These aerosol layers were mostly confined in the lowest parts of the troposphere, on average at altitudes between 740 m and 2 km. However, there are several lofted layers observed at high altitudes, especially between 2 and 4 km but also above 4 km. At this point, it should be mentioned that one layer that was observed in spring 2018 at an altitude between 11 and 12 km is not shown in Fig. 6.2 (for visual clarity of the figure) but is considered in the analysis.

The layers of Fig. 6.2 have been color-coded according to the aerosol component that
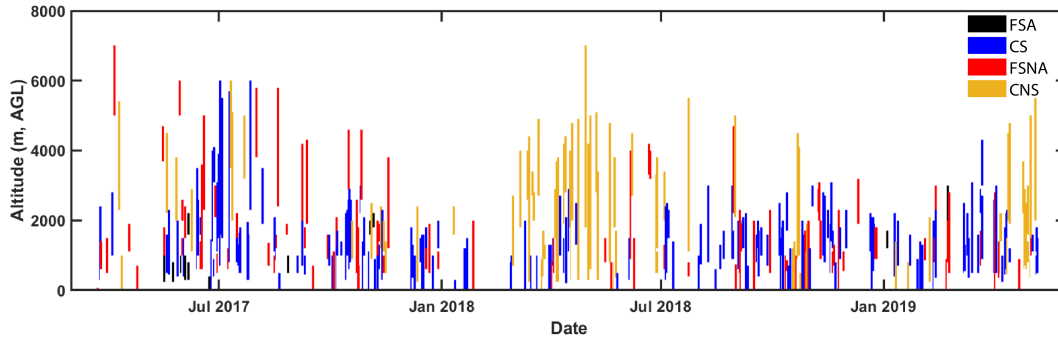
Figure 6.2: Overview of the altitude end extent of the aerosol layers above Haifa, Israel, between March 2017 and May 2019, as observed with a Polly$^{\mathrm{XT}}$ lidar. Each layer is colored by the aerosol component that is present with at least 50% of the total relative volume of the layer. Black, blue, red and yellow bars correspond to FSA, CS, FSNA and CNS, respectively.

is contributing to the mixture at least with 50% in terms of relative volume contribution. This rather large threshold is not able to reveal the layer composition, but nevertheless it provides a first impression of the overall aerosol load composition. Most of the lower aerosol layers (below 2 km) have large contributions of CS particles. CS particles can be attributed mainly to marine particles, especially after taking into account the location of Haifa (coastal city) as well as the altitude of the layers (lowest 2 km, MBL). Above 2 km, most of the lofted layers are composed of FSNA and CNS aerosol. These aerosol components can be attributed to particles of anthropogenic origin and desert dust, respectively. Given the altitude of these layers, we can assume that these aerosol types were both subject to long-range transport. Apart from an overview of the aerosol load above Haifa, Fig. 6.2 provides a quick overview of the seasonality of the observed aerosol components. Overall, the CS aerosol component is exhibiting high contributions to the aerosol mixtures observed for all seasons (2017–2019), however, during the spring season we observe high contributions from the CNS component. The reason for that, along with a more detailed analysis of the seasonality of the different aerosol components, is discussed in Sec. 6.3.

To fully characterise the mixing state of the aerosol layers observed above Haifa, the composition of all the aerosol layers needs to be carefully examined. Figure 6.3 shows the relative volume contribution of the different aerosol components of each considered aerosol layer. From Fig. 6.3 one can see that most layers are composed of two or more contributing aerosol components (277 layers) and that the existence of pure aerosol layers is rather rare. To examine the aerosol composition of the layers, we defined a contribution greater than 80% (in terms of relative volume) as "clear component predominance" and a contribution greater than 95% as "pure layers". Out of the 474 examined aerosol layers, in 83 layers
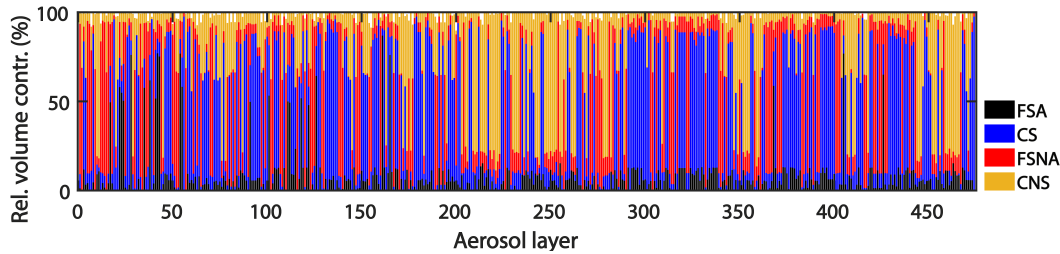
Figure 6.3: Overview of the aerosol layer composition in terms of relative volume contribution per aerosol component, above Haifa, Israel, as retrieved with the OEM-based typing scheme. The layers (474 in total) are presented in chronological order (March 2017 until May 2019).

(17.5%) the aerosol component with clear predominance was the CS, in 67 layers (14.1%) the FSNA component, in 46 layers (9.7%) the CNS component and only 1 layer was found to be clearly dominated by the FSA component (0.2%). In addition, pure layers were observed only 8 times (1.7%) in this period (23.03.2017–08.05.2019) and consisted of the CS component.

To retrieve information regarding the dominant mixture per aerosol layer, it is sufficient to examine the two dominant aerosol components present in the aerosol layer. Not surprisingly, it was found that the most commonly occurring aerosol mixture is dominated by CS particles (in 229 layers, mixed with either FSNA, FSA or CNS), followed by FSNA, CNS and FSA dominated aerosol mixtures (131, 90 and 24 layers, respectively). Regardless of the dominance of one of the two aerosol components present in the mixture, in total 120 aerosol mixtures were characterized as CS/FSA mixture, 85 as CS/FSNA, 114 as CS/CNS, 18 as FSNA/FSA, 115 as FSNA/CNS and 22 as CNS/FSA aerosol mixture.

The most commonly occurring aerosol mixture contains the CS and FSA aerosol components. Aerosol mixtures of CS/CNS and FSNA/CNS are also commonly encountered. All the aforementioned mixtures are expected at Haifa given its location and the governing circulation patters (Sec. 6.1). The least occurring aerosol mixture was found to be the mixture of the CNS and FSA components.

6.3 Seasonal variability

In this section, the seasonal variability of the different aerosol components found in the aerosol layers above Haifa is examined. In total 64, 143, 122 and 145 layers were analysed for winter, spring, summer and autumn months, respectively. For each season the frequency of occurrence of each layer's aerosol component relative volume contribution is studied and presented. By evaluating the times during which a specific range of relative volume contri-

bution of each aerosol component occurs, the seasonal aerosol distribution can be evaluated. To further support the typing scheme and identify the aerosol sources, the temporally and vertically resolved air-mass source attribution TRACE (Radenz et al., 2021b) was used.

TRACE combines backward trajectories or particle positions from a Lagrangian Particle Dispersion Model (LPDM) with a simplified land cover classification (MODIS land cover (Friedl et al., 2002) and manually defined areas) to obtain a continuous and vertically resolved estimate of an air-mass source above a certain location. Meteorological data is obtained from the Global Forecast System (GFS) analysis at a horizontal resolution of 1° (NCEP, 2000). In this study, the LPDM approach was used with the most recent version of FLEXPART (Stohl et al., 2005; Pisso et al., 2019). 500 particles are traced with their corresponding positions being stored every 3 hours. These simulations run for every 3 hours with height steps of 500 m for the period of interest (in this case, for the periods of the layer observations). The reception height (i.e., the proximity of an air mass to the surface) is set to 2 km, which is a widely applicable reception height threshold (Val Martin et al., 2018).
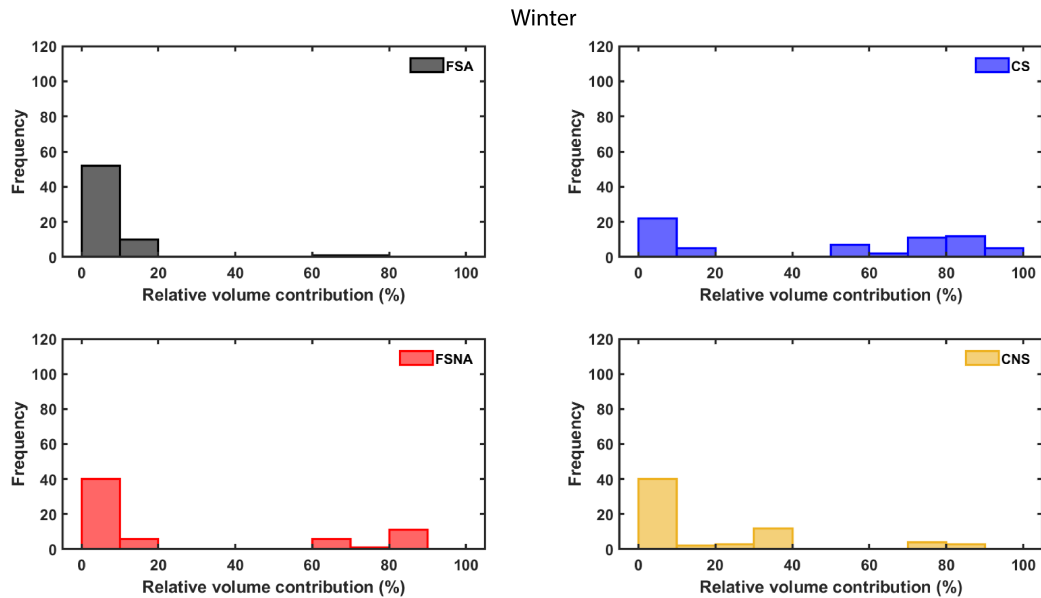
Figure 6.4: Frequency of occurrence (absolute) of the different aerosol components, in terms of relative volume contribution, during winter months (December, January and February 2017–2019).

Winter

The frequency of occurrence of the different aerosol components in terms of relative volume during winter months is depicted in Fig. 6.4. The aerosol component with the largest relative
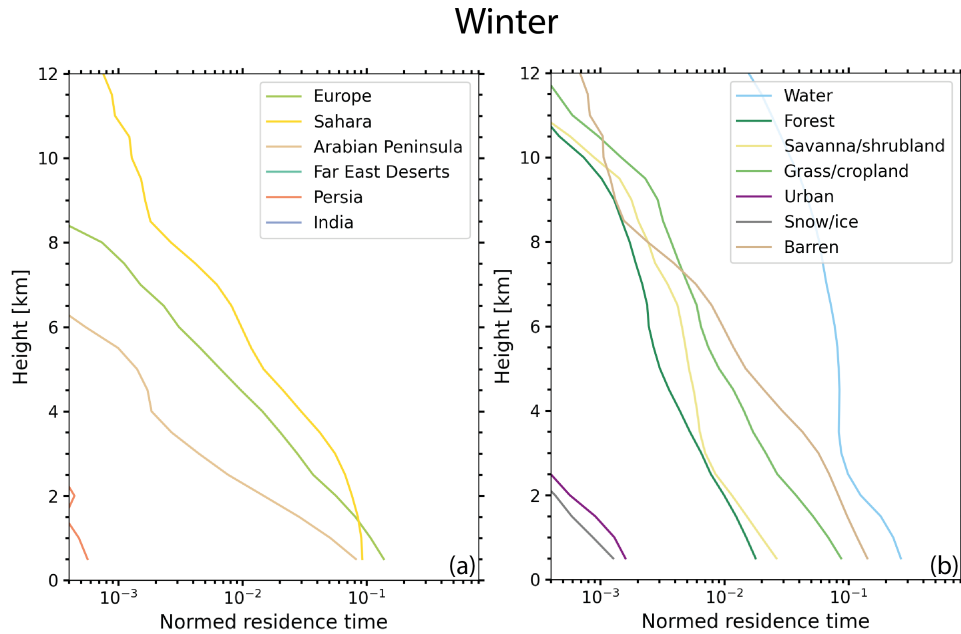
Figure 6.5: Air-mass source estimate based on FLEXPART particle positions for the Haifa station in December, January and February 2017–2019 (winter season). The named geographical areas (a) and the land surface classification (b) are only for the periods with available lidar data (sub-sampled). The reception height threshold is 2 km.

volume contribution to the aerosol layers was the CS component, as indicated by the high frequency of occurrence of the rather large relative volume contributions (e.g., above > 50%). The second aerosol component with large contributions (in terms of relative volume) to the aerosol layers was the FSNA one, followed by the CNS component. The aerosol component least present (relative volume contribution < 20%) in the aerosol layers was FSA. On average, aerosol mixtures during the winter months of 2017–2019 were composed of $47 \pm 21\%$ CS, $26 \pm 20\%$ FSNA, $19 \pm 20\%$ CNS and $7 \pm 15\%$ FSA.

Profiles of the air-mass sources for the aerosol layers observed above Haifa are shown in Fig. 6.5, with a reception height threshold of 2 km. In terms of geographical areas (Fig. 6.5a), below 2 km (where most layers were located) the most dominant air-mass sources were Europe and the Sahara, followed closely by the Arabian Peninsula. Air masses from Persia were less frequent and only up to 2 km. Above 2 km and up to 12 km, the most predominant source is the Saharan desert. No air masses from the Far East deserts (manually defined area containing the Gobi and Taklamakan deserts) and India fulfilled the 2 km reception height criterion. Most dominant land surface categories (Fig. 6.5b) were water, barren and grass/cropland, with water having significantly higher normed residence time (i.e., the time that an air mass spends above a specific land category) compared to all the other land surface categories (especially above 4 km). The residence time of all other categories

decreased monotonically with height.

Combining the aerosol information obtained from the OEM-based typing scheme and the air-mass sources from TRACE, it can be concluded that the CS particles found in aerosol mixtures above Haifa are most likely marine aerosol, primarily composed of water-soluble, coarse sea-salt particles. The air masses travelling from the Sahara and Europe towards Haifa spent significant time above water masses (Mediterranean Sea), thus allowing the uptake of sea-salt particles. The FSNA aerosol component can be attributed to European and Arabian sources and it is most likely composed of low or moderately absorbing aerosol, typically found in urban background and in general aerosol of anthropogenic origin. CNS particles observed are clearly desert dust particles, originating mainly from the Saharan region (long-range transport) but also from the Arabian Peninsula (locally produced or transported over short distances). Finally, the small contribution of FSA aerosol can be attributed to particles emitted from direct combustion processes, either produced regionally (Arabian Peninsula) or observed above Haifa after long-range transport events (e.g., from Europe).
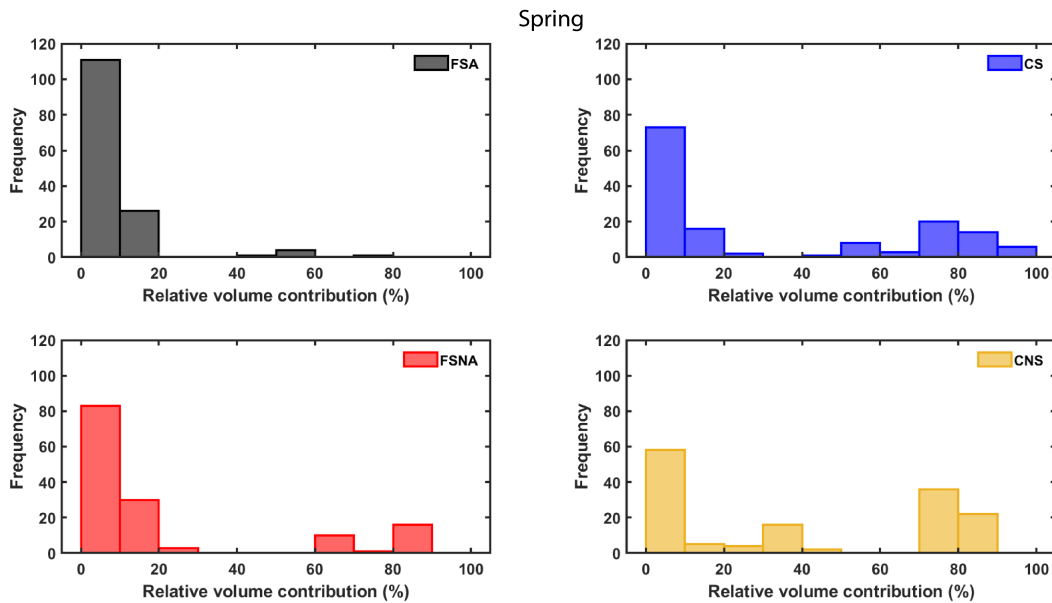
Figure 6.6: Frequency of occurrence (absolute) of the different aerosol components, in terms of relative volume contribution, during spring months (March, April and May 2017–2019).

Spring

Similarly to Fig. 6.4, Fig. 6.6 shows the the frequency of occurrence of the different aerosol components in the aerosol layers observed during the spring months between 2017 and

2019. The dominating aerosol component for the aerosol layers observed in the spring season was the CNS component. As Fig. 6.6 indicates, in many aerosol layers the CNS component contribution was above 70% (in terms of relative volume). High occurrence of large contributions (relative volume above 60%) is also observed for the CS component. The aerosol component least present (mostly with relative volume contributions lower than 20%) was FSA. On average, aerosol mixtures are composed of $40 \pm 21\%$ CNS, $31 \pm 19\%$ CS, $21 \pm 18\%$ FSNA and $7 \pm 14\%$ FSA.

The height profiles of the air-mass sources for the Haifa station are shown in Fig. 6.7, again using a reception height threshold of 2 km. In the lowest 1 km, the most dominant air-mass sources in terms of geographical area (6.7a) were Europe and the Sahara, followed closely by the Arabian Peninsula. Air masses from Persia were less frequent and were observed up to 4.5 km. Above 1 km, the Saharan desert is the predominant source. During the spring months, air masses originating over water were the most frequent ones below 2 km and then again above 3.5 km. The second most frequent land cover was barren, reaching a peak between 2.5 and 3 km. Other land surfaces with significant air-mass residence times were grass/cropland, savanna/shrubland and forest. With the exception of water, the residence time of all other land categories decreased monotonically with height.

The predominant CNS component contribution to the aerosol layers during the spring season (2017–2019) can be attributed to desert dust particles transported to the measurement site mainly from the Saharan desert and the Arabian Peninsula. Air masses from Europe most likely carried FSNA and FSA particles, associated with smoke, urban background and pollution particles. The latter two are expected to have local sources as well, given their low frequency of occurrence of high relative volume contributions. The overall high contribution of CS particles, usually associated with sea-salt particles, can be explained by the vicinity of the city to the eastern part of the Mediterranean basin and prevailing sea-breeze conditions at the measurement site.

Summer

The frequency of occurrence of the different aerosol components during the summer months (June, July and August 2017–2018) is shown in Fig. 6.8. The aerosol layers were dominated by the CS component, as indicated by the high occurrence of large relative volume contributions (above 70%), followed by the FSNA and the CNS component. The FSA component was the least present in the mixtures observed. Similar aerosol composition was also observed during winter months (Fig. 6.4). On average, the observed aerosol mixtures consisted of $41 \pm 21\%$ CS, $29 \pm 20\%$ FSNA, $20 \pm 20\%$ CNS and $9 \pm 15\%$ FSA particles.

The height profiles of the air-mass sources for the Haifa station are shown in Fig. 6.9 and reveal that in the lowest 4 km the most dominant air-mass source in terms of geographical
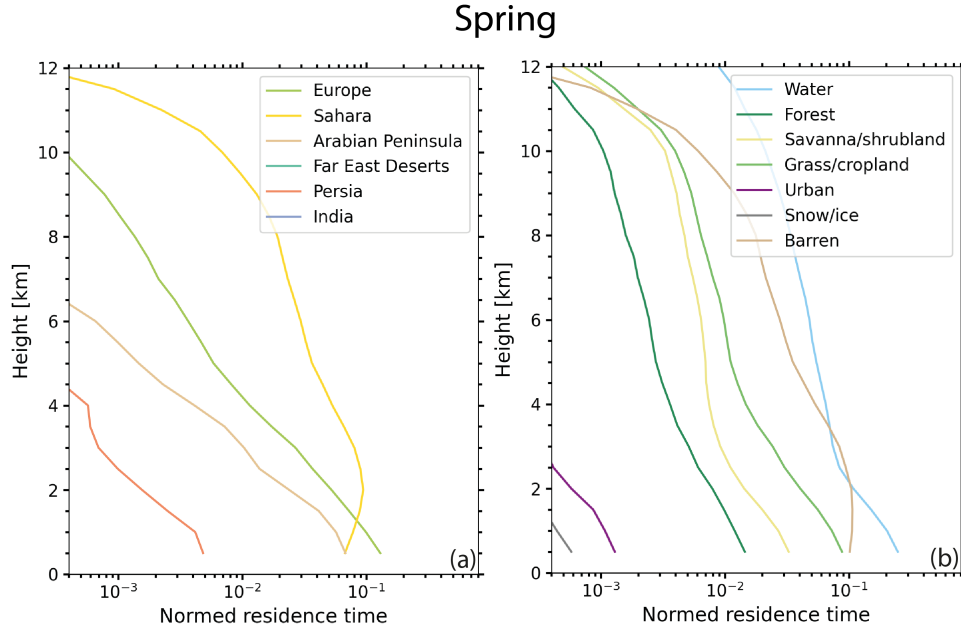
Spring

Figure 6.7: Same as Fig. 6.5, but for March, April and May (spring season) 2017–2019.

area was Europe and in terms of land cover it was water. The Saharan desert and Arabian Peninsula were also significant sources of air masses below 4 km, with the Sahara being the most prominent source between 6 and 9 km. Below 6 km, air-mass sources included also Persia. Apart from water, dominant land surface categories included barren (peaking between 3.5 and 6 km), grass/cropland, savanna/shrubland and forest. Land cover categorized as urban was only apparent below 3.5 km.

The air-mass source attribution correlates well with the aerosol typing results. The dominance of the CS component can be associated most likely with sea-salt particles picked up, while air masses were travelling towards Haifa crossing over the Mediterranean Sea. FSNA particles originated most likely from Europe, along with local contributions, while air masses originating from the Sahara and the Arabian Peninsula justify the relatively high contribution of CNS particles in the aerosol mixtures observed.

Autumn

Fig. 6.10 depicts the frequency of occurrence of the different aerosol components during autumn months (September, October and November 2017–2018). The aerosol component with the highest occurrence of large relative volume contributions (above 60%) was the FSNA one, followed closely by CS. Similarly to all previously discussed seasons, the FSA component was the least present in the mixtures observed. The aerosol mixtures observed in the aerosol layers during autumn consisted of $44 \pm 21\%$ CS, $28 \pm 20\%$ FSNA, $19 \pm 20\%$
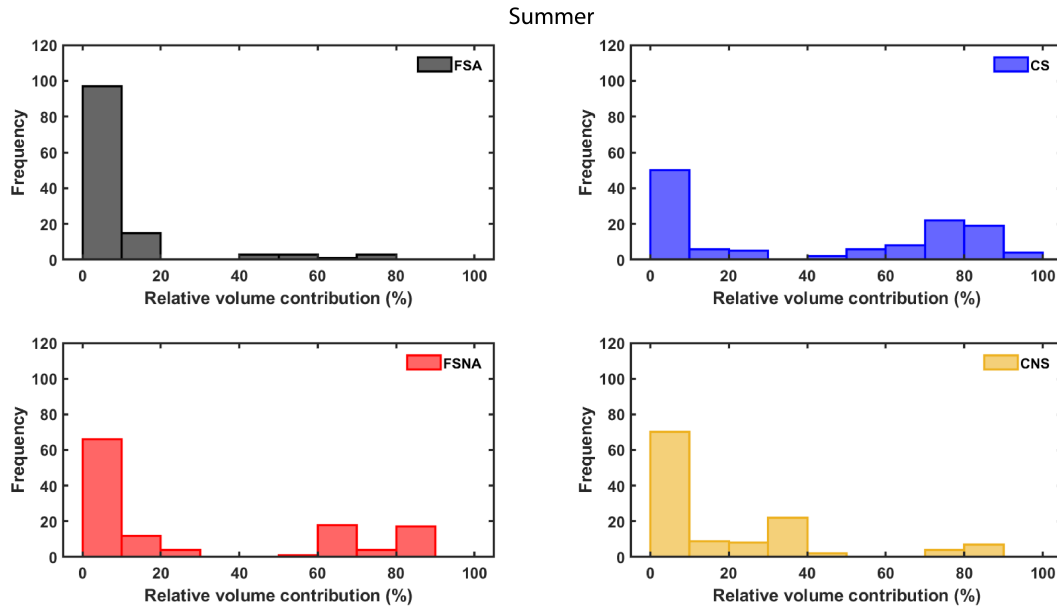
Figure 6.8: Frequency of occurrence (absolute) of the different aerosol components, in terms of relative volume contribution, during summer months (June, July and August 2017–2018).

CNS and $8 \pm 15\%$ FSA.

The most predominant air-mass sources for autumn (Fig. 6.11) were the Arabian Peninsula and Europe below 2 km, while above that altitude Sahara dominated in terms of geographical area. In terms of land surface, air masses spent significant time above areas characterized as water, barren, forest, savanna/shrubland and grass/cropland.

The strong presence of FSNA can be correlated with the predominance of local and transported pollution and other anthropogenic aerosol from the Arabian Peninsula and Europe, respectively (Fig. 6.11a). The CNS component of the aerosol mixtures can be attributed to desert dust particles originating mainly from the Arabian Peninsula for altitudes below 2 km, and mainly from the Saharan desert for altitudes above 2 km.

Seasonal statistics

As a means to summarize the findings of Sec. 6.3, the statistics of the seasonal relative volume contribution per aerosol component are presented in Fig. 6.12 in the form of boxplots. In general, no major differences were observed between the seasons. The aerosol component with the highest contribution (for all seasons except spring) to the aerosol mixtures observed was the CS component, followed by the FSNA, the CNS and finally the FSA component, which exhibited the lowest contributions. The high contribution of the CS aerosol component can be attributed to the close vicinity of Haifa to the Mediterranean Sea as discussed

Summary
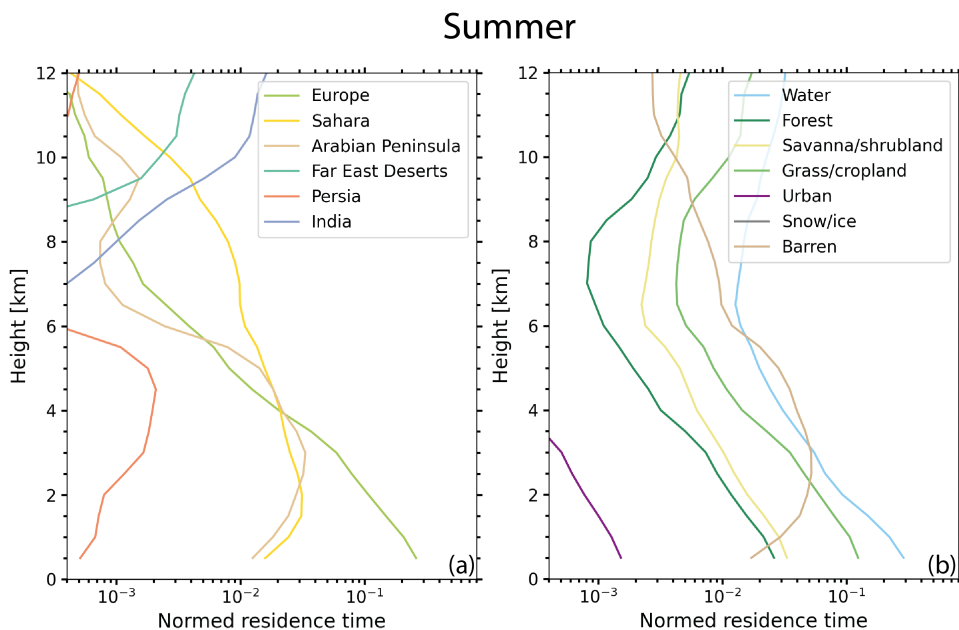
Figure 6.9: Same as Fig. 6.5, but for June, July and August (summer season) 2017–2018.

above, as well as the altitude of the investigated layers. In addition, the presence of the CS aerosol component is further supported by the air-mass source attribution analysis. These CS particles can be correlated either with sea-salt particles or with other aerosol particle types (e.g., continental aerosol) that due to hygroscopic growth increased in size (Laskina et al., 2015) and therefore were identified as large spherical particles by the OEM typing scheme. The contribution of FSNA can be partially explained by the fact that Haifa is surrounded by industries, including large petrochemical plants, an oil-fueled power station, a large cement factory and petroleum refineries among other smaller industries and workshops (Ganor et al., 1998). However, the sources of the FSNA aerosol are not only local, as FSNA aerosol is often subject to long-range transport (e.g., sulfate-related aerosol particles originating from Europe) (Luria et al., 1996). During spring, the aerosol component with the highest contribution to the observed aerosol mixtures was the CNS one, typically associated with desert dust particles. Desert dust particles were transported to Haifa by the prevailing synoptic patterns from the Sahara (Gkikas et al., 2016). In particular, thermal Saharan lows developing south of the Atlas Mountains and moving eastwards across the North African coast induce south-southwesterly winds, favoring thus the transport of dust particles towards the Mediterranean (basin) and Europe (Moulin et al., 1997; Gkikas et al., 2015). The central and eastern parts of the Mediterranean Sea are commonly affected by dust transports under the aforementioned synoptic conditions as shown in previous studies (Barnaba and Gobbi, 2004; Papadimas et al., 2008; Gkikas et al., 2013; Floutsi et al., 2016).
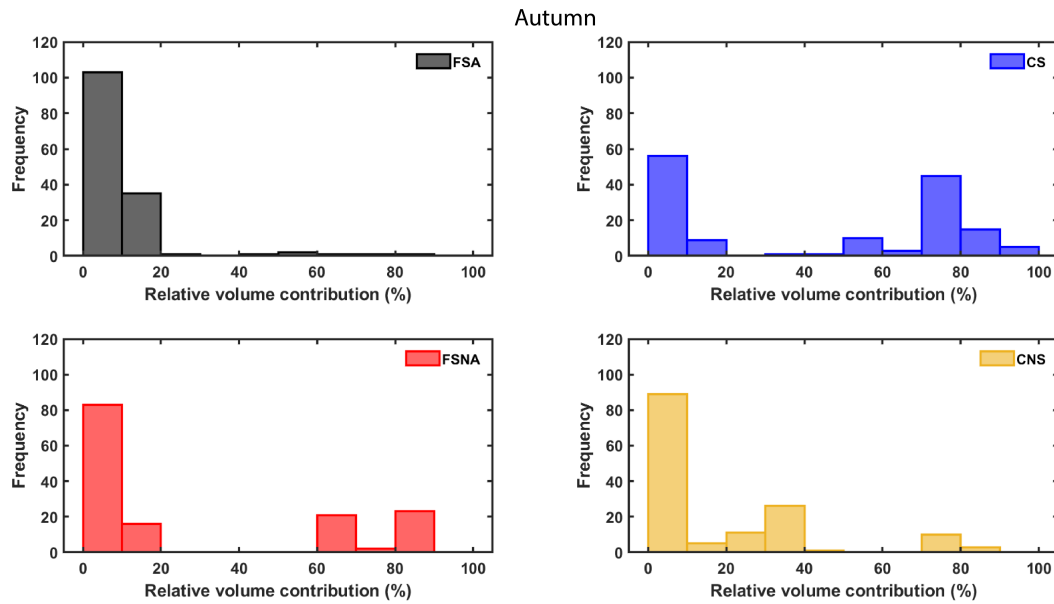
Figure 6.10: Frequency of occurrence (absolute) of the different aerosol components, in terms of relative volume contribution, during autumn months (September, October and November 2017–2018).
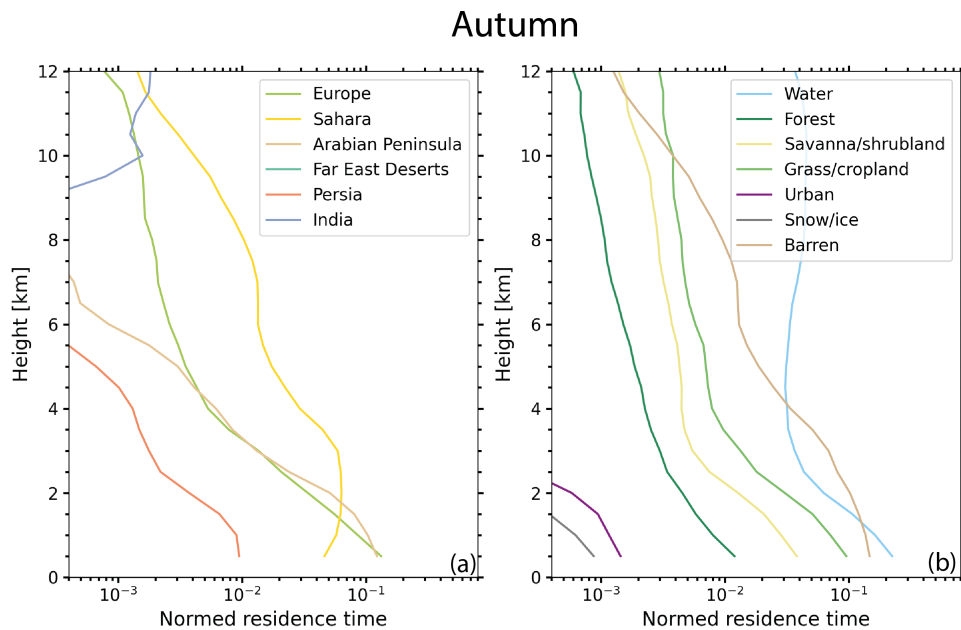
Figure 6.11: Same as Fig. 6.5, but for September, October and November (autumn season) 2017–2018.
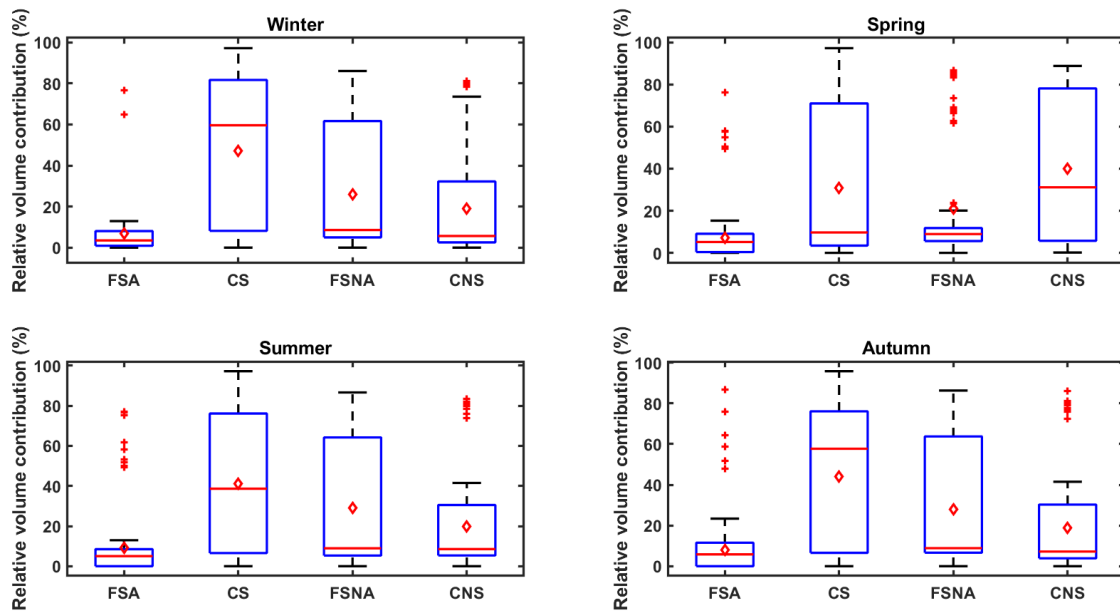
Figure 6.12: Seasonal statistics of the relative volume contribution of the different aerosol components. Minimum and maximum values are indicated by the lower and upper whisker, median and mean values by the red lines and rhombuses, respectively. The lower part of each box indicates the 25% percentile and the upper part the 75% percentile. Red crosses show the outliers.

Chapter 7

Summary, conclusions and outlook

Within the scope of this work, an aerosol typing methodology was developed and applied to several case studies as well as to a long-term lidar dataset to demonstrate its strengths and limitations. The novel aerosol typing scheme was developed based on the optimal estimation method (OEM) and allows the identification of up to four different aerosol components present in an aerosol mixture as well as the quantification of their contribution to the aerosol mixture in terms of relative volume. The four aerosol components considered in this typing scheme represent the most commonly observed aerosol particles in nature and are assumed to be physically separated from each other and, therefore, can create external mixtures. Two components represent fine-mode particles, absorbing (FSA) and less absorbing (FSNA), and the remaining two aerosol components represent coarse-mode particles, spherical (CS) and non-spherical (CNS). These components can adequately represent the most frequently observed aerosol types in the atmosphere: combustion- and pollution-related aerosol, sea salt and desert dust, respectively.

The "optical footprints" of the different aerosol types allow (in most cases) their correct identification and separation. The optical parameters used exhibit a high discriminatory power (Ch. 2) and are the lidar ratio and the particle linear depolarization ratio at two distinctive wavelengths (355 and 532 nm), the backscatter-related color ratio (for the wavelength pair of 532/1064 nm) and the extinction-related Ångström exponent (for the wavelength pair of 355/532 nm).

By using the OEM-based typing scheme, the mixing ratio of the different aerosol components (expressed in terms of relative volume) can be determined. From a mathematical point of view, the mixing ratio is the most-likely state that can reproduce the measurement vector, which contains the optical parameters measured by the lidar. Once retrieved and statistically significant within the 95% confidence interval, the mixing ratio of the four different aerosol components can be used for component-separated calculations of other quantities, e.g., the volume and number concentration (Sec. 4.2).

Overall, the algorithm performed well and as expected. The information content of the measurement vector (expressed by the degrees of freedom, see Ch. 4.1.6) was one of the driving forces behind the different available forward-model configurations (i.e., different combinations of intensive properties, see Tab. 4.4). In principle, each retrieval mode is able to produce statistically significant results and there is no evidence suggesting that one mode is superior to another systematically. In retrospect, it can be concluded that the retrieval modes that include two input parameters (i.e., retrieval modes 1 and 2) tend to converge faster (but not necessarily more reliably) than those that include more input parameters. The conclusion is coherent, since the more input parameters the more challenging the task of the forward model to simultaneously reproduce all the input parameters within the measurement error.

The retrieved contributions were sometimes accompanied by rather high values of the respective retrieval error. The *a posteriori* uncertainty (i.e., the covariance matrix of the optimally estimated state vector, see Sec. 4.1.6) is directly linked to the *a priori* uncertainty, meaning that the larger the *a priori* uncertainty the larger the retrieved *a posteriori* uncertainty. The retrieval is, therefore, strongly driven by the *a priori* uncertainty, which essentially constrains the retrieval solution space. Hence, the continuation of the lidar measurements of the different aerosol types as well as the expansion of the experimental basis (Ch. 3) is of great importance for this typing scheme.

The functionality of the typing scheme was demonstrated by its application to case studies of known aerosol load (see Sec. 4.4) as well as to cases of non-characterized aerosol load (see Ch. 5). The algorithm was also applied to a long-term dataset to provide a seasonal characterization of the aerosol load (see Ch. 6). It was shown that the OEM is an effective methodology that can be also applied for aerosol typing purposes, which provided a satisfactory answer to the first research question of this thesis (see Sec. 2.6). As a refresher the research questions are listed here again.

1. Can an optimal-estimation-based methodology be applied effectively for aerosol typing purposes?

2. Is it possible to describe an aerosol mixture observed by means of basic aerosol components?

3. Can we retrieve the contribution of each aerosol component, in terms of relative volume, leading thus to a quantitative characterization of the aerosol mixture?

4. Is the proposed methodology sufficient for ground-based validation purposes and further support of EarthCARE's aerosol classification scheme (HETEAC)?

In addition, the application of the OEM-based algorithm to the aforementioned case studies

and to a long-term dataset showed that, in most of the cases, the aerosol mixture observed could be adequately described as a mixture of the four predefined aerosol components (Sec. 4.1.3) and that their contribution (in terms of relative volume) could be determined, answering thus the second and third research questions. The ability of the algorithm to quantitatively characterize the aerosol mixture was well demonstrated for the A-LIFE case (Sec. 5.1). The volume and number size distributions that were obtained based on the OEM output (relative volume contributions of each aerosol component) compared well with the respective size distributions that were obtained via in-situ aircraft measurements, especially when one takes into account that the OEM-based typing scheme is based on parameterization and cannot be used for detailed microphysical retrievals. The aforementioned comparison also proved that the microphysical assumptions that are currently in use, even though very generic, can sufficiently capture the aerosol load observed.

Furthermore, the successful application of the OEM-based aerosol typing scheme to various case studies under various aerosol-load scenarios answers the last research question. The algorithm is flexible enough allowing, thus, its application to lidar systems with different configurations (e.g., single wavelength or multiwavelength, Sec. 4.1.5), a feature that is desired for ground-based validation purposes (e.g., when different lidar systems are deployed in different locations). To ensure compatibility with HETEAC, EarthCARE's aerosol classification scheme, the microphysical and optical properties of the predefined aerosol components were kept almost identical (only a few minor changes in the optical properties were applied, as described in Sec. 4.1.3). At its current status (four basic aerosol components), the OEM-based aerosol typing methodology can provide sufficient support to HETEAC as it can be utilized to compare the aerosol mixtures determined by EarthCARE with those retrieved from ground-based lidar systems. In the following, an outlook along with some suggestions that could be carried out in a follow-up project are provided.

Outlook

The current version of the OEM-based typing scheme is planned to be implemented in the processing chain of PollyNET (Baars et al., 2016) soon, resulting, thus, in near-real-time information regarding the aerosol components of the aerosol mixture observed. In addition to that, all the historical data acquired within the framework of previous measurement campaigns (e.g., CADEX) or routinely at the different PollyNET stations will also be processed with the typing scheme, yielding, thus, to the characterization of the aerosol load in terms of mixing ratio of the four aerosol components in addition to the already existing aerosol optical parameters. Such information will facilitate long-term studies focused on aerosol typing (i.e., for stations with long datasets such as the station in Leipzig, Germany, which has data for more than 14 years). The following propositions emerged during the

development and usage of the OEM-based typing scheme, but they were beyond the scope of this dissertation. Nevertheless, they will be briefly discussed to give an idea about the possibilities that the typing scheme allows.

The aerosol typing scheme can be refined and developed even further. For instance, it would be interesting to explore the possibility of adding more constraints and evaluate the benefits of it. Currently, there is only one constraint in place (introduced in Sec. 4.1.6), which ensures physically meaningful solutions. This constraint is very important for the quality of retrievals, however, there are other constraints that potentially could further improve the retrievals. Relative humidity (RH) is a useful quantity for the determination of the aerosol growth, as it plays an important role in the physical, chemical and optical properties of aerosol particles due to their increased water content. Therefore, the connection between RH and the observed optical properties needs to be further investigated. For instance, in the current version of the typing scheme, the CS aerosol component, which resembles marine aerosol particles, is assumed to be spherical, regardless of the actual relative humidity. Haarig et al. (2017) and later on Bohlmann et al. (2018) observed that under dry conditions (RH<45%) marine particles adopt a cubic-like shape and can produce much higher particle linear depolarization ratios. It becomes obvious that the typing scheme would benefit from an additional constraint based on routinely-performed RH measurements. The constraint could work as an *a priori* regulator, adjusting appropriately the *a priori* information of the different aerosol components based on the RH.

As already discussed in Sec. 5.3, the introduction of additional aerosol components in the typing scheme is possible, and it can increase the accuracy of the algorithm. To name a few, such aerosol components could represent aged biomass-burning aerosol, aerosol of volcanic origin and pollen. In addition, the existing aerosol components could be further refined, for instance the CS component that resembles the marine particles could include two modes, a coarse mode (sea salt) and an accumulation mode (non-sea-salt sulphates). Such potential refinements were not attempted here as they would not allow for direct comparisons with HETEAC, but their potential should be further investigated.

Appendix A

Experimental Basis

The experimental basis presented in Ch. 3, was focused on the lidar ratio and particle linear depolarization ratio at the wavelengths of 355 and 532 nm. However, the complete collection of ground-based observations includes more intensive parameters, i.e., the extinction- and backscatter-related Ångström exponents. Therefore, in addition to Figures 3.1 and 3.2, all the intensive optical parameters collected are presented here in all possible combinations. Furthermore, a statistical analysis of the extinction- and backscatter-related Ångström exponents of the four aerosol types is also provided (see Fig. A.20).
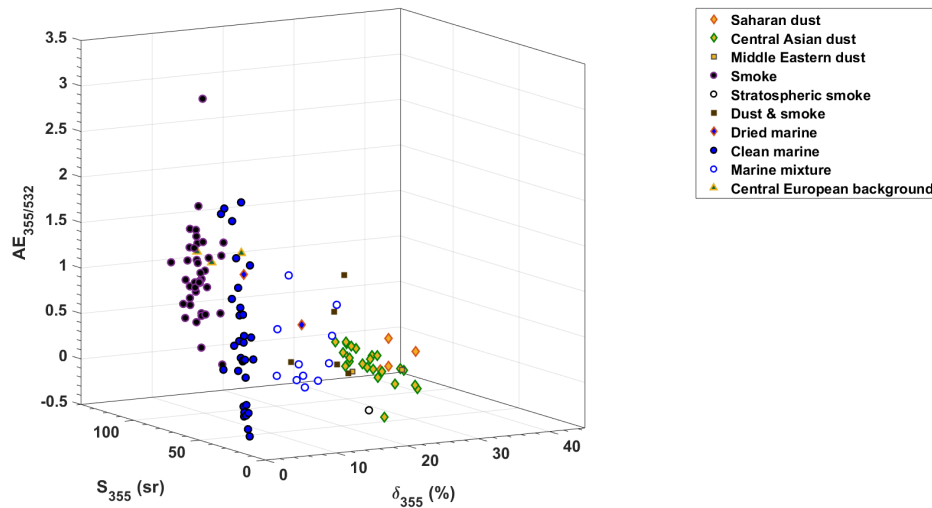
Figure A.1: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), both at 355 nm, and extinction-related Ångström exponent (AE).

It has been already demonstrated (Sec. 3.2, Fig. 3.3) that adding a third dimension (in this case the extinction- or backscatter-related Ångström exponent) to the 2D domain of lidar ratio and particle linear depolarization ratio at 355 nm allows the separation of
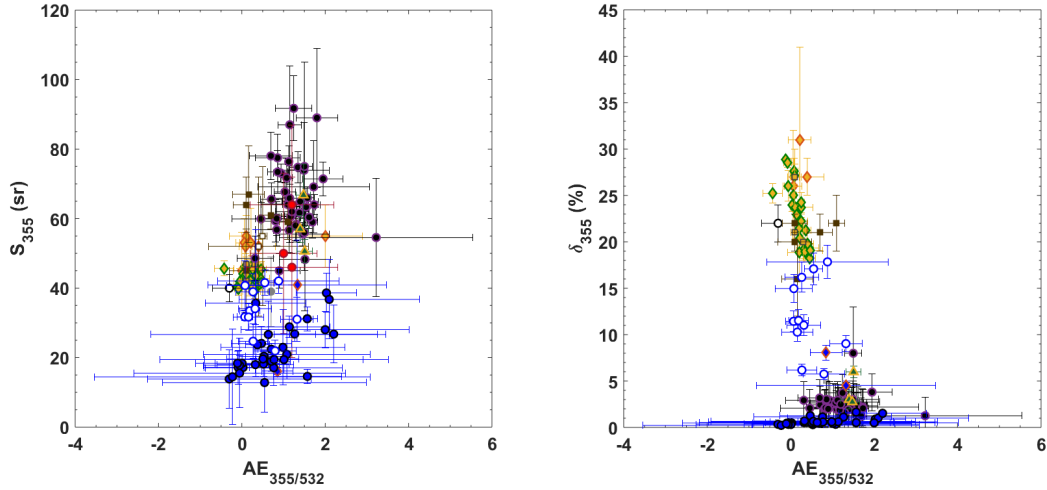
Figure A.2: 2D domains of the intensive optical parameters corresponding to Fig. A.1. Point markers are uniform throughout the thesis (for a complete list see Fig. 3.1).

aerosol mixtures that otherwise appear to have similar optical properties. By comparing the 2D spaces (Fig. A.2, A.4 and Fig. A.6) it becomes clear that when the backscatter-related Ångström exponent (both at 355/532 and 532/1064 nm) is combined with either the lidar ratio or the particle linear depolarization ratio it tends to create more "well-defined" aerosol clusters in comparison to the extinction-related Ångström exponent, and that is due to the high uncertainties of the later.
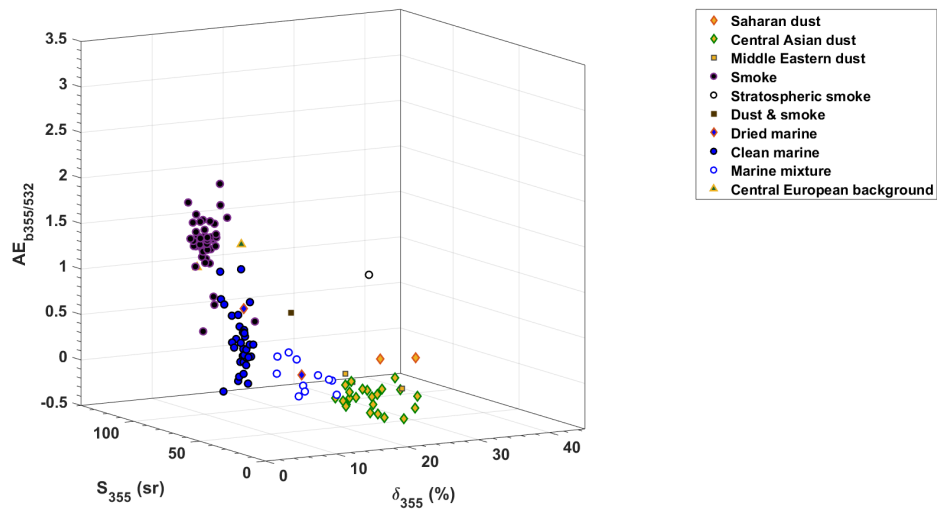
Figure A.3: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), both at 355 nm, and backscatter-related Ångström exponent at the wavelength pair 355/532 nm (AE).
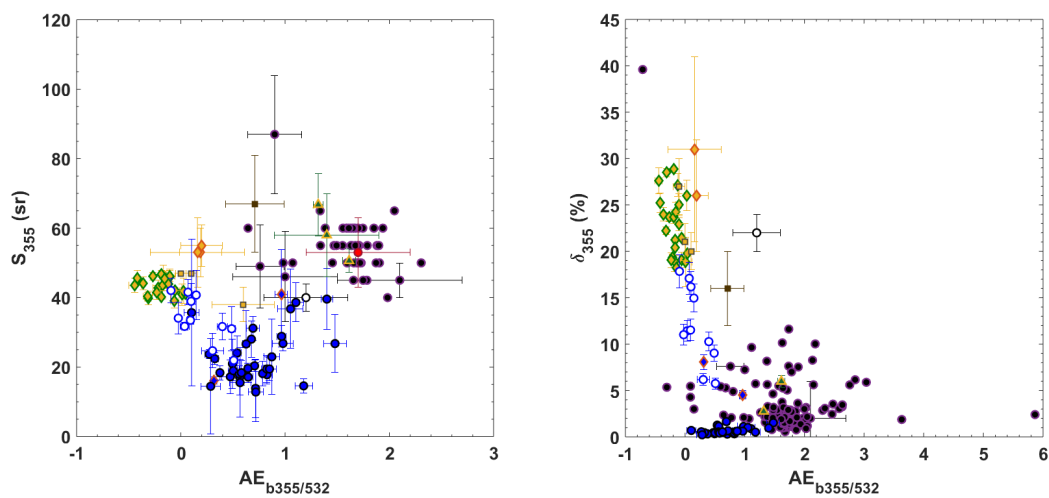
Figure A.4: Same as Fig. A.2, but corresponding to Fig. A.3.
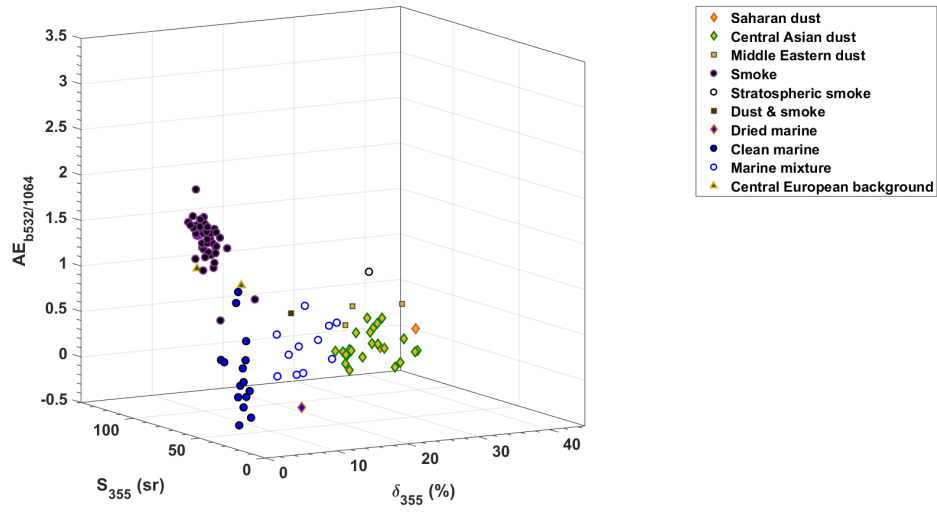
Figure A.5: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), both at 355 nm, and backscatter-related Ångström exponent at the wavelength pair 532/1064 nm (AE).
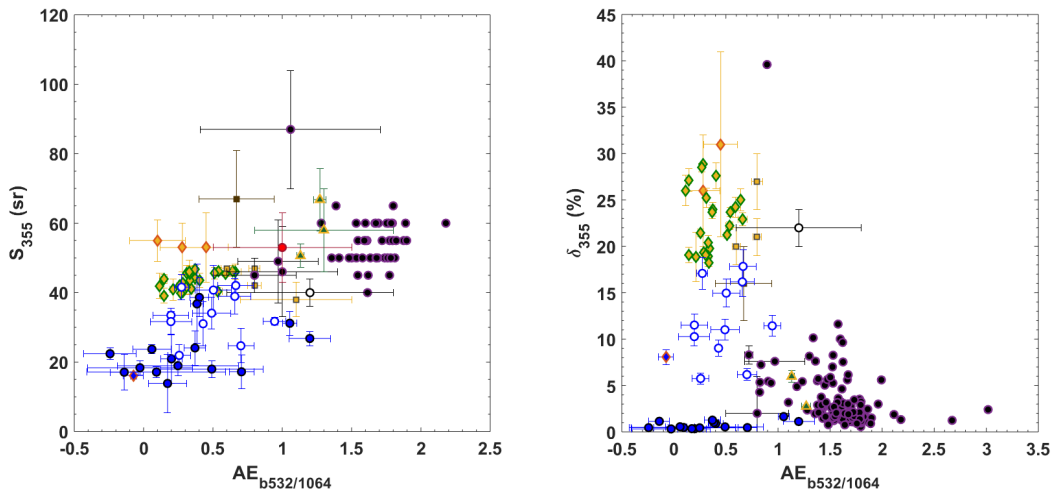
Figure A.6: Same as Fig. A.2, but corresponding to Fig. A.5.
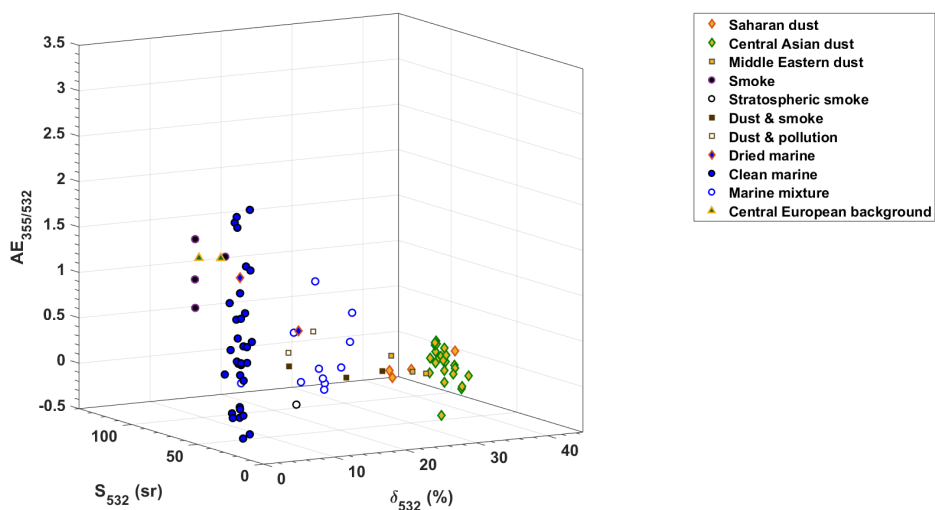
Figure A.7: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), both at 532 nm, and extinction-related Ångström exponent (AE).
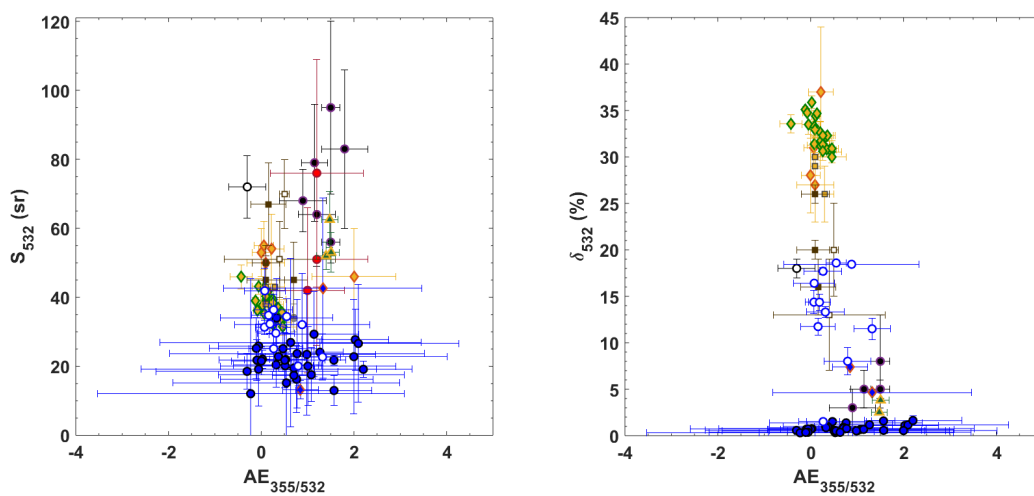
Figure A.8: Same as Fig. A.2, but corresponding to Fig. A.7.
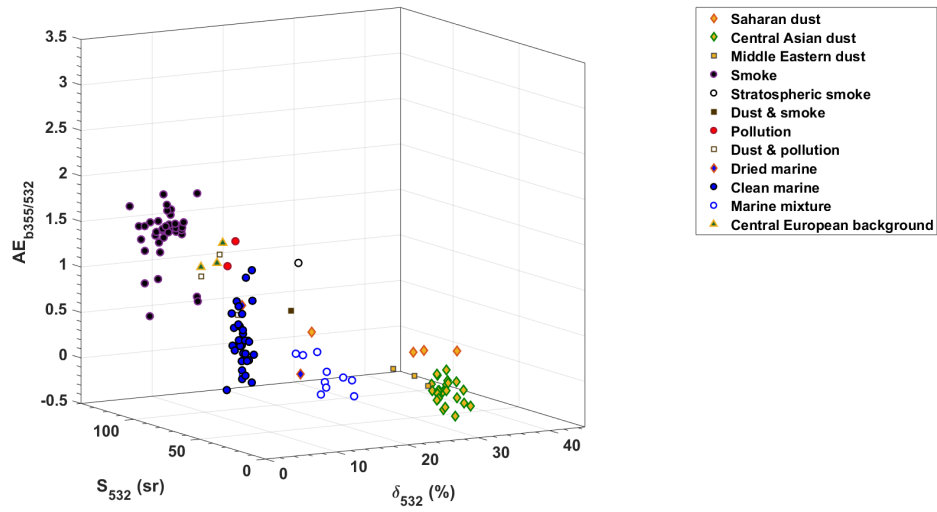
Figure A.9: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), both at 532 nm, and backscatter-related Ångström exponent at the wavelength pair 355/532 nm (AE).
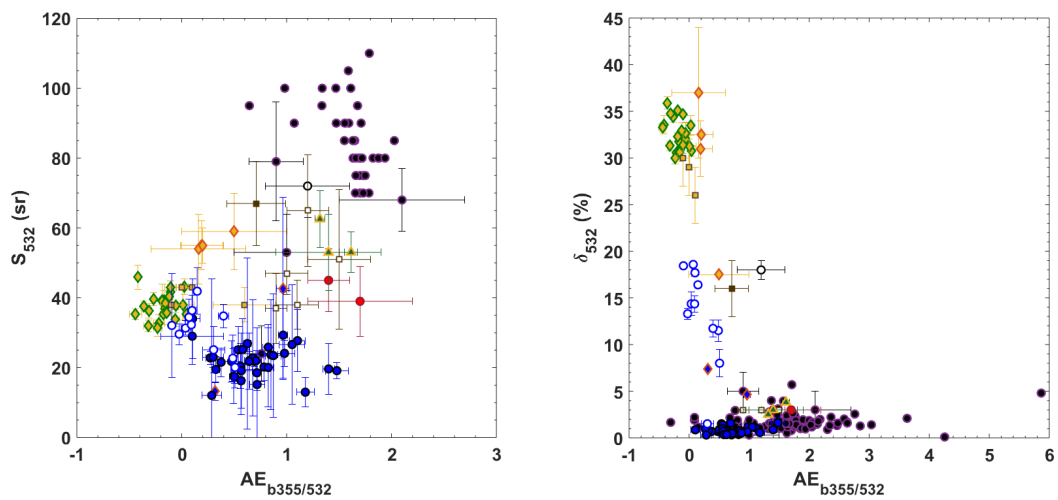
Figure A.10: Same as Fig. A.2, but corresponding to Fig. A.9.
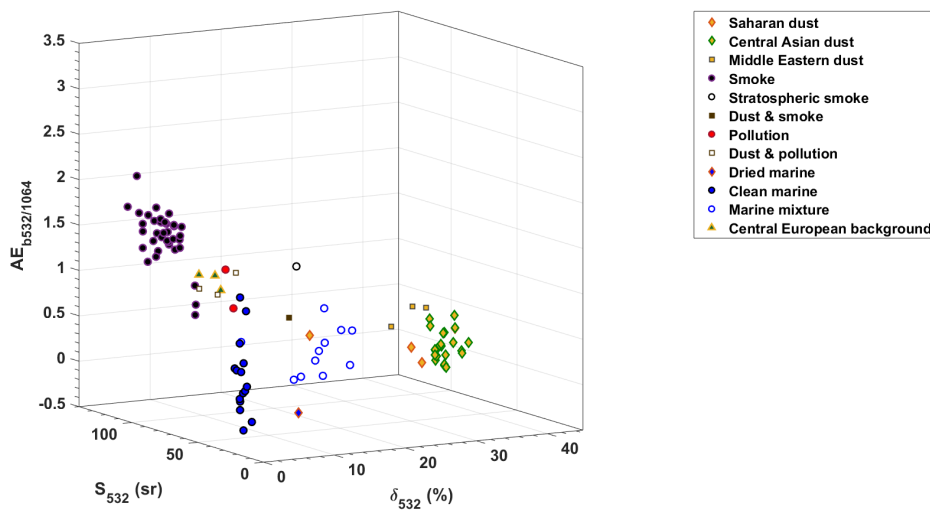
Figure A.11: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), both at 532 nm, and backscatter-related Ångström exponent at the wavelength pair 532/1064 nm (AE).
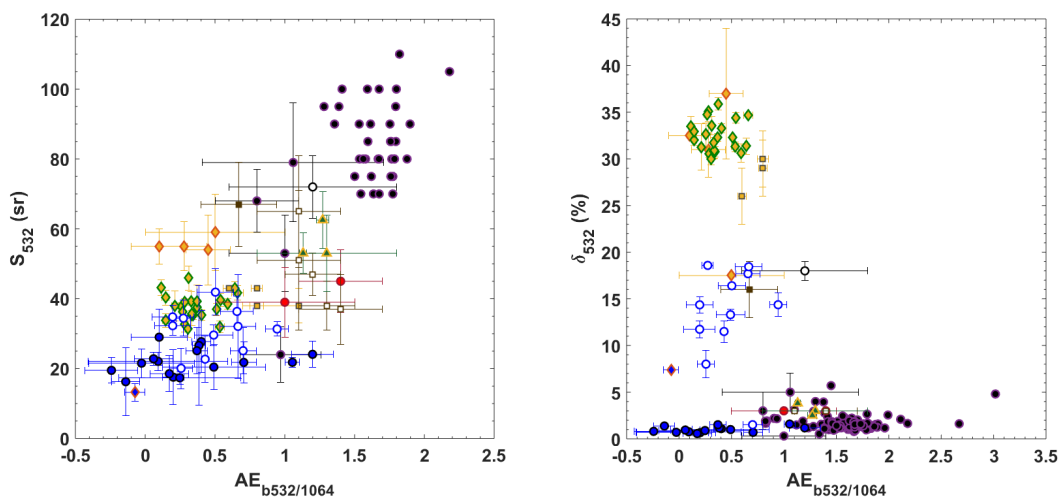
Figure A.12: Same as Fig. A.2, but corresponding to Fig. A.11.
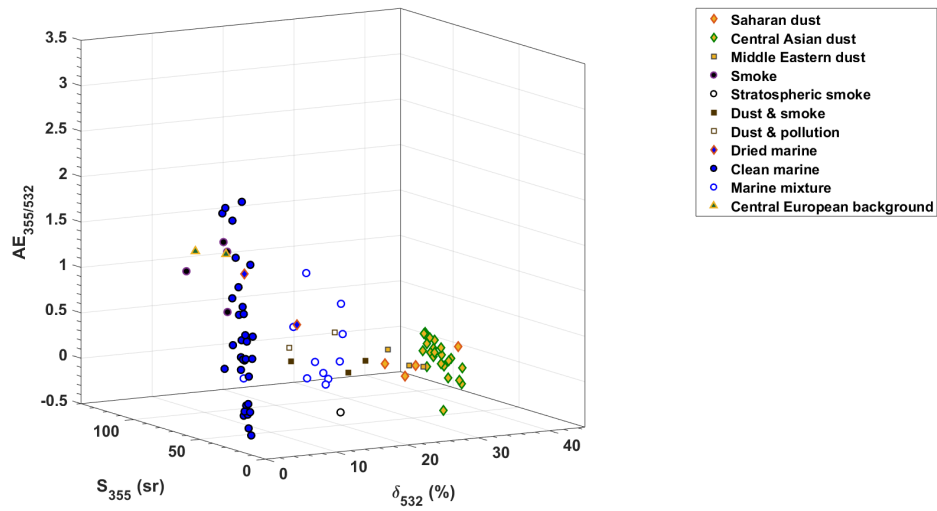
Figure A.13: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), at 355 and 532 nm, respectively, and extinction-related Ångström exponent (AE).
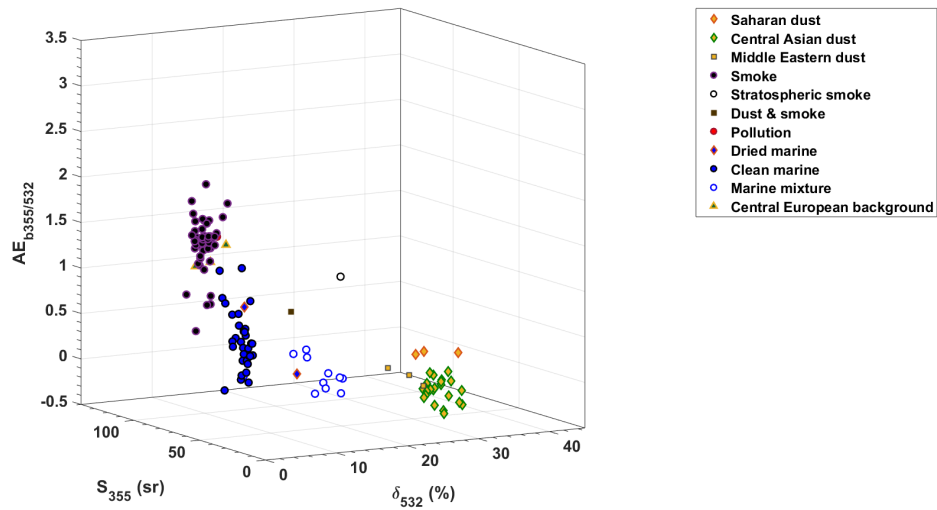
Figure A.14: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), at 355 and 532 nm, respectively, and backscatter-related Ångström exponent at the wavelength pair 355/532 nm (AE).
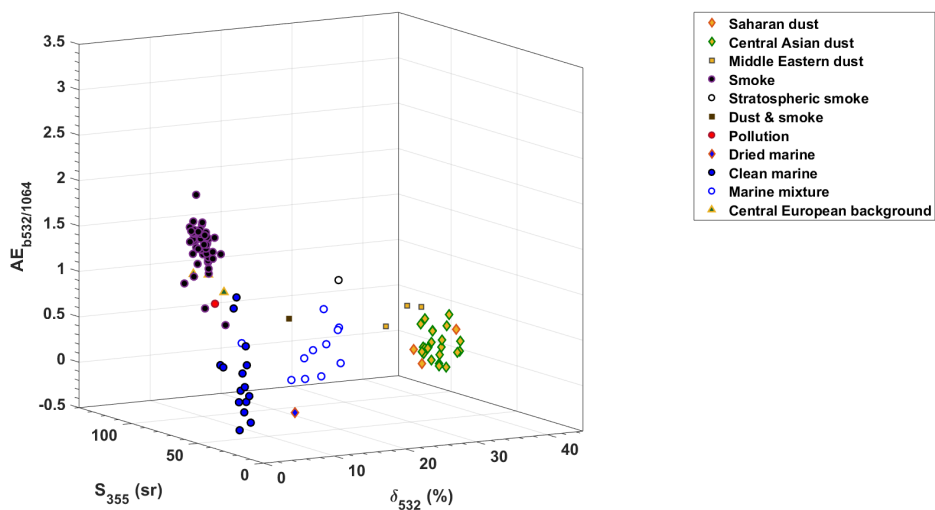
Figure A.15: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), at 355 and 532 nm, respectively, and backscatter-related Ångström exponent at the wavelength pair 532/1064 nm (AE).
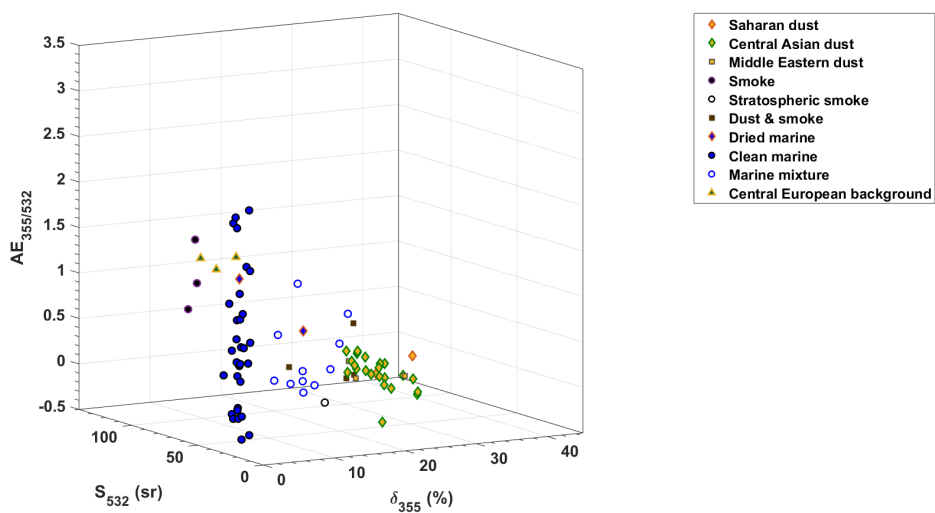
Figure A.16: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), at 532 and 355 nm, respectively, and extinction-related Ångström exponent (AE).
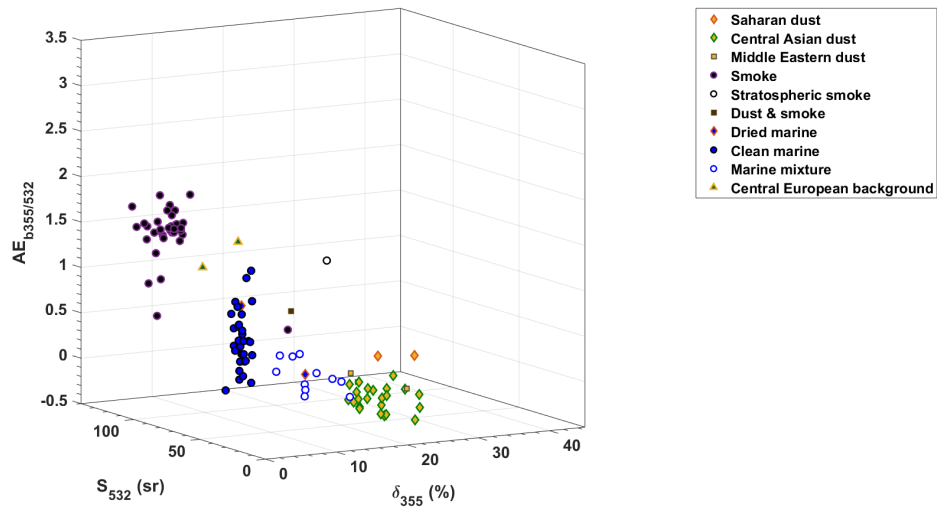
Figure A.17: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), at 532 and 355 nm, respectively, and backscatter-related Ångström exponent at the wavelength pair 355/532 nm (AE).
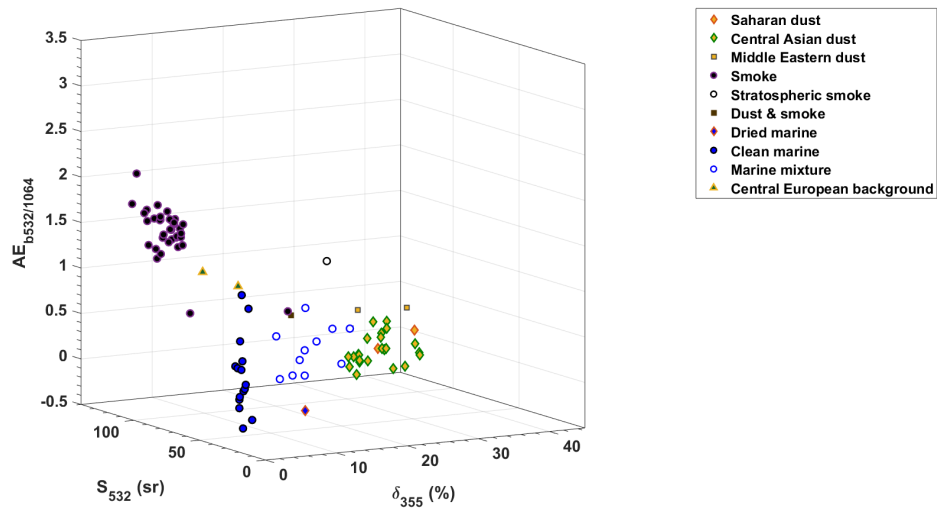
Figure A.18: 3D domain of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ), at 532 and 355 nm, respectively, and backscatter-related Ångström exponent at the wavelength pair 532/1064 nm (AE).
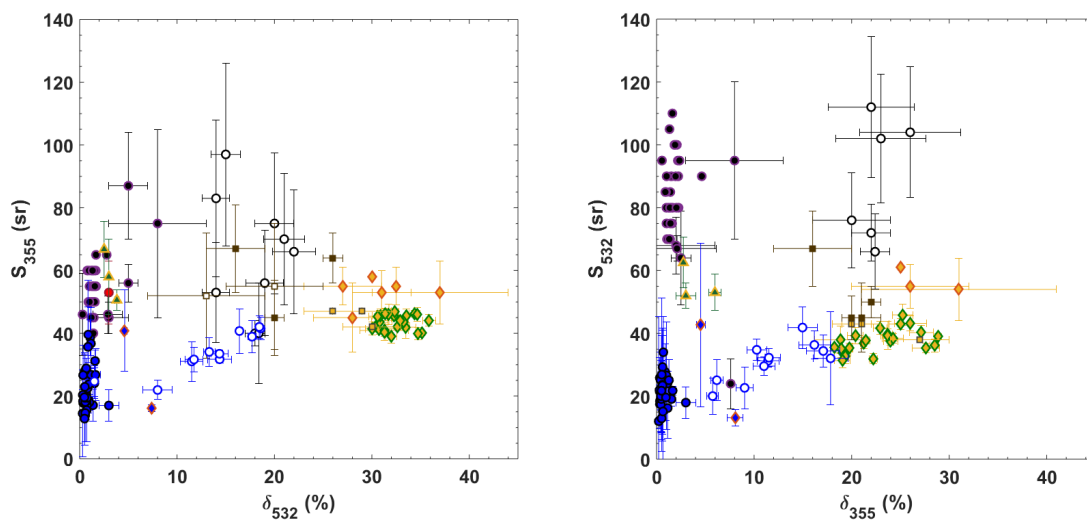
Figure A.19: 2D domains of the intensive optical parameters lidar ratio (S) and particle linear depolarization ratio (δ) at 355 and 532 nm, respectively (left) and vice versa (right).

Statistical analysis of the Ångström exponent

Similar to Sec. 3.2.2, a statistical analysis on the extinction- and backscatter-related Ångström exponents was performed (Fig. A.20). The number of samples for the extinction-related Ångström exponent were 43, 32, 2 and 33 for smoke, marine, pollution and dust category, respectively. Similarly, 141, 35, 2 and 31 samples were considered for the backscatter-related Ångström exponent at the wavelength pair of 355/532 nm. For the backscatter-related Ångström exponent at the wavelength pair of 532/1064 nm the number of samples were identical with the ones for the backscatter-related Ångström exponent at the wavelength pair of 355/532 nm. The mean values of the Ångström exponents of the four aerosol types are presented in Tab. A.1. However, due to the limited number of samples for the pollution category, the results are more indicative than qualitative. Regardless of the undersampling of the pollution category, the Ångström exponents reflect well (in a qualitative way) the aerosol particle size of the different aerosol types. Smoke and pollution, which consist of mainly fine mode aerosol particles, exhibit Ångström exponents (both extinction- and backscatter-related) greater than one. On the other hand, marine and dust, which consist of mainly coarse mode aerosol particles, exhibit lower values of Ångström exponents (both extinction- and backscatter-related).
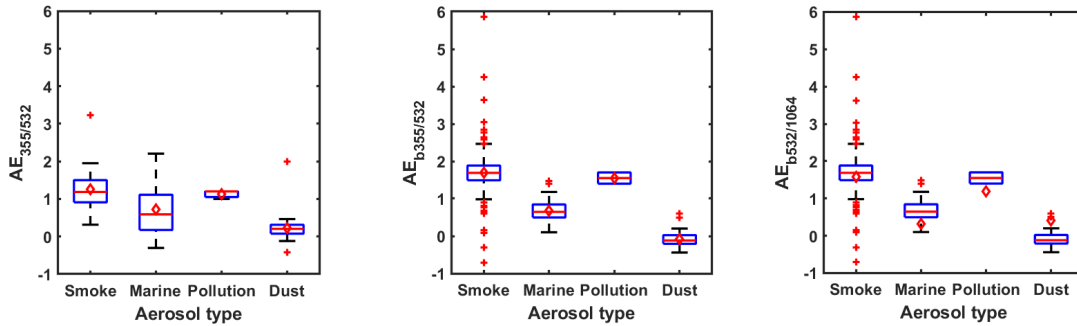
Figure A.20: Statistics of the extinction-related Ångström exponent (left), backscatter-related Ångström exponent at the wavelength pair 355/532 nm (middle) and backscatter-related Ångström exponent at the wavelength pair 532/1064 nm (right) for the four most commonly found pure aerosol types. It should be noted that for the dust category all dust measurements, regardless of the source origin, have been considered.

Table A.1: Mean values of the extinction- and backscatter-related Ångström exponents for the different pure aerosol types, as derived from the respective experimental basis.

Aerosol type	$AE_{355/532}$	$AE_{b355/532}$	$AE_{b532/1064}$
Smoke	1.3 ± 0.5	1.7 ± 0.7	1.6 ± 0.3
Marine	0.7 ± 0.7	1.7 ± 0.3	0.3 ± 0.4
Pollution	1.1 ± 0.1	1.5 ± 0.2	1.2 ± 0.3
Dust	0.2 ± 0.4	0 ± 0.2	0.4 ± 0.2

Appendix B

A-CARE: additional information

In addition to the A-LIFE case study presented in detail in Sec. 5.1, the results of the rest of the A-LIFE cases are provided here. The A-LIFE cases considered for evaluation with the OEM-based typing scheme were listed in Tab. 5.1. The optical properties for each case were used as an input to the OEM algorithm. Since both the particle linear depolarization ratio and the lidar ratio were available at two wavelengths (355 and 532 nm) three retrieval modes were available; 1, 2 and 5 (see retrieval mode configurations in Tab. 4.4). Out of all the retrievals that were conducted only those that produced a statistically significant (within the 95% confidence interval) optimal state were considered for further evaluation (e.g., for additional products derivations). In case of multiple statistically significant solutions per case study, the one that produced the closest forward-modelled values to the measurement (and hence, smaller outcome in the chi-squared test) was kept. The retrieved optimal state as well as the retrieval mode used and the statistical significance of the solution are shown in Tab. B.1. The optimal solution for the cases 1, 2 and 8 was not statistically significant and therefore not considered further or listed here.

The volume and number concentration per aerosol component were calculated as described in Sec. 4.2 and are presented in Tab. B.2 and Tab. B.3, respectively. These quantities are needed for the transformation from the relative volume per aerosol component to size distributions. As described in Sec. 5.1.4, in the next step the volume and number concentrations can be converted into the respective size distributions by using each mode's width and radius (Tab. 4.1). The effective radius of the observed aerosol mixture was also calculated (as Eq. (4.26)) for each case and is presented in Tab. B.4.

The results presented in the tables below (Tab. B.2, B.3, B.4) were compared against the corresponding parameters derived from the in-situ Falcon 20E measurements. The results for the case studies presented in Tab. B.1 are summarized in Fig. B.1, with the exception of case 5 that was analytically discussed in Sec. 5.1. Overall, the agreement between the size distributions obtained using the in-situ measurements and the lidar/OEM approach are

Table B.1: Contribution and uncertainties of the four basic aerosol components to the aerosol mixture observed in the selected cases for A-CARE. For each case, the retrieval mode is provided along with the statistical significance level. The cases left empty produced no statistically significant solutions.

Case #	OEM output (%)				Retrieval mode	Stat. significance
	FSA	CS	FSNA	CNS		
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	**66** ± 21	19 ± 15	12 ± 20	3 ± 5	2	95%
4	7 ± 11	**78** ± 22	5 ± 19	10 ± 14	1	95%
5	0 ± 8	4 ± 18	10 ± 11	**86** ± 22	1	95%
6	2 ± 8	2 ± 18	9 ± 10	**87** ± 22	1	95%
7	3 ± 9	28 ± 17	28 ± 18	**41** ± 16	5	95%
8	-	-	-	-	-	-

Table B.2: Volume concentrations ($\mu m^3/cm^3$) of each aerosol component along with uncertainties, as calculated from the optimal solutions of Tab. B.1.

Case #	Volume Concentration ($\mu m^3/cm^3$)			
	FSA	CS	FSNA	CNS
3	1.1 ± 0.15	0.3 ± 0.12	0.2 ± 0.13	0.05 ± 0.04
4	2.6 ± 1.8	29.4 ± 2.8	1.9 ± 2.3	3.8 ± 2.3
5	0 ± 1	1.7 ± 2.4	4.4 ± 2.3	37.9 ± 2.8
6	1 ± 1.4	1 ± 2.5	4.4 ± 2.4	42.4 ± 3
7	0.5 ± 0.5	4.3 ± 1.3	4.3 ± 1.3	6.3 ± 1.2

Table B.3: Number concentrations ($\#/cm^3$) of each aerosol component along with uncertainties, as calculated from the optimal solutions of Tab. B.1.

Case #	Number Concentration ($\#/cm^3$)			
	FSA	CS	FSNA	CNS
3	209.7 ± 29.3	0.03 ± 0.01	38.1 ± 26	0.005 ± 0.003
4	519.2 ± 364.4	2.8 ± 0.3	370.9 ± 452.1	0.4 ± 0.2
5	0 ± 195.6	0.2 ± 0.2	865.7 ± 464	3.7 ± 0.3
6	191.6 ± 268.9	0.09 ± 0.2	862.4 ± 466.8	4.1 ± 0.3
7	90.3 ± 102.2	0.4 ± 0.1	843.1 ± 264.4	0.6 ± 0.1

satisfactory. Case study 3 (11/04/2017) was dominated by fine mode particles with total contributions of 78% (66% FSA and 12% FSNA) in terms of relative volume. The volume

Table B.4: Effective radius for the aerosol mixtures of Tab. 5.1, as calculated from the corresponding optimal solutions of Tab. B.1.

Case #	Effective radius (µm)
3	0.18 ± 0.02
4	0.80 ± 0.25
5	0.85 ± 0.27
6	0.81 ± 0.26
7	0.40 ± 0.10

and number distributions (Fig. B.1(a) and (b), respectively) obtained with the two different methods compare very well with each other for the fine-mode particles, while for the coarse-mode particles a small underestimation (1 order of magnitude) by the lidar/OEM approach is observed. A similar behavior is observed for case study 4 (14/04/2017, Fig. B.1(c, d)) that was dominated by coarse-mode particles with a total relative volume contribution of 88%. For case study 6 (21/04/2017) both the volume and number distribution of the fine-mode particles (Fig. B.1e and f, respectively) are overestimated by the lidar/OEM approach. A rather large underestimation of the volume and number size distribution (up to three and two orders of magnitude, respectively) is observed for coarse mode particles with a diameter larger than 8 µm. Case 7 (25/04/2017) was the only case with roughly equal contributions of fine- and coarse-mode particles (56% and 44%, respectively). For that case the volume and number size distributions obtained with both methods fit almost perfectly with each other (Fig. B.1(g, h)). For the same case, the effective radius obtained with the two methods also fits perfectly. For the rest of the cases the effective radius derived with the lidar/OEM approach is overall larger than the one derived from the Falcon 20E in-situ measurements.
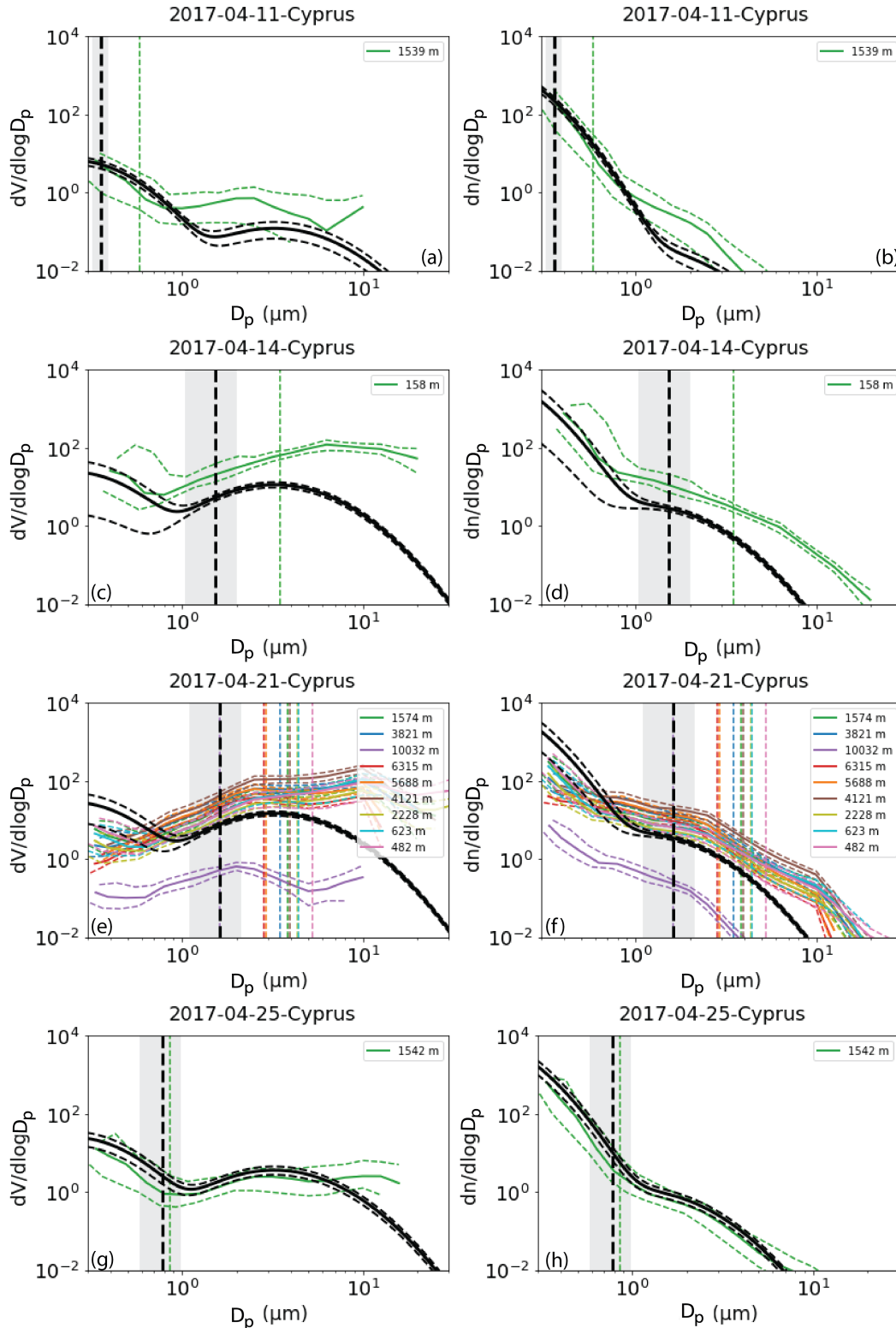
Figure B.1: Volume (left row) and number (right row) size distributions derived from lidar measurements following the lidar/OEM approach (black line, dashed lines indicate uncertainties) compared to the corresponding size distributions measured on board of Falcon 20E at various altitudes (colored lines and dashed lines), for the case studies 3, 4, 6 and 7 (top to bottom, respectively). The vertical lines (dashed) indicate the effective diameter (black for the lidar/OEM approach and colored for Falcon 20E).

Bibliography

Ackerman, A. S., Toon, O. B., Taylor, J. P., Johnson, D. W., Hobbs, P. V., and Ferek, R. J.: Effects of Aerosols on Cloud Albedo: Evaluation of Twomey's Parameterization of Cloud Susceptibility Using Measurements of Ship Tracks, journal = Journal of the Atmospheric Sciences, 57, 2684–2695, https://doi.org/10.1175/1520-0469(2000)057⟨2684: eoaoca⟩2.0.co;2, 2000.

Albrecht, B. A.: Aerosols, Cloud Microphysics, and Fractional Cloudiness, Science, 245, 1227–1230, https://doi.org/10.1126/science.245.4923.1227, 1989.

Alpert, P., Abramsky, R., and Neeman, B. U.: The prevailing summer synoptic system in Israel—Subtropical High, not Persian Trough, Isr. J. Earth Sci, 39, 93–102, 1990.

Althausen, D., Engelmann, R., Baars, H., Heese, B., Ansmann, A., Muller, D., and Komppula, M.: Portable Raman Lidar Polly(XT) for Automated Profiling of Aerosol Backscatter, Extinction, and Depolarization, Journal of Atmospheric and Oceanic Technology, 26, 2366–2378, https://doi.org/10.1175/2009jtecha1304.1, 2009.

Amiridis, V., Balis, D. S., Giannakaki, E., Stohl, A., Kazadzis, S., Koukouli, M. E., and Zanis, P.: Optical characteristics of biomass burning aerosols over Southeastern Europe determined from UV-Raman lidar measurements, Atmospheric Chemistry and Physics, 9, 2431–2440, https://doi.org/10.5194/acp-9-2431-2009, 2009.

Ansmann, A. and Müller, D.: Lidar and Atmospheric Aerosol Particles, in: Lidar: range-resolved optical remote sensing of the atmosphere, edited by Weitkamp, C., vol. 102 of *Springer Series in Optical Sciences*, pp. 105–141, Springer, New York, New York, USA, 1 edn., 2005.

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, Applied Optics, 31, 7113–7131, https://doi.org/10.1364/AO.31.007113, 1992.

Ansmann, A., Wagner, F., Althausen, D., Muller, D., Herber, A., and Wandinger, U.: European pollution outbreaks during ACE 2: Lofted aerosol plumes observed with Raman lidar at the Portuguese coast, Journal of Geophysical Research-Atmospheres, 106, 20 725–20 733, https://doi.org/10.1029/2000jd000091, 2001.

Ansmann, A., Wagner, F., Müller, D., Althausen, D., Herber, A., von Hoyningen-Huene, W., and Wandinger, U.: European pollution outbreaks during ACE 2: Optical particle properties inferred from multiwavelength lidar and star-Sun photometry, Journal of Geophysical Research: Atmospheres, 107, AAC 8–1–AAC 8–14, https://doi.org/10.1029/2001JD001109, 2002.

Ansmann, A., Wandinger, U., Le Rille, O., Lajas, D., and Straume, A. G.: Particle backscatter and extinction profiling with the spaceborne high-spectral-resolution Doppler lidar ALADIN: methodology and simulations, Applied Optics, 46, 6606–6622, https://doi.org/10.1364/AO.46.006606, 2007.

Ansmann, A., Petzold, A., Kandler, K., Tegen, I., Wendisch, M., Müller, D., Weinzierl, B., Müller, T., and Heintzenberg, J.: Saharan Mineral Dust Experiments SAMUM–1 and SAMUM–2: what have we learned?, Tellus B: Chemical and Physical Meteorology, 63, 403–429, https://doi.org/10.1111/j.1600-0889.2011.00555.x, 2011a.

Ansmann, A., Tesche, M., Seifert, P., Groß, S., Freudenthaler, V., Apituley, A., Wilson, K. M., Serikov, I., Linné, H., Heinold, B., Hiebsch, A., Schnell, F., Schmidt, J., Mattis, I., Wandinger, U., and Wiegner, M.: Ash and fine-mode particle mass profiles from EARLINET-AERONET observations over central Europe after the eruptions of the Eyjafjallajökull volcano in 2010, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2010JD015567, 2011b.

Ansmann, A., Seifert, P., Tesche, M., and Wandinger, U.: Profiling of fine and coarse particle mass: case studies of Saharan dust and Eyjafjallajökull/Grimsvötn volcanic plumes, Atmos. Chem. Phys., 12, 9399–9415, https://doi.org/10.5194/acp-12-9399-2012, 2012.

Ansmann, A., Baars, H., Chudnovsky, A., Mattis, I., Veselovskii, I., Haarig, M., Seifert, P., Engelmann, R., and Wandinger, U.: Extreme levels of Canadian wildfire smoke in the stratosphere over central Europe on 21-22 August 2017, Atmospheric Chemistry and Physics, 18, 11 831–11 845, https://doi.org/10.5194/acp-18-11831-2018, 2018.

Ansmann, A., Mamouri, R. E., Bühl, J., Seifert, P., Engelmann, R., Hofer, J., Nisantzi, A., Atkinson, J. D., Kanji, Z. A., Sierau, B., Vrekoussis, M., and Sciare, J.: Ice-nucleating particle versus ice crystal number concentrationin altocumulus and cirrus layers embedded

in Saharan dust:a closure study, Atmos. Chem. Phys., 19, 15 087–15 115, https://doi.org/ 10.5194/acp-19-15087-2019, 2019.

Antoine, D. and Morel, A.: Relative importance of multiple scattering by air molecules and aerosols in forming the atmospheric path radiance in the visible and near-infrared parts of the spectrum, Applied Optics, 37, 2245–2259, https://doi.org/10.1364/AO.37.002245, 1998.

Baars, H.: Aerosol profiling with lidar in the Amazon Basin during wet and dry season, PhD dissertation, Leipzig University, URL `https://nbn-resolving.org/urn:nbn:de: bsz:15-qucosa-98757`, 2011.

Baars, H., Ansmann, A., Althausen, D., Engelmann, R., Heese, B., Muller, D., Artaxo, P., Paixao, M., Pauliquevis, T., and Souza, R.: Aerosol profiling with lidar in the Amazon Basin during the wet and dry season, Journal of Geophysical Research-Atmospheres, 117, 16, https://doi.org/10.1029/2012jd018338, 2012.

Baars, H., Kanitz, T., Engelmann, R., Althausen, D., Heese, B., Komppula, M., Preissler, J., Tesche, M., Ansmann, A., Wandinger, U., Lim, J. H., Ahn, J. Y., Stachlewska, I. S., Amiridis, V., Marinou, E., Seifert, P., Hofer, J., Skupin, A., Schneider, F., Bohlmann, S., Foth, A., Bley, S., Pfuller, A., Giannakaki, E., Lihavainen, H., Viisanen, Y., Hooda, R. K., Pereira, S. N., Bortoli, D., Wagner, F., Mattis, I., Janicka, L., Markowicz, K. M., Achtert, P., Artaxo, P., Pauliquevis, T., Souza, R. A. F., Sharma, V. P., van Zyl, P. G., Beukes, J. P., Sun, J. Y., Rohwer, E. G., Deng, R. R., Mamouri, R. E., and Zamorano, F.: An overview of the first decade of Polly(NET): an emerging network of automated Raman-polarization lidars for continuous aerosol profiling, Atmos. Chem. Phys., 16, 5111–5137, https://doi.org/10.5194/acp-16-5111-2016, 2016.

Baars, H., Radenz, M., Floutsi, A. A., Engelmann, R., Althausen, D., Heese, B., Ansmann, A., Flament, T., Dabas, A., Trapon, D., Reitebuch, O., Bley, S., and Wandinger, U.: Californian Wildfire Smoke Over Europe: A First Example of the Aerosol Observing Capabilities of Aeolus Compared to Ground-Based Lidar, Geophysical Research Letters, 48, https://doi.org/10.1029/2020GL092194, 2021.

Barnaba, F. and Gobbi, G. P.: Aerosol seasonal variability over the Mediterranean region and relative impact of maritime, continental and Saharan dust particles over the basin from MODIS data in the year 2001, Atmos. Chem. Phys., 4, 2367–2391, https://doi.org/ 10.5194/acp-4-2367-2004, 2004.

Belegante, L., Nicolae, D., Nemuc, A., Talianu, C., and Derognat, C.: Retrieval of the

boundary layer height from active and passive remote sensors. Comparison with a NWP model, Acta Geophysica, 62, 276–289, https://doi.org/10.2478/s11600-013-0167-4, 2014.

Belegante, L., Bravo-Aranda, J. A., Freudenthaler, V., Nicolae, D., Nemuc, A., Ene, D., Alados-Arboledas, L., Amodeo, A., Pappalardo, G., D'Amico, G., Amato, F., Engelmann, R., Baars, H., Wandinger, U., Papayannis, A., Kokkalis, P., and Pereira, S. N.: Experimental techniques for the calibration of lidar depolarization channels in EARLINET, Atmospheric Measurement Techniques, 11, 1119–1141, https://doi.org/10.5194/amt-11-1119-2018, 2018.

Bohlmann, S., Baars, H., Radenz, M., Engelmann, R., and Macke, A.: Ship-borne aerosol profiling with lidar over the Atlantic Ocean: from pure marine conditions to complex dust-smoke mixtures, Atmos. Chem. Phys., 18, 9661–9679, https://doi.org/10.5194/acp-18-9661-2018, 2018.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., and Lohmann, U.: Clouds and aerosols, in: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 571–657, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Bréon, F.-M., Vermeulen, A., and Descloitres, J.: An evaluation of satellite aerosol products against sunphotometer measurements, Remote Sensing of Environment, 115, 3102–3111, 2011.

Burton, S. P., Ferrare, R. A., Hostetler, C. A., Hair, J. W., Rogers, R. R., Obland, M. D., Butler, C. F., Cook, A. L., Harper, D. B., and Froyd, K. D.: Aerosol classification using airborne High Spectral Resolution Lidar measurements - methodology and examples, Atmospheric Measurement Techniques, 5, 73–98, https://doi.org/10.5194/amt-5-73-2012, 2012.

Burton, S. P., Hair, J. W., Kahnert, M., Ferrare, R. A., Hostetler, C. A., Cook, A. L., Harper, D. B., Berkoff, T. A., Seaman, S. T., Collins, J. E., Fenn, M. A., and Rogers, R. R.: Observations of the spectral dependence of linear particle depolarization ratio of aerosols using NASA Langley airborne High Spectral Resolution Lidar, Atmos. Chem. Phys., 15, 13 453–13 473, https://doi.org/10.5194/acp-15-13453-2015, 2015.

Cattrall, C., Reagan, J., Thome, K., and Dubovik, O.: Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2004JD005124, 2005.

Chi-Square Table: in: The Concise Encyclopedia of Statistics, pp. 76–77, Springer New York, New York, NY, https://doi.org/10.1007/978-0-387-32833-1_56, 2008.

Clarke, A. D., Howell, S., Quinn, P. K., Bates, T. S., Ogren, J. A., Andrews, E., Jefferson, A., Massling, A., Mayol-Bracero, O., Maring, H., Savoie, D., and Cass, G.: INDOEX aerosol: A comparison and summary of chemical, microphysical, and optical properties observed from land, ship, and aircraft, Journal of Geophysical Research: Atmospheres, 107, https://doi.org/10.1029/2001JD000572, 2002.

D'Amico, G., Amodeo, A., Baars, H., Binietoglou, I., Freudenthaler, V., Mattis, I., Wandinger, U., and Pappalardo, G.: EARLINET Single Calculus Chain - overview on methodology and strategy, Atmospheric Measurement Techniques, 8, 4891–4916, https://doi.org/10.5194/amt-8-4891-2015, 2015.

do Carmo, J. P., de Villele, G., Wallace, K., Lefebvre, A., Ghose, K., Kanitz, T., Chassat, F., Corselle, B., Belhadj, T., and Bravetti, P.: ATmospheric LIDar (ATLID): Pre-Launch Testing and Calibration of the European Space Agency Instrument That Will Measure Aerosols and Thin Clouds in the Atmosphere, Atmosphere, 12, 76, 2021.

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F., Volten, H., Muñoz, O., Veihelmann, B., van der Zande, W. J., Leon, J.-F., Sorokin, M., and Slutsker, I.: Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust, Journal of Geophysical Research: Atmospheres, 111, https://doi.org/10.1029/2005JD006619, 2006.

Eck, T. F., Holben, B. N., Reid, J. S., Dubovik, O., Smirnov, A., O'Neill, N. T., Slutsker, I., and Kinne, S.: Wavelength dependence of the optical depth of biomass burning, urban, and desert dust aerosols, Journal of Geophysical Research: Atmospheres, 104, 31 333–31 349, https://doi.org/10.1029/1999JD900923, 1999.

Engelmann, R., Kanitz, T., Baars, H., Heese, B., Althausen, D., Skupin, A., Wandinger, U., Komppula, M., Stachlewska, I. S., Amiridis, V., Marinou, E., Mattis, I., Linne, H., and Ansmann, A.: The automated multiwavelength Raman polarization and water-vapor lidar Polly(XT): the neXT generation, Atmospheric Measurement Techniques, 9, 1767–1784, https://doi.org/10.5194/amt-9-1767-2016, 2016.

Flamant, P., Cuesta, J., Denneulin, M.-L., Dabas, A., and Huber, D.: ADM-Aeolus retrieval algorithms for aerosol and cloud products, Tellus A: Dynamic Meteorology and Oceanography, 60, 273–286, https://doi.org/10.1111/j.1600-0870.2007.00287.x, 2008.

Floutsi, A. A., Korras-Carraca, M. B., Matsoukas, C., Hatzianastassiou, N., and Biskos, G.: Climatology and trends of aerosol optical depth over the Mediterranean basin during

the last 12 years (2002-2014) based on Collection 006 MODIS-Aqua data, Science of the Total Environment, 551, 292–303, https://doi.org/10.1016/j.scitotenv.2016.01.192, 2016.

Floutsi, A. A., Baars, H., Radenz, M., Haarig, M., Yin, Z., Seifert, P., Jimenez, C., Ansmann, A., Engelmann, R., Barja, B., Zamorano, F., and Wandinger, U.: Advection of Biomass Burning Aerosols towards the Southern Hemispheric Mid-Latitude Station of Punta Arenas as Observed with Multiwavelength Polarization Raman Lidar, Remote Sensing, 13, 138, 2021.

Foth, A. and Pospichal, B.: Optimal estimation of water vapour profiles using a combination of Raman lidar and microwave radiometer, Atmospheric Measurement Techniques, 10, 3325–3344, https://doi.org/10.5194/amt-10-3325-2017, 2017.

Freudenthaler, V.: About the effects of polarising optics on lidar signals and the Delta 90 calibration, Atmospheric Measurement Techniques, 9, 4181–4255, https://doi.org/10.5194/amt-9-4181-2016, 2016.

Freudenthaler, V., Esselborn, M., Wiegner, M., Heese, B., Tesche, M., Ansmann, A., Müller, D., Althausen, D., Wirth, M., Fix, A., Ehret, G., Knippertz, P., Toledano, C., Gasteiger, J., Garhammer, M., and Seefeldner, M.: Depolarization ratio profiling at several wavelengths in pure Saharan dust during SAMUM 2006, Tellus Series B-Chemical and Physical Meteorology, 61, 165–179, https://doi.org/10.1111/j.1600-0889.2008.00396.x, 2009.

Friedl, M. A., McIver, D. K., Hodges, J. C., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., and Cooper, A.: Global land cover mapping from MODIS: algorithms and early results, Remote sensing of Environment, 83, 287–302, 2002.

Ganor, E., Levin, Z., and Van Grieken, R.: Composition of individual aerosol particles above the Israelian Mediterranean coast during the summer time, Atmospheric Environment, 32, 1631–1642, https://doi.org/10.1016/S1352-2310(97)00397-X, 1998.

Gasteiger, J., Wiegner, M., Groß, S., Freudenthaler, V., Toledano, C., Tesche, M., and Kandler, K.: Modelling lidar-relevant optical properties of complex mineral dust aerosols, Tellus B: Chemical and Physical Meteorology, 63, 725–741, https://doi.org/10.1111/j.1600-0889.2011.00559.x, 2011.

Giannakaki, E., Pfüller, A., Korhonen, K., Mielonen, T., Laakso, L., Vakkari, V., Baars, H., Engelmann, R., Beukes, J. P., Van Zyl, P. G., Josipovic, M., Tiitta, P., Chiloane, K., Piketh, S., Lihavainen, H., Lehtinen, K. E. J., and Komppula, M.: One year of Raman lidar observations of free-tropospheric aerosol layers over South Africa, Atmos. Chem. Phys., 15, 5429–5442, https://doi.org/10.5194/acp-15-5429-2015, 2015.

Giannakaki, E., van Zyl, P. G., Müller, D., Balis, D., and Komppula, M.: Optical and microphysical characterization of aerosol layers over South Africa by means of multi-wavelength depolarization and Raman lidar measurements, Atmos. Chem. Phys., 16, 8109–8123, https://doi.org/10.5194/acp-16-8109-2016, 2016.

Gkikas, A., Hatzianastassiou, N., Mihalopoulos, N., Katsoulis, V., Kazadzis, S., Pey, J., Querol, X., and Torres, O.: The regime of intense desert dust episodes in the Mediterranean based on contemporary satellite observations and ground measurements, Atmos. Chem. Phys., 13, 12 135–12 154, https://doi.org/10.5194/acp-13-12135-2013, 2013.

Gkikas, A., Houssos, E., Lolis, C., Bartzokas, A., Mihalopoulos, N., and Hatzianastassiou, N.: Atmospheric circulation evolution related to desert-dust episodes over the Mediterranean, Quarterly Journal of the Royal Meteorological Society, 141, 1634–1645, 2015.

Gkikas, A., Basart, S., Hatzianastassiou, N., Marinou, E., Amiridis, V., Kazadzis, S., Pey, J., Querol, X., Jorba, O., Gassó, S., and Baldasano, J. M.: Mediterranean intense desert dust outbreaks and their vertical structure based on remote sensing data, Atmos. Chem. Phys., 16, 8609–8642, https://doi.org/10.5194/acp-16-8609-2016, 2016.

Groß, S., Tesche, M., Freudenthaler, V., Toledano, C., Wiegner, M., Ansmann, A., Althausen, D., and Seefeldner, M.: Characterization of Saharan dust, marine aerosols and mixtures of biomass-burning aerosols and dust by means of multi-wavelength depolarization and Raman lidar measurements during SAMUM–2, Tellus Series B-Chemical and Physical Meteorology, 63, 706–724, https://doi.org/10.1111/j.1600-0889.2011.00556.x, 2011.

Groß, S., Freudenthaler, V., Wiegner, M., Gasteiger, J., Geiß, A., and Schnell, F.: Dual-wavelength linear depolarization ratio of volcanic aerosols: Lidar measurements of the Eyjafjallajökull plume over Maisach, Germany, Atmospheric Environment, 48, 85–96, https://doi.org/10.1016/j.atmosenv.2011.06.017, 2012.

Groß, S., Esselborn, M., Weinzierl, B., Wirth, M., Fix, A., and Petzold, A.: Aerosol classification by airborne high spectral resolution lidar observations, Atmospheric Chemistry and Physics, 13, 2487–2505, https://doi.org/10.5194/acp-13-2487-2013, 2013.

Haarig, M., Ansmann, A., Engelmann, R., Baars, H., Toledano, C., Torres, B., Althausen, D., Radenz, M., and Wandinger, U.: First triple-wavelength lidar observations of depolarization and extinction-to-backscatter ratios of Saharan dust, Atmos. Chem. Phys., 22, 355–369.

Haarig, M., Ansmann, A., Gasteiger, J., Kandler, K., Althausen, D., Baars, H., Radenz, M., and Farrell, D. A.: Dry versus wet marine particle optical properties: RH dependence of depolarization ratio, backscatter, and extinction from multiwavelength lidar measurements during SALTRACE, Atmos. Chem. Phys., 17, 14 199–14 217, https://doi.org/10.5194/acp-17-14199-2017, 2017.

Haarig, M., Ansmann, A., Baars, H., Jimenez, C., Veselovskii, I., Engelmann, R., and Althausen, D.: Depolarization and lidar ratios at 355, 532, and 1064 nm and microphysical properties of aged tropospheric and stratospheric Canadian wildfire smoke, Atmos. Chem. Phys., 18, 11 847–11 861, https://doi.org/10.5194/acp-18-11847-2018, 2018.

Hansen, J., Sato, M., and Ruedy, R.: Radiative forcing and climate response, Journal of Geophysical Research: Atmospheres, 102, 6831–6864, https://doi.org/10.1029/96JD03436, 1997.

Heese, B., Baars, H., Bohlmann, S., Althausen, D., and Deng, R.: Continuous vertical aerosol profiling with a multi-wavelength Raman polarization lidar over the Pearl River Delta, China, Atmos. Chem. Phys., 17, 6679–6691, https://doi.org/10.5194/acp-17-6679-2017, 2017.

Heese, B., Floutsi, A. A., Baars, H., Althausen, D., Hofer, J., Herzog, A., Mewes, S., Radenz, M., and Schechner, Y. Y.: The vertical aerosol type distribution above Israel – 2 years of lidar observations at the coastal city of Haifa, Atmos. Chem. Phys., 22, 1633–1648, https://doi.org/10.5194/acp-22-1633-2022, 2022.

Herman, J. R., Bhartia, P. K., Torres, O., Hsu, C., Seftor, C., and Celarier, E.: Global distribution of UV-absorbing aerosols from Nimbus 7/TOMS data, Journal of Geophysical Research: Atmospheres, 102, 16 911–16 922, https://doi.org/10.1029/96JD03680, 1997.

Hess, M., Koepke, P., and Schult, I.: Optical Properties of Aerosols and Clouds: The Software Package OPAC, Bulletin of the American Meteorological Society, 79, 831–844, https://doi.org/10.1175/1520-0477(1998)079⟨0831:opoaac⟩2.0.co;2, 1998.

Hill, T. and Lewicki, P.: Statistics methods and applications., 2007.

Hofer, J., Althausen, D., Abdullaev, S. F., Makhmudov, A. N., Nazarov, B. I., Schettler, G., Engelmann, R., Baars, H., Fomba, K. W., Müller, K., Heinold, B., Kandler, K., and Ansmann, A.: Long-term profiling of mineral dust and pollution aerosol with multiwavelength polarization Raman lidar at the Central Asian site of Dushanbe, Tajikistan: case studies, Atmos. Chem. Phys., 17, 14 559–14 577, https://doi.org/10.5194/acp-17-14559-2017, 2017.

Hofer, J., Ansmann, A., Althausen, D., Engelmann, R., Baars, H., Fomba, K. W., Wandinger, U., Abdullaev, S. F., and Makhmudov, A. N.: Optical properties of Central Asian aerosol relevant for spaceborne lidar applications and aerosol typing at 355 and 532 nm, Atmos. Chem. Phys., 20, 9265–9280, https://doi.org/10.5194/acp-20-9265-2020, 2020.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET - A federated instrument network and data archive for aerosol characterization, Remote Sensing of Environment, 66, 1–16, https://doi.org/10.1016/s0034-4257(98)00031-5, 1998.

Holben, B. N., Tanre, D., Smirnov, A., Eck, T. F., Slutsker, I., Abuhassan, N., Newcomb, W. W., Schafer, J. S., Chatenet, B., Lavenu, F., Kaufman, Y. J., Castle, J. V., Setzer, A., Markham, B., Clark, D., Frouin, R., Halthore, R., Karneli, A., O'Neill, N. T., Pietras, C., Pinker, R. T., Voss, K., and Zibordi, G.: An emerging ground-based aerosol climatology: Aerosol optical depth from AERONET, Journal of Geophysical Research-Atmospheres, 106, 12 067–12 097, https://doi.org/10.1029/2001jd900014, 2001.

Holzer-Popp, T., de Leeuw, G., Griesfeller, J., Martynenko, D., Klüser, L., Bevan, S., Davies, W., Ducos, F., Deuzé, J. L., Graigner, R. G., Heckel, A., von Hoyningen-Hüne, W., Kolmonen, P., Litvinov, P., North, P., Poulsen, C. A., Ramon, D., Siddans, R., Sogacheva, L., Tanre, D., Thomas, G. E., Vountas, M., Descloitres, J., Griesfeller, J., Kinne, S., Schulz, M., and Pinnock, S.: Aerosol retrieval experiments in the ESA Aerosol_cci project, Atmos. Meas. Tech., 6, 1919–1957, https://doi.org/10.5194/amt-6-1919-2013, 2013.

Hu, Q., Goloub, P., Veselovskii, I., Bravo-Aranda, J. A., Popovici, I. E., Podvin, T., Haeffelin, M., Lopatin, A., Dubovik, O., Pietras, C., Huang, X., Torres, B., and Chen, C.: Long-range-transported Canadian smoke plumes in the lower stratosphere over northern France, Atmos. Chem. Phys., 19, 1173–1193, https://doi.org/10.5194/acp-19-1173-2019, 2019.

Illingworth, A. J., Barker, H. W., Beljaars, A., Ceccaldi, M., Chepfer, H., Clerbaux, N., Cole, J., Delanoe, J., Domenech, C., Donovan, D. P., Fukuda, S., Hirakata, M., Hogan, R. J., Huenerbein, A., Kollias, P., Kubota, T., Nakajima, T., Nakajima, T. Y., Nishizawa, T., Ohno, Y., Okamoto, H., Oki, R., Sato, K., Satoh, M., Shephard, M. W., Velazquez-Blazquez, A., Wandinger, U., Wehr, T., and van Zadelhoff, G. J.: THE EARTHCARE SATELLITE The Next Step Forward in Global Measurements of Clouds, Aerosols, Pre-

cipitation, and Radiation, Bulletin of the American Meteorological Society, 96, 1311–1332, https://doi.org/10.1175/bams-d-12-00227.1, 2015.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Kanitz, T., Ansmann, A., Engelmann, R., and Althausen, D.: North-south cross sections of the vertical aerosol distribution over the Atlantic Ocean from multiwavelength Raman/polarization lidar during Polarstern cruises, Journal of Geophysical Research-Atmospheres, 118, 2643–2655, https://doi.org/10.1002/jgrd.50273, 2013.

Kanitz, T., Ansmann, A., Foth, A., Seifert, P., Wandinger, U., Engelmann, R., Baars, H., Althausen, D., Casiccia, C., and Zamorano, F.: Surface matters: limitations of CALIPSO V3 aerosol typing in coastal regions, Atmospheric Measurement Techniques, 7, 2061–2072, https://doi.org/10.5194/amt-7-2061-2014, 2014.

Kaufman, Y. J., Tanré, D., and Boucher, O.: A satellite view of aerosols in the climate system, Nature, 419, 215–223, https://doi.org/10.1038/nature01091, 2002.

Kim, M. H., Omar, A. H., Tackett, J. L., Vaughan, M. A., Winker, D. M., Trepte, C. R., Hu, Y. X., Liu, Z. Y., Poole, L. R., Pitts, M. C., Kar, J., and Magill, B. E.: The CALIPSO version 4 automated aerosol classification and lidar ratio selection algorithm, Atmospheric Measurement Techniques, 11, 6107–6135, https://doi.org/10.5194/amt-11-6107-2018, 2018.

Koepke, P., Hess, M., Schult, I., and Shettle, E.: Global Aerosol Data Set, URL `http://hdl.handle.net/10068/256361`, 1997.

Koçak, M., Nimmo, M., Kubilay, N., and Herut, B.: Spatio-temporal aerosol trace metal concentrations and sources in the Levantine Basin of the Eastern Mediterranean, Atmospheric Environment, 38, 2133–2144, https://doi.org/10.1016/j.atmosenv.2004.01.020, 2004.

Kubilay, N., Nickovic, S., Moulin, C., and Dulac, F.: An illustration of the transport and deposition of mineral dust onto the eastern Mediterranean, Atmospheric Environment, 34, 1293–1303, https://doi.org/10.1016/S1352-2310(99)00179-X, 2000.

Kumar, K. R., Kang, N., and Yin, Y.: Classification of key aerosol types and their frequency distributions based on satellite remote sensing data at an industrially polluted

city in the Yangtze River Delta, China, International Journal of Climatology, 38, 320–336, https://doi.org/10.1002/joc.5178, 2018.

Laskina, O., Morris, H. S., Grandquist, J. R., Qin, Z., Stone, E. A., Tivanski, A. V., and Grassian, V. H.: Size Matters in the Water Uptake and Hygroscopic Growth of Atmospherically Relevant Multicomponent Aerosol Particles, The Journal of Physical Chemistry A, 119, 4489–4497, https://doi.org/10.1021/jp510268p, 2015.

Luria, M., Peleg, M., Sharf, G., Tov-Alper, D. S., Spitz, N., Ben Ami, Y., Gawii, Z., Lifschitz, B., Yitzchaki, A., and Seter, I.: Atmospheric sulfur over the east Mediterranean region, Journal of Geophysical Research: Atmospheres, 101, 25 917–25 930, https://doi.org/10.1029/96JD01579, 1996.

Mahalanobis, P. C.: On the generalized distance in statistics, National Institute of Science of India, 1936.

Mamouri, R. E. and Ansmann, A.: Fine and coarse dust separation with polarization lidar, Atmos. Meas. Tech., 7, 3717–3735, https://doi.org/10.5194/amt-7-3717-2014, 2014.

Mamouri, R. E. and Ansmann, A.: Potential of polarization lidar to provide profiles of CCN- and INP-relevant aerosol parameters, Atmospheric Chemistry and Physics, 16, 5905–5931, https://doi.org/10.5194/acp-16-5905-2016, 2016.

Mamouri, R. E. and Ansmann, A.: Potential of polarization/Raman lidar to separate fine dust, coarse dust, maritime, and anthropogenic aerosol profiles, Atmospheric Measurement Techniques, 10, 3403–3427, https://doi.org/10.5194/amt-10-3403-2017, 2017.

Marinou, E., Amiridis, V., Binietoglou, I., Tsikerdekis, A., Solomos, S., Proestakis, E., Konsta, D., Papagiannopoulos, N., Tsekeri, A., Vlastou, G., Zanis, P., Balis, D., Wandinger, U., and Ansmann, A.: Three-dimensional evolution of Saharan dust transport towards Europe based on a 9-year EARLINET-optimized CALIPSO dataset, Atmos. Chem. Phys., 17, 5893–5919, https://doi.org/10.5194/acp-17-5893-2017, 2017.

MATLAB: Version 9.4.0.813654 (R2018a), The MathWorks Inc., Natick, Massachusetts, 2018.

Mattis, I., Ansmann, A., Müller, D., Wandinger, U., and Althausen, D.: Dual-wavelength Raman lidar observations of the extinction-to-backscatter ratio of Saharan dust, Geophys. Res. Lett., 29, https://doi.org/10.1029/2002GL014721, 2002.

Mattis, I., Ansmann, A., Müller, D., Wandinger, U., and Althausen, D.: Multiyear aerosol observations with dual-wavelength Raman lidar in the framework of EARLINET, Jour-

nal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2004JD004600, 2004.

Mattis, I., D'Amico, G., Baars, H., Amodeo, A., Madonna, F., and Iarlori, M.: EARLINET Single Calculus Chain - technical - Part 2: Calculation of optical products, Atmospheric Measurement Techniques, 9, 3009–3029, https://doi.org/10.5194/amt-9-3009-2016, 2016.

McCormick, R. A. and Ludwig, J. H.: Climate Modification by Atmospheric Aerosols, Science, 156, 1358–1359, https://doi.org/10.1126/science.156.3780.1358, 1967.

Mewes, S.: Characterization of aerosol properties by lidar measurements at Haifa, Israel, Master's thesis, Leipzig Institute for Meteorology, Faculty of Physics and Earth Sciences, University of Leipzig, Leipzig, Germany, 2018.

Mishchenko, M. I., Geogdzhayev, I. V., Cairns, B., Rossow, W. B., and Lacis, A. A.: Aerosol retrievals over the ocean by use of channels 1 and 2 AVHRR data: sensitivity analysis and preliminary results, Applied Optics, 38, 7325–7341, https://doi.org/10.1364/AO.38.007325, 1999.

Moulin, C., Lambert, C. E., Dulac, F., and Dayan, U.: Control of atmospheric export of dust from North Africa by the North Atlantic Oscillation, Nature, 387, 691–694, 1997.

Murayama, T., Müller, D., Wada, K., Shimizu, A., Sekiguchi, M., and Tsukamoto, T.: Characterization of Asian dust and Siberian smoke with multi-wavelength Raman lidar over Tokyo, Japan in spring 2003, Geophysical Research Letters, 31, https://doi.org/10.1029/2004GL021105, 2004.

Müller, D., Ansmann, A., Wagner, F., Franke, K., and Althausen, D.: European pollution outbreaks during ACE 2: Microphysical particle properties and single-scattering albedo inferred from multiwavelength lidar observations, Journal of Geophysical Research: Atmospheres, 107, https://doi.org/10.1029/2001JD001110, 2002.

Müller, D., Franke, K., Ansmann, A., Althausen, D., and Wagner, F.: Indo-Asian pollution during INDOEX: Microphysical particle properties and single-scattering albedo inferred from multiwavelength lidar observations, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2003JD003538, 2003.

Müller, D., Mattis, I., Wandinger, U., Ansmann, A., Althausen, D., and Stohl, A.: Raman lidar observations of aged Siberian and Canadian forest fire smoke in the free troposphere over Germany in 2003: Microphysical particle characterization, Journal of Geophysical Research-Atmospheres, 110, 16, https://doi.org/10.1029/2004jd005756, 2005.

Müller, D., Ansmann, A., Mattis, I., Tesche, M., Wandinger, U., Althausen, D., and Pisani, G.: Aerosol-type-dependent lidar ratios observed with Raman lidar, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006jd008292, 2007.

Nabat, P., Somot, S., Mallet, M., Chiapello, I., Morcrette, J. J., Solmon, F., Szopa, S., Dulac, F., Collins, W., Ghan, S., Horowitz, L. W., Lamarque, J. F., Lee, Y. H., Naik, V., Nagashima, T., Shindell, D., and Skeie, R.: A 4-D climatology (1979–2009) of the monthly tropospheric aerosol optical depth distribution over the Mediterranean region from a comparative evaluation and blending of remote sensing and model products, Atmos. Meas. Tech., 6, 1287–1314, https://doi.org/10.5194/amt-6-1287-2013, 2013.

NCEP: National Centers For Environmental Prediction/National Weather Service/NOAA/U.S. Department Of Commerce: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999, https://doi.org/10.5065/D6M043C6, 2000.

Nicolae, D., Vasilescu, J., Talianu, C., Binietoglou, I., Nicolae, V., Andrei, S., and Antonescu, B.: A neural network aerosol-typing algorithm based on lidar data, Atmos. Chem. Phys., 18, 14 511–14 537, https://doi.org/10.5194/acp-18-14511-2018, 2018.

Nisantzi, A., Mamouri, R. E., Ansmann, A., and Hadjimitsis, D.: Injection of mineral dust into the free troposphere during fire events observed with polarization lidar at Limassol, Cyprus, Atmospheric Chemistry and Physics, 14, 12 155–12 165, https://doi.org/10.5194/acp-14-12155-2014, 2014.

Ohneiser, K., Ansmann, A., Baars, H., Seifert, P., Barja, B., Jimenez, C., Radenz, M., Teisseire, A., Floutsi, A., Haarig, M., Foth, A., Chudnovsky, A., Engelmann, R., Zamorano, F., Bühl, J., and Wandinger, U.: Smoke of extreme Australian bushfires observed in the stratosphere over Punta Arenas, Chile, in January 2020: optical thickness, lidar ratios, and depolarization ratios at 355 and 532 nm, Atmos. Chem. Phys., 20, 8003–8015, https://doi.org/10.5194/acp-20-8003-2020, 2020.

Ohneiser, K., Ansmann, A., Engelmann, R., Ritter, C., Chudnovsky, A., Veselovskii, I., Baars, H., Gebauer, H., Griesche, H., Radenz, M., Hofer, J., Althausen, D., Dahlke, S., and Maturilli, M.: Siberian fire smoke in the High-Arctic winter stratosphere observed during MOSAiC 2019-2020, Atmos. Chem. Phys. Discuss., 2021, 1–36, https://doi.org/10.5194/acp-2021-117, 2021.

Omar, A. H., Won, J.-G., Winker, D. M., Yoon, S.-C., Dubovik, O., and McCormick, M. P.: Development of global aerosol models using cluster analysis of Aerosol Robotic

Network (AERONET) measurements, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/10.1029/2004JD004874, 2005.

Omar, A. H., Winker, D. M., Vaughan, M. A., Hu, Y., Trepte, C. R., Ferrare, R. A., Lee, K.-P., Hostetler, C. A., Kittaka, C., Rogers, R. R., Kuehn, R. E., and Liu, Z.: The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, Journal of Atmospheric and Oceanic Technology, 26, 1994–2014, https://doi.org/10.1175/2009jtecha1231.1, 2009.

Papadimas, C., Hatzianastassiou, N., Mihalopoulos, N., Querol, X., and Vardavas, I.: Spatial and temporal variability in aerosol properties over the Mediterranean basin based on 6-year (2000–2006) MODIS data, Journal of Geophysical Research: Atmospheres, 113, 2008.

Papagiannopoulos, N., Mona, L., Amodeo, A., D'Amico, G., Gumà Claramunt, P., Pappalardo, G., Alados-Arboledas, L., Guerrero-Rascado, J. L., Amiridis, V., Kokkalis, P., Apituley, A., Baars, H., Schwarz, A., Wandinger, U., Binietoglou, I., Nicolae, D., Bortoli, D., Comerón, A., Rodríguez-Gómez, A., Sicard, M., Papayannis, A., and Wiegner, M.: An automatic observation-based aerosol typing method for EARLINET, Atmos. Chem. Phys., 18, 15 879–15 901, https://doi.org/10.5194/acp-18-15879-2018, 2018.

Pappalardo, G., Amodeo, A., Apituley, A., Comeron, A., Freudenthaler, V., Linné, H., Ansmann, A., Bösenberg, J., D'Amico, G., Mattis, I., Mona, L., Wandinger, U., Amiridis, V., Alados-Arboledas, L., Nicolae, D., and Wiegner, M.: EARLINET: towards an advanced sustainable European aerosol lidar network, Atmos. Meas. Tech., 7, 2389–2409, https://doi.org/10.5194/amt-7-2389-2014, 2014.

Pisso, I., Sollum, E., Grythe, H., Kristiansen, N. I., Cassiani, M., Eckhardt, S., Arnold, D., Morton, D., Thompson, R. L., and Groot Zwaaftink, C. D.: The Lagrangian particle dispersion model FLEXPART version 10.4, Geoscientific Model Development, 12, 4955–4997, 2019.

Preißler, J., Wagner, F., Pereira, S. N., and Guerrero-Rascado, J. L.: Multi-instrumental observation of an exceptionally strong Saharan dust outbreak over Portugal, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2011JD016527, 2011.

Radenz, M., Bühl, J., Seifert, P., Baars, H., Engelmann, R., Barja González, B., Mamouri, R. E., Zamorano, F., and Ansmann, A.: Hemispheric contrasts in ice formation in stratiform mixed-phase clouds: disentangling the role of aerosol and dynamics with ground-based remote sensing, Atmos. Chem. Phys., 21, 17 969–17 994, https://doi.org/10.5194/acp-21-17969-2021, 2021a.

Radenz, M., Seifert, P., Baars, H., Floutsi, A. A., Yin, Z., and Bühl, J.: Automated time–height-resolved air mass source attribution for profiling remote sensing applications, Atmos. Chem. Phys., 21, 3015–3033, https://doi.org/10.5194/acp-21-3015-2021, 2021b.

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, vol. 2, World scientific, 2000.

Sasano, Y. and Browell, E. V.: Light scattering characteristics of various aerosol types derived from multiple wavelength lidar observations, Applied Optics, 28, 1670–1679, https://doi.org/10.1364/AO.28.001670, 1989.

Sassen, K.: The Polarization Lidar Technique for Cloud Research: A Review and Current Assessment., Bulletin of the American Meteorological Society, 72, 1848–1866, https://doi.org/10.1175/1520-0477(1991)072⟨1848:TPLTFC⟩2.0.CO;2, 1991.

Schotland, R. M., Sassen, K., and Stone, R.: Observations by Lidar of Linear Depolarization Ratios for Hydrometeors, Journal of Applied Meteorology and Climatology, 10, 1011–1017, https://doi.org/10.1175/1520-0450(1971)010⟨1011:oblold⟩2.0.co;2, 1971.

Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J. B., Cohen, M. D., and Ngan, F.: NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System, Bulletin of the American Meteorological Society, 96, 2059–2077, https://doi.org/10.1175/bams-d-14-00110.1, 2015.

Stoffelen, A., Marseille, G. J., Bouttier, F., Vasiljevic, D., de Haan, S., and Cardinali, C.: ADM-Aeolus Doppler wind lidar Observing System Simulation Experiment, Quarterly Journal of the Royal Meteorological Society, 132, 1927–1947, https://doi.org/10.1256/qj.05.83, 2006.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmospheric Chemistry and Physics, 5, 2461–2474, 2005.

Straume, A., Rennie, M., Isaksen, L., de Kloe, J., Marseille, G.-J., Stoffelen, A., Flament, T., Stieglitz, H., Dabas, A., Huber, D., Reitebuch, O., Lemmerz, C., Lux, O., Marksteiner, U., Weiler, F., Witschas, B., Meringer, M., Schmidt, K., Nikolaus, I., Geiss, A., Flamant, P., Kanitz, T., Wernham, D., von Bismarck, J., Bley, S., Fehr, T., Floberghagen, R., and Parinello, T.: ESA's Space-Based Doppler Wind Lidar Mission Aeolus – First Wind and Aerosol Product Assessment Results, EPJ Web Conf., 237, 01 007, URL 10.1051/epjconf/202023701007, 2020.

Sugimoto, N., Matsui, I., Shimizu, A., Uno, I., Asai, K., Endoh, T., and Nakajima, T.: Observation of dust and anthropogenic aerosol plumes in the Northwest Pacific with a two-wavelength polarization lidar on board the research vessel Mirai, Geophysical Research Letters, 29, 7–1–7–4, https://doi.org/10.1029/2002GL015112, 2002.

Sugimoto, N., Nishizawa, T., Shimizu, A., Matsui, I., and Jin, Y.: Characterization of aerosols in East Asia with the Asian Dust and Aerosol Lidar Observation Network (AD-Net), vol. 9262 of *SPIE Asia-Pacific Remote Sensing*, SPIE, 2014.

Tesche, M.: Vertical profiling of aerosol optical properties with multiwavelength aerosol lidar during the Saharan Mineral Dust Experiments, PhD dissertation, University of Leipzig, URL `https://nbn-resolving.org/urn:nbn:de:bsz:15-qucosa-71257`, last access: 14 September 2021, 2011.

Tesche, M., Ansmann, A., Muller, D., Althausen, D., Mattis, I., Heese, B., Freudenthaler, V., Wiegner, M., Esselborn, M., Pisani, G., and Knippertz, P.: Vertical profiling of Saharan dust with Raman lidars and airborne HSRL in southern Morocco during SA-MUM, Tellus Series B-Chemical and Physical Meteorology, 61, 144–164, https://doi.org/10.1111/j.1600-0889.2008.00390.x, 2009a.

Tesche, M., Ansmann, A., Muller, D., Althausen, D., Engelmann, R., Freudenthaler, V., and Groß, S.: Vertically resolved separation of dust and smoke over Cape Verde using multiwavelength Raman and polarization lidars during Saharan Mineral Dust Experiment 2008, Journal of Geophysical Research-Atmospheres, 114, 14, https://doi.org/10.1029/2009jd011862, 2009b.

Tesche, M., Groß, S., Ansmann, A., Muller, D., Althausen, D., Freudenthaler, V., and Esselborn, M.: Profiling of Saharan dust and biomass-burning smoke with multiwavelength polarization Raman lidar at Cape Verde, Tellus Series B-Chemical and Physical Meteorology, 63, 649–676, https://doi.org/10.1111/j.1600-0889.2011.00548.x, 2011a.

Tesche, M., Müller, D., Groß, S., Ansmann, A., Althausen, D., Freudenthaler, V., Weinzierl, B., Veira, A., and Petzold, A.: Optical and microphysical properties of smoke over Cape Verde inferred from multiwavelength lidar measurements, Tellus B: Chemical and Physical Meteorology, 63, 677–694, https://doi.org/10.1111/j.1600-0889.2011.00549.x, 2011b.

Torres, O., Bhartia, P. K., Herman, J. R., Ahmad, Z., and Gleason, J.: Derivation of aerosol properties from satellite measurements of backscattered ultraviolet radiation: heoretical basis.

Torres, O., Bhartia, P. K., Herman, J. R., Sinyuk, A., Ginoux, P., and Holben, B.: A Long-Term Record of Aerosol Optical Depth from TOMS Observations and Comparison to AERONET Measurements, Journal of the Atmospheric Sciences, 59, 398–413, https://doi.org/10.1175/1520-0469(2002)059⟨0398:altroa⟩2.0.co;2, 2002.

Twomey, S.: The Influence of Pollution on the Shortwave Albedo of Clouds, Journal of Atmospheric Sciences, 34, 1149–1152, https://doi.org/10.1175/1520-0469(1977)034⟨1149:tiopot⟩2.0.co;2, 1977.

Val Martin, M., Kahn, R. A., and Tosca, M. G.: A Global Analysis of Wildfire Smoke Injection Heights Derived from Space-Based Multi-Angle Imaging, Remote Sensing, 10, https://doi.org/10.3390/rs10101609, 2018.

Vardavas, I. M. and Taylor, F.: Radiation and Climate: Atmospheric energy budget from satellite remote sensing, vol. 138, Oxford University Press, 2011.

Wandinger, U.: Multiple-scattering influence on extinction- and backscatter-coefficient measurements with Raman and high-spectral-resolution lidars, Applied Optics, 37, 417–427, https://doi.org/10.1364/AO.37.000417, 1998.

Wandinger, U., Müller, D., Böckmann, C., Althausen, D., Matthias, V., Bösenberg, J., Weiß, V., Fiebig, M., Wendisch, M., Stohl, A., and Ansmann, A.: Optical and microphysical characterization of biomass- burning and industrial-pollution aerosols from-multiwavelength lidar and aircraft measurements, Journal of Geophysical Research: Atmospheres, 107, https://doi.org/10.1029/2000JD000202, 2002.

Wandinger, U., Baars, H., Engelmann, R., Hünerbein, A., Horn, S., Kanitz, T., Donovan, D., van Zadelhoff, G.-J., Daou, D., Fischer, J., von Bismarck, J., Filipitsch, F., Docter, N., Eisinger, M., Lajas, D., and Wehr, T.: HETEAC: The Aerosol Classification Model for EarthCARE, EPJ Web of Conferences, 119, 01 004, https://doi.org/10.1051/epjconf/201611901004, 2016a.

Wandinger, U., Freudenthaler, V., Baars, H., Amodeo, A., Engelmann, R., Mattis, I., Groß, S., Pappalardo, G., Giunta, A., D'Amico, G., Chaikovsky, A., Osipenko, F., Slesar, A., Nicolae, D., Belegante, L., Talianu, C., Serikov, I., Linne, H., Jansen, F., Apituley, A., Wilson, K. M., de Graaf, M., Trickl, T., Giehl, H., Adam, M., Comeron, A., Munoz-Porcar, C., Rocadenbosch, F., Sicard, M., Tomas, S., Lange, D., Kumar, D., Pujadas, M., Molero, F., Fernandez, A. J., Alados-Arboledas, L., Bravo-Aranda, J. A., Navas-Guzman, F., Guerrero-Rascado, J. L., Granados-Munoz, M. J., Preissler, J., Wagner, F., Gausa, M., Grigorov, I., Stoyanov, D., Iarlori, M., Rizi, V., Spinelli, N., Boselli, A., Wang, X., Lo Feudo, T., Perrone, M. R., De Tomasi, F., and Burlizzi, P.: EARLINET instrument

intercomparison campaigns: overview on strategy and results, Atmospheric Measurement Techniques, 9, 1001–1023, https://doi.org/10.5194/amt-9-1001-2016, 2016b.

Weinzierl, B., Sauer, D., Esselborn, M., Petzold, A., Veira, A., Rose, M., Mund, S., Wirth, M., Ansmann, A., Tesche, M., Groß, S., and Freudenthaler, V.: Microphysical and optical properties of dust and tropical biomass burning aerosol layers in the Cape Verde region—an overview of the airborne in situ and lidar measurements during SAMUM–2, Tellus B, 63, 589–618, https://doi.org/10.1111/j.1600-0889.2011.00566.x, 2011.

Welton, E. J., Campbell, J. R., Berkoff, T. A., Spinhirne, J. D., Tsay, S.-C., Holben, B., Shiobara, M., and Starr, D. O.: The micro-pulse lidar network (MPLNET), 2002.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO mission and CALIOP data processing algorithms, Journal of Atmospheric and Oceanic Technology, 26, 2310–2323, 2009.

Young, S. A. and Vaughan, M. A.: The Retrieval of Profiles of Particulate Extinction from Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) Data: Algorithm Description, Journal of Atmospheric and Oceanic Technology, 26, 1105–1119, https://doi.org/10.1175/2008jtecha1221.1, 2009.

Zhang, R., Khalizov, A. F., Pagels, J., Zhang, D., Xue, H., and McMurry, P. H.: Variability in morphology, hygroscopicity, and optical properties of soot aerosols during atmospheric processing, Proceedings of the National Academy of Sciences, 105, 10 291–10 296, https://doi.org/10.1073/pnas.0804860105, 2008.

Ångström, A.: The parameters of atmospheric turbidity, Tellus, 16, 64–75, 1964.

List of Abbreviations

A-CARE A-LIFE and EarthCARE joint reseach program. 66, 68, 69, 124, 157, 158

A-LIFE Absorbing aerosol layers in a changing climate: aging, LIFEtime and dynamics. 65–69, 71, 73, 75, 107, 123, 150

AD-Net Asian Dust and Aerosol Lidar Observation Network. 2

Aeolus Wind Lidar Mission Aeolus. 3

AERONET Aerosol Robotic Network. 1, 10, 14

AGL Above Ground Level. 92

ALADIN Atmospheric Laser Doppler Instrument. 3

ANN Artificial Neural Network. 12–15, 17

AOD Aerosol Optical Depth. 2, 10

ASL Above Sea Level. 92

ATLID ATmospheric LIDar. 4, 14, 15, 39

AVHRR Advanced Very High Resolution Radiometer. 2

BBA Biomass-burning aerosol. 28

BBR Broad-Band Radiometer. 4, 14

CADEX Central Asian Dust EXperiment. 15, 24, 25, 27, 107

CALIOP Cloud-Aerosol Lidar with Orthogonal Polarization. 3, 9, 10

CALIPSO Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations. 3, 4, 9–11, 13, 15, 28

CAS-DPOL Cloud and Aerosol Spectrometer with Detector for Polarization Detection. 66, 73

CCI Climate Change Initiative. 14

CCN Cloud Condensation Nuclei. 1, 9

CNRS French National Centre for Scientific Research. 1

CNS Coarse Non–Spherical aerosol component. 32, 37–41, 43, 52, 54–57, 59–62, 70–73, 75, 78–82, 86, 87, 94, 95, 97–102, 105, 124, 152

CPC Condensation Particle Counter. 66

CPR Cloud Profiling Radar. 4

CS Coarse Spherical aerosol component. 32, 37–42, 52, 54–60, 62, 70, 72, 78, 79, 81, 82, 86–88, 94, 95, 97–102, 105, 108, 124, 151, 152

CyCARE Cyprus Clouds Aerosols and Rain Experiment. 24, 66

DACAPO-PESO Dynamics, Aerosol, Cloud and Precipitation Observations in the Pristine Environment of the Southern Ocean . 24, 27

DLR Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center). 65, 66, 69, 74

EARLINET European Aerosol Research Lidar Network. 2, 12, 13, 15, 92

EarthCARE Cloud, Aerosol and Radiation Explorer. 4, 14, 15, 17, 38, 39, 66, 71, 106, 107

ESA European Space Agency. 3, 4, 14

EUCAARI European Integrated Project on Aerosol Cloud Climate and Air Quality interactions. 23, 27

FLEXPART FLEXible PARTicle dispersion model. 96

FSA Fine Spherical Absorbing aerosol component. 32, 37–41, 52–57, 59, 60, 62, 63, 70, 72, 78–83, 86–89, 94, 95, 97–101, 105, 124, 152

FSNA Fine Spherical Non–Absorbing aerosol component. 32, 37–41, 52, 54–62, 70, 72, 73, 78–83, 86–88, 94, 95, 97–102, 105, 124, 152

GADS Global Aerosol Data Set. 12

GFS Global Forecast System. 96

HETEAC Hybrid End-To-End Aerosol Classification. 4, 7, 14, 15, 17, 19, 23–25, 32, 38, 39, 66, 75, 106–108, 157

HSRL High Spectral Resolution Lidar. 3, 4, 11

HYSPLIT Hybrid Single Particle Lagrangian Integrated Trajectory Model. 69

INP Ice-Nucleating Particles. 1, 9

JAXA Japanese Aerospace Exploration Agency. 4

lidar Light Detection And Ranging. 2, 66

LM Levenberg-Marquardt method. 45, 46

LPDM Lagrangian Particle Dispersion Model. 96

LUT Lookup table. 15

MBL Marine Boundary Layer. 55, 57, 94

MCS Monte Carlo Simulations. 49

MODIS Moderate Resolution Imaging Spectroradiometer. 3, 96

MOSAiC Multidisciplinary drifting Observatory for the Study of Arctic Climate. 24, 27

MPLNET Micro-Pulse Lidar Network. 2

MSI Multi-Spectral Imager. 4, 14

NASA National Aeronautics and Space Administration. 1

NATALI Neural Network Aerosol Typing Algorithm Based on Lidar Data. 12–14, 16

OEM Optimal Estimation Method. 35, 37–39, 41–43, 46, 47, 49, 50, 52, 53, 55, 58, 61, 62, 65, 66, 68–70, 72–75, 78–83, 85, 87, 88, 90–92, 98, 102, 105–108, 123, 125, 126, 149, 151, 152, 155, 157, 158, 161

OPAC Optical Properties of Aerosols and Clouds. 12, 14

OPC Optical Particle Counters. 66, 73

PBL Planetary Boundary Layer. 20

PDF Probability Density Function. 36

POLIPHON Polarization-Lidar Photometer Networking. 9, 16, 65, 80–83, 151

PS Polarstern research vessel. 25

RH Relative Humidity. 12, 108

SAMUM Saharan Mineral Dust Experiment. 7, 9, 15, 20, 23–27, 60, 61, 65, 76, 149, 150

TOMS Total Ozone Mapping Spectromete. 2

List of Figures

List of Tables

List of Symbols

α	Extinction coefficient	Mm^{-1}
α^*	*A priori* extinction coefficient	$\mathrm{Mm}^{-1}/\frac{\mu\mathrm{m}^3}{\mathrm{cm}^3}$
α'_{j}	Aerosol-component attributable extinction coefficient	$\mathrm{Mm}^{-1}/\frac{\mu\mathrm{m}^3}{\mathrm{cm}^3}$
α^*_{tot}	Total extinction coefficient	$\mathrm{Mm}^{-1}/\frac{\mu\mathrm{m}^3}{\mathrm{cm}^3}$
α_{j}	Absolute aerosol-component attributable extinction coefficient	Mm^{-1}
α_{tot}	Lidar-measured extinction coefficient	Mm^{-1}
β	Backscatter coefficient	$\mathrm{Mm}^{-1}\mathrm{sr}^{-1}$
β^*	*A priori* backscatter coefficient	$\mathrm{Mm}^{-1}\mathrm{sr}^{-1}/\frac{\mu\mathrm{m}^3}{\mathrm{cm}^3}$
β'_{j}	Aerosol-component attributable backscatter coefficient	$\mathrm{Mm}^{-1}\mathrm{sr}^{-1}/\frac{\mu\mathrm{m}^3}{\mathrm{cm}^3}$
β^*_{tot}	Total backscatter coefficient	$\mathrm{Mm}^{-1}\mathrm{sr}^{-1}/\frac{\mu\mathrm{m}^3}{\mathrm{cm}^3}$
β_{j}	Absolute aerosol-component attributable backscatter coefficient	$\mathrm{Mm}^{-1}\mathrm{sr}^{-1}$
β_{tot}	Lidar-measured backscatter coefficient	$\mathrm{Mm}^{-1}\mathrm{sr}^{-1}$
γ	Levenberg - Marquardt parameter	m^{-1}
δ	Linear particle depolarization ratio	1
δ^*	*A priori* linear particle depolarization ratio	1
ζ	Constant related to constraints	$-$
λ	Wavelength	nm
σ	Aerosol mode width (variance)	1
χ^2	Statistical quantity	$-$

χ_{thr}	Threshold value for χ^2 test	1
$\parallel$	Co–polarized component of a quantity	–
$\perp$	Cross–polarized component of a quantity	–
$\widehat{\boldsymbol{S}}$	Covariance matrix of the retrieved state (a posteriori)	–
$\widehat{\boldsymbol{x}}$	Optimal state vector or optimal solution	–
$\boldsymbol{F}$	Forward model	–
$\boldsymbol{J_\alpha}$	Initial guess costs	–
$\boldsymbol{J_y}$	Observation costs	–
$\boldsymbol{J}$	Cost function	–
$\boldsymbol{J}_{con}$	Constraint costs	–
$\boldsymbol{K}$	Weighting function matrix, or Kernel or Jacobian	–
$\boldsymbol{S}_\alpha$	Covariance matrix of the initial guess of the state vector	–
$\boldsymbol{S}_{\delta\widehat{y}}$	Covariance matrix between $\boldsymbol{y}$ and $\boldsymbol{F}(\widehat{\boldsymbol{x}})$	–
$\boldsymbol{S}_\epsilon$	Error covariance matrix related to the lidar measurements	–
$\boldsymbol{x}$	State vector	–
$\boldsymbol{x}_\alpha$	Initial guess of the state vector	–
$\boldsymbol{y}$	Measurement vector	–
$\mathring{A}$	Ångström exponent	–
A_{j}	Aerosol-component attributable surface area	$\mu\mathrm{m}^2/\mathrm{cm}^3$
A_{tot}	Total surface area	$\mu\mathrm{m}^2/\mathrm{cm}^3$
$C_{\beta_{\lambda_1/\lambda_2}}$	Backscatter-related color ratio	1
$D\mathrm{p}$	Particle diameter	m
d_{f}	Degrees of freedom of the measurement	–
i	Iteration step	–
j	Aerosol-component index	–

m	Refractive index	–
m_I	Imaginary part of the refractive index	–
m_R	Real part of the refractive index	–
N_j	Aerosol-component attributable number concentration	$\#/\mathrm{cm}^3$
N_tot	OEM-derived total number concentration	$\#/\mathrm{cm}^3$
R	Range	m
$r_{0,\mathrm{N}}$	Number-distribution-related aerosol mode radius	$\mu\mathrm{m}$
$r_{0,\mathrm{V}}$	Volume-distribution-related aerosol mode radius	$\mu\mathrm{m}$
R_eff	Effective radius	$\mu\mathrm{m}$
S	Lidar ratio	sr
V_j	Aerosol-component attributable volume concentration	$\mu\mathrm{m}^3/\mathrm{cm}^3$
V_tot	OEM-derived total volume concentration	$\mu\mathrm{m}^3/\mathrm{cm}^3$

Acknowledgements

At this point, I would like to thank several people that have helped me in one way or another during the last years and have enabled the realization of this dissertation.

First of all I would like to express my gratitude to the two reviewers of my dissertation Prof. Dr. Andreas Macke and Prof. Dr. Dimitrios Balis. In addition, I would like thank my whole Ph.D. commission, Prof. Dr. Andreas Macke, Prof. Ina Tegen and Prof. Johannes Quass for their critical input and support.

I am grateful to my entire supervision committee and especially to Dr. Ulla Wandinger and Dr. Holger Baars. Ulla has been an inspiring scientist for me, with vast knowledge on diverse research fields. From every discussion we shared I learned something new or acquired a different point of view on a topic. She has always kept me on the right path, and her valuable input (from scattering theory to punctuation rules) and support have made this thesis possible. I would like to thank Holger for all the scientific discussions, troubleshooting, for answering all my questions and for pushing me further, whenever it was necessary. I am deeply grateful for the help he provided during this thesis as well as for providing valuable feedback on my work. I would also like to thank Dr. Moritz Haarig, for our fruitful collaboration on the A-CARE project, for teaching me a lot about in-situ data, and for being patient during our small chats in German.

I am thankful to the entire lidar group of TROPOS for the fruitful working environment, the never-ending-emerging great ideas as well as all the (not only) science-related German vocabulary that they helped me acquire during the coffee breaks. Special thanks go to my fellow Ph.D. students (some of which have graduated by now), Marios, Vasilis, Julian, Audrey, Martin, Hannes, Cristofer, Willi, Matthias as well as to those who have passed by room 2.15. And of course I would like to thank Christian for his support, understanding and everything else.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια και τους φίλους μου για την ατελείωτη υποστήριξη, υπομονή και την πίστη τους σε μένα. Ιδιαίτερα ευχαριστώ την Κατερίνα και τον Γιάννη με τους οποίους μεγαλώσαμε παρέα και με έχουν βοηθήσει σε διαφορετικές καταστάσεις (αδύνατον να τις απαριθμήσω όλες). Είμαι ευγνώμων στον Κώστα για την ατελείωτη προθυμία και ευελιξία του στο να με βοηθήσει πάντα, καθώς και για τις μουσικές μας κουβέντες. Ευχαριστώ την γιαγιά μου Ελένη, για την αγάπη, την υποστήριξη και την βοήθεια της οπου και αν βρίσκομαι. Φυσικά τον παππού μου Γιάννη και τον πατέρα μου Μιχάλη, οι οποίοι δυστυχώς δεν πρόλαβαν να με δουν να μεγαλώνω. Τέλος, θα ήθελα να ευχαριστήσω για όλα την μητέρα μου Πηγή, στην οποία και αφιερώνω αυτή τη διατριβή. Χωρίς την καθοδήγηση, στοργή και αγάπη της όχι μόνο δεν θα βρισκόμουν εδώ που είμαι σήμερα αλλά και δεν θα ήμουν το άτομο που είμαι σήμερα.