Rigorous dynamic error estimates in diverse remote sensing applications : concept, validation and realization in GRASP algorithm

Milagros Estefania Herrera

▶ **To cite this version:**

Milagros Estefania Herrera. Rigorous dynamic error estimates in diverse remote sensing applications : concept, validation and realization in GRASP algorithm. Earth Sciences. Université de Lille; Universidad Tecnológica Nacional. Facultad Regional Buenos Aires (Buenos Aires, Argentine), 2022. English. NNT : 2022ULILR012 . tel-03902068

HAL Id: tel-03902068

https://theses.hal.science/tel-03902068

Submitted on 15 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Lille

École doctorale des Sciences de la Matière, du Rayonnement et de l'Environnement

Thèse pour obtenir le grade de:

Docteur de l'Université de Lille

Spécialité:

Terre, Enveloppes fluides

Soutenue publiquement le 12 avril 2022 par

Milagros Estefanía Herrera

Rigorous dynamic error estimates in diverse remote sensing applications: concept, validation and realization in GRASP algorithm

---

devant le jury composé de :

Alexander Sinyuk, Directeur de Recherche, NASA/GSFC	**Rapporteur**
Julian Gröbner, Directeur de Recherche, PMOD/WRF	**Rapporteur**
Philippe Dubuisson, Professeur des Universités, Université de Lille	**Président**
Africa Barreto, Chargée de Recherches, Izaña Atmospheric Research Center	**Examinatrice**
Walter Legnani, Professeur des Universités, UTN.BA	**Examinateur**
Oleg Dubovik, Directeur de Recherche, CNRS/Université de Lille	**Directeur**
Benjamin Torres, Maître de Conférences, Université de Lille	**Co-encadrant**
Pablo Ristori, Professeur des Universités, CITEDEF/UTN.BA	**Co-directeur**
Lucas Bali, Maître de Conférences, CONICET	**Co-encadrent**

Titre de la thèse en français:

Estimations rigoureuses des erreurs dynamiques dans diverses applications de télédétection : concept, validation et réalisation dans l'algorithme GRASP

La thèse a été effectuée en cotutelle au sein du Laboratoire d'Optique Atmosphérique de l'Université de Lille et la Universidad Tecnológica Nacional.

Laboratoire d'Optique Atmosphérique
LOA UMR 8518 CNRS / Université de Lille
Bâtiment P5 - Université Lille - Sciences et Technologies
59655 Villeneuve d'Ascq Cedex

Universidad Tecnológica Nacional
Facultad Regional Buenos Aires
Medrano 951 (C1179AAQ) CABA

Abstract

The understanding of the uncertainties in the retrieval of the aerosol and surface properties is very important for an adequate characterization of the processes that occur in the atmosphere. However, the reliable characterization of the error budget of the retrieval products is a very challenging aspect that currently remains not fully resolved in most remote sensing approaches. The level of uncertainties for the majority of the remote sensing products relies mostly on post-processing validations and inter comparisons with other data while the dynamic errors are rarely provided. This study describes, discuses and evaluates a concept realized in GRASP (Generalized Retrieval of Atmosphere and Surface Properties) algorithm for providing the dynamic estimates of uncertainties for retrieved parameter. The approach employs rigorous concept of statistical optimization for estimating the effects of measurement uncertainties propagation to the retrieval results. The approach accounts for the effect of both random and systematic uncertainties in the initial data and provides error estimates both for directly retrieved parameters included in the retrieval state vector and for the characteristics derived from these parameters. The efficiency of realized error estimation concept is extensively analyzed for GRASP applications for aerosol retrieval from ground-based observations by sun/sky photometer and lidar. The diverse aspects of the generations and evaluations of the error estimates are discussed and illustrated. The evaluation of the error estimates was realized using the series of comprehensive sensitivity tests when simulated sun/sky photometer measurements and lidar data are perturbed by random and systematic errors and inverted. The results of the retrievals and their error estimations obtained in the tests are analyzed and evaluated. The tests are conducted for the different observations of several types of aerosols including biomass burning, urban, dust and their mixtures. The study considers popular observations by AERONET sun/sky radiometer at 440, 675, 870 and 1020 nm and multi-wavelength elastic lidar at 355, 532 and 1064 nm. The sun/sky radiometer data are inverted alone or together with lidar data. The analysis shows that the generated error estimates overall satisfactory of the uncertainties of different retrieved aerosol characteristics including aerosol size distribution, complex refractive index, single scattering albedo, lidar ratios, aerosol vertical profiles, etc. Also, the analysis shows that the main ob-

served error dynamic agrees well with the errors tendencies commonly known from the retrieval experience. For example, the serious retrieval accuracy limitations for all aerosol types are associated with the situations with low optical depth. Also, for observations of multi-component aerosol mixtures, the reliable characterization of each component is possible only in limited situations, for example from radiometric data obtained for low solar zenith angle observations or from a combination of radiometric and lidar data. At the same time, total optical properties of aerosol mixtures tend to be always retrieved satisfactorily. In addition, the study includes the analysis of the detailed structure of correlation matrices for the retrieval errors of mono- and multi-component aerosols. The conducted analysis of error correlation appears to be a useful approach for optimizing observations schemes and retrieval setups. The illustration of the developed approach application to real data is provided for co-located observations of sun/sky photometer and lidar over Buenos Aires. Furthermore, the preliminary results for utilizing the error estimates for the retrieval of aerosol from satellite data are provided.

Résumé

La compréhension des incertitudes dans la restitution des propriétés des aérosols et des surfaces est très importante pour une caractérisation adéquate des processus qui se produisent dans l'atmosphère. Cependant, la caractérisation fiable du budget d'erreur des produits de restitution est un aspect complexe qui n'est pas encore entièrement résolu dans la plupart des approches de la télédétection. Le niveau d'incertitude pour la majorité des produits de télédétection repose principalement sur des validations post-traitement et des comparaisons avec d'autres données, alors que les erreurs dynamiques sont rarement fournies. Cette étude décrit, analyse et évalue un concept réalisé dans l'algorithme GRASP (Generalized Retrieval of Atmosphere and Surface Properties) qui a pour objectif de fournir les estimations dynamiques des incertitudes pour les paramètres restitués. L'approche utilise un concept rigoureux d'optimisation statistique pour estimer les effets de la propagation des incertitudes de mesure sur les résultats de la restitution. L'approche tient compte de l'effet des incertitudes aléatoires et systématiques dans les données initiales et fournit des estimations d'erreurs tant pour les paramètres directement restitués inclus dans le vecteur d'état de restitution, que pour les caractéristiques dérivées de ces paramètres. L'efficacité du concept d'estimation des erreurs réalisées est analysée en profondeur pour les applications GRASP pour la restitution d'aérosols à partir d'observations depuis le sol par le photomètre et le lidar.

Les divers aspects des générations et des évaluations des estimations d'erreurs sont discutés et illustrés. L'évaluation des estimations d'erreurs a été réalisée à l'aide de la série exhaustive de tests de sensibilité lorsque des mesures photométriques solaires simulées et des données lidar sont perturbées par des erreurs aléatoires et systématiques et inversées. Les résultats des restitutions et leurs estimations d'erreurs obtenues dans les tests sont analysés et évalués. Les essais sont effectués pour les différentes observations de plusieurs types d'aérosols, issus de la combustion de biomasse, urbains, de poussières et leurs mélanges. L'étude tient compte des observations faites par les photomètres solaires AERONET effectuées à 440, 675, 870 et 1020 nm et lidar élastique multi-longueurs d'onde à 355, 532 et 1064 nm. Les données du photomètre solaire sont inversées seules ou avec les données lidar. L'analyse montre que l'erreur générée estime

globalement de façon satisfaisante les incertitudes des différentes caractéristiques des aérosols restitués, y compris la distribution en taille des aérosols, l'indice de réfraction complexe, l'albédo à diffusion simple, les rapports du lidar, les profils verticaux des aérosols, etc. En outre, l'analyse montre que les principales erreurs dynamiques observées concorde bien avec les tendances d'erreurs communément connues par les expériences de restitution. Par exemple, les limites de précision de restitution pour tous les types d'aérosols sont associées aux situations de faible épaisseur optique. En outre, pour les observations de mélanges d'aérosols multi-composants, la caractérisation fiable de chaque composant n'est possible que dans des situations limitées, par exemple à partir de données radiométriques obtenues pour des observations à faible angle zénithal solaire ou à partir d'une combinaison de données radiométriques et lidar. Dans le même temps, les propriétés optiques totales des mélanges d'aérosols ont tendance à être toujours restituées de manière satisfaisante. En outre, l'étude comprend l'analyse de la structure détaillée des matrices de corrélation pour les erreurs de restitution des aérosols mono et multi-composants. L'analyse de la corrélation des erreurs semble être une approche utile pour optimiser les schémas d'observations et les configurations de récupération. L'illustration de l'application de l'approche développée aux données réelles est fournie pour les observations co-localisées du photomètre solaire et du lidar au-dessus de Buenos Aires. De plus, les résultats préliminaires de l'utilisation des estimations d'erreurs pour la restitution des aérosols à partir des données satellitaires sont fournis.

Resumen

La comprensión de las incertidumbres para el cálculo de las propiedades de los aerosoles y de la superficie es de suma importancia para una adecuada caracterización de los procesos que ocurren en la atmósfera. Sin embargo, una caracterización confiable de la estimación de los errores totales de los productos obtenidos es un aspecto muy desafiante que actualmente no está totalmente resuelto en la mayoría de los enfoques de teledetección. El nivel de incertidumbre provisto por la mayoría de los productos de teledetección está basado principalmente en validaciones posteriores al procesamiento y en intercomparaciones con otros datos, mientras que los errores dinámicos rara vez son suministrados. Este estudio describe, discute y evalúa un concepto realizado en el algoritmo GRASP (Generalized Retrieval of Atmosphere and Surface Properties) para proporcionar las estimaciones dinámicas de las incertidumbres de los parámetros obtenidos. Este enfoque emplea un riguroso concepto de optimización estadística para estimar los efectos de la propagación de las incertidumbres de la medición a los resultados calculados. Además, tiene en cuenta el efecto de las incertidumbres tanto aleatorias como sistemáticas de los datos iniciales y proporciona estimaciones de error tanto para los parámetros calculados directamente, incluidos en el vector de estado de recuperación, como para las características derivadas a partir de estos parámetros. La eficiencia del concepto de estimación de errores realizado es analizado extensamente para las aplicaciones de GRASP para la recuperación de aerosoles a partir de observaciones terrestres mediante el uso de fotómetro solar y lidar. También se discuten e ilustran los diversos aspectos de las generaciones y evaluaciones de las estimaciones de error. La evaluación de las estimaciones de error se llevó a cabo mediante una serie exhaustiva de pruebas de sensibilidad en las cuales las mediciones simuladas del fotómetro solar y los datos del lidar fueron perturbadas por errores aleatorios y sistemáticos para luego invertirlas. A continuación se analizaron y evaluaron los resultados de las inversiones y las estimaciones de error obtenidas en las pruebas realizadas. Las mismas se realizaron para las diferentes observaciones de varios tipos de aerosoles, incluyendo los de quema de biomasa, urbanos, polvo y la mezcla entre ambos. El estudio considera las observaciones del fotómetro solar de AERONET a 440, 675, 870 y 1020 nm y el lidar elástico de múltiples longitudes de onda a 355, 532 y 1064 nm. Los datos del

fotómetro solar se invirtieron solos o junto con los datos del lidar. El análisis muestra que las estimaciones del error generado estima en general de forma satisfactoria las incertidumbres de las características de los distintos aerosoles calculados, incluyendo la distribución del tamaño de los aerosoles, el índice de refracción complejo, el albedo de dispersión simple, la razón lidar, los perfiles verticales de los aerosoles, etc. Por otro lado, el análisis muestra que los principales errores dinámicos observados coinciden con las tendencias de error comúnmente conocidas por la experiencia en los parámetros calculados. Por ejemplo, las mayores limitaciones en la precisión de las propiedades obtenidas a partir de la inversión para todos los tipos de aerosoles, están asociadas a condiciones de observación con bajos espesores ópticos de los aerosoles. Además, para las observaciones de mezclas de múltiples componentes de aerosoles, la caracterización fiable de cada componente sólo es posible en un limitado número de casos, por ejemplo, cuando se cuenta con datos radiométricos obtenidos durante observaciones de bajo ángulo cenital solar o a partir de e una combinación de datos radiométricos combinados con observaciones lidar. Al mismo tiempo, las propiedades ópticas totales de las mezclas de aerosoles tienden a recuperarse siempre de forma satisfactoria. El estudio incluye el análisis de la estructura detallada de las matrices de correlación para los errores calculados de los aerosoles monocomponentes y multicomponentes. El análisis realizado de la correlación de errores parece ser un enfoque útil para optimizar los esquemas de observación y las configuraciones para las inversiones. La ilustración de la aplicación del enfoque desarrollado a los datos reales es provista para las observaciones co-localizadas del fotómetro solar y el lidar sobre Buenos Aires. Finalmente, se presentan los resultados preliminares del uso de las estimaciones de error de las propiedades de los aerosoles obtenidas a partir de datos satelitales.

Acknowledgements

I would like to thank my advisors, Oleg and Benja, for everything they have taught me over all these years, for their patience and commitment, for all those scientific discussions, and for all that personal support always present whenever I needed them. This thesis would not have been the same without their help! I will also be eternally grateful because my life in Lille was easier with people like them. They have helped me manage a new language and integrate into a new culture by always being so thoughtful to me and ensuring that I had everything I needed.

I also want to thank my advisors from Argentina, Pablo, and Lucas, who have placed their trust in me and made the thesis under joint supervision possible. A very special thanks to Eduardo because from the very first moment, he has motivated me to embark on this research adventure and helped me with each procedure.

I sincerely appreciate the comments and discussion of the reviewers Alexander Sinyuk and Julian Gröbner together with Africa Barreto, Philippe Dubuisson and Walter Legnani.

D'autre part, je veux remercier les personnes du laboratoire qui ont toujours été présentes pour m'aider dès le premier jour de mon arrivée. Un merci particulier à Anne, Marie-Lyse et Christophe Van Brussel!

Special thanks to Tatsiana and Fabrice for all the advice and help I have received from them all this time, for their daily presence, for an open discussion, or just for a heart-warming smile!

Aux amis rencontrés à l'Université, Lucie, Simonne, Antoine, Fer, Chong, car sans vous cela n'aurait pas été pareil. Merci d'avoir été là quand j'ai eu besoin de vous, mais surtout merci beaucoup pour ces beaux moments que nous avons passés. En particulier, un grand merci à Lucie et Antoine, car grâce à leur patience avec moi, aujourd'hui je parle français! Y también a Fer y Simonne, porque todo este último tiempo fueron un gran sostén.

I want to thank GRASP! and here I will not say all the names, but I want to thank you because you are like a family. I remember when I arrived, Anton and Yana helped me with the whole installation. They waited for me and made sure everything was in order. Thank you both very much!! Thanks to Cheng, for always being present and

for having also helped me throughout the thesis with good pieces of advice. And I particularly want to thank David, because you have always been there, present, even from a distance, giving encouragement, advice and support throughout the work.

A mis colegas y amigos españoles, con los que he compartido no solo trabajos sino también muy lindos momentos. En particular a Dani y Jose! Gracias por todas esas charlas que hemos tenido entre tapas y por los buenos recuerdos!

A mis queridos amigos Andres, Carolina, Juanma, Consu, por todo el aguante en este tiempo. Es difícil tener un amigo haciendo un doctorado, y ustedes lo saben muy bien! Y especialmente a Caro, mi amiga de viajes y doctorado. Como hemos disfrutado estos años!! Y que buenas escapadas de fin de semana! Gracias Caro, por todos esos momentos que pasamos juntas!

Quiero agradecer a mi familia argentina del corazón (incluso si no todos son argentinos... pero casi...), a Pablo, Mati, Esteban, Pao, Santi, Angela, por hacerme sentir en casa aun estando a miles de kilómetros. Gracias por estar siempre! Y a vos Pablo, infinitas gracias por todo!

A mi familia, gracias por todo el apoyo, porque ustedes saben todo lo que me ha costado cada meta y siempre me han motivado a seguir para adelante, a ir un poco más, a no bajar los brazos. Et merci, merci beaucoup, à Romain, car tu es arrivé au moment le plus difficile et que tu as toujours été là pour me soutenir, pour m'aider en tout et pour m'encourager, tout en profitant de chaque nouvelle étape ensemble! Un grand merci!!

A mi abuela, Martha, la persona que desde muy chica me alentó a todo, me enseñó los valores que tengo, me mostró que podemos cumplir lo que queremos si trabajamos duro para ello. Ya no podré volver a verte, ni abrazarte, pero gracias abuela porque siempre me diste fuerzas y hoy soy esta persona gracias a vos!

To my parents and my grandmother Martha.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Importance of studying aerosols and cloud in the atmosphere and motivations of the thesis

Atmospheric aerosols have an important influence on the radiative balance and the climate due to their loading and their different physical, chemical and optical properties. However, according to the Intergovernmental Panel on Climate Change (IPCC) assessments (Boucher et al. (2013), Masson-Delmotte et al. (2021)), the uncertainties in understanding aerosol radiative and climate effects are high and remain among the largest contributors to the overall uncertainties in understating the climate change effects. The main limitations in aerosol characterization are related to the challenges to provide accurate monitoring at global scale of a highly complex aerosol microphysical and optical properties that have generally higher temporal and spatial heterogeneity compared to greenhouse gases.

Remote sensing is a major tool for studying the interactions of solar radiation with the atmosphere and the surface, and their influences on the Earth's radiative balance. Over the past five decades, the radiances measured by satellite sensors, from airborne instruments or from the surface (Holben et al. (1998);King et al. (1999); Takamura (2004); Nakajima et al. (2007); Li et al. (2009b); Bréon et al. (2011); Tanré et al. (2011); Kokhanovsky et al. (2015); Dubovik et al. (2019)), have been successfully used to characterize the radiative properties of the Earth, the oceans, atmospheric gases, aerosols, clouds, etc. One of the challenges in implementing remote sensing is the development of a reliable inversion algorithm required to derive information about the atmospheric or surface component interaction with the measured radiation. Therefore, in recent years, a significant progress in the development of reliable inversion methods has been made as part of overall evolution of remote sensing technique as discussed in several studies, such us Kokhanovsky and Leeuw (2009); Lenoble et al. (2013). At

the same time, there is still a significant need to further advance in several aspects of the retrieval algorithms (e.g. see Dubovik et al. (2021); Dubovik et al. (2019)). One of the most challenging and important aspect, while underdeveloped, is the evaluation of the errors in the retrieval products. For example, the review by Sayer et al. (2020) emphasises that the quality of the retrieval uncertainty estimates has not been routinely assessed in most aerosol satellite retrievals.

There are two most common approaches to provide the uncertainties of retrieved parameters. The first one is based on conducting different sensitivity tests and validations and estimating the potential effects of different possible uncertainties in the measurements or used models. For example, the estimation of the retrieval uncertainties in AERONET Version 3 (Giles et al. (2019), Sinyuk et al. (2020)) and LiRIC (Lidar and Radiometer Inversion Code, Chaikovsky et al. (2016)) algorithms are based on the analysis of the influence of the possible measurements perturbations or retrieval assumptions (mainly in measurements) on the retrieval results. However, modeling all the factors and circumstances that can affect the retrieval in all situations is theoretically impossible, and practically difficult, within the limited series of perturbed runs especially for the situations when a large number of parameters is retrieved. Indeed, a large number of factors affect the retrievals, whose variation is complex and nonlinear and addressing them adequately in a concise series of perturbation test is very challenging. In these regards, the error propagations approaches based on statistical estimation theory and described in numerous textbooks (e.g., Fourgeaud and Fuchs (1967); Edie et al. (1971); Rodgers (2000), etc.) provide asymptotically comprehensive estimates for random retrieval errors. At the same time, it should be noted that both the result of perturbation tests and statistical estimates of propagated error rely on the forward model employed may not fully represent inaccuracies related to the limitations of this model. Some additional evaluations and considerations are always desirable for accessing the adequacy of chosen forward model and its potential limitations.

In this respect, the studies in the scope of thesis are focused on the analysis and evaluation of the approach realized for the estimating of the errors of the retrieved parameters in framework of the GRASP (Generalized Retrieval of Atmosphere and Surface Properties, Dubovik et al. (2011, 2014, 2021)). The approach employed in GRASP relies on rigorously realized concept of statistical estimations and designed to account for the propagation of both random and systematic errors. Therefore, in this thesis the performance of the GRASP error estimates for aerosol parameters retrieved from mainly ground-based observations is evaluated. Moreover, GRASP generates the full covariance matrices that are used to provide error bars for retrieved parameters and also an interesting inside for understanding retrieval tendencies. Therefore, this study evaluates not only the quantitative reliability of the obtained standard deviations of retrieval error (provided by diagonal elements of covariance matrices) but it also

provides an analysis of the structure of the correlation coefficients of covariance matrices which often reveal important retrieval tendencies that can be used for optimization of the algorithms.

Finally, it is important to mention that, in a contrast with common practices when the algorithms are mainly designed for specific missions (Kokhanovsky and Leeuw (2009); Lenoble et al. (2013); Dubovik et al. (2019)), GRASP is highly versatile algorithms based on several generalized principles (e.g., see Dubovik et al. (2021)). Therefore, most of the results obtained in current study provide useful information for methodological and practical point of view of wide diversity of remote sensing applications.

1.2 Objectives and thesis outline

The goal of this PhD aims to realize in depth analysis the evaluation and optimization of the retrieval error estimation approach which has been implemented in the GRASP algorithm with the final objective of providing a solid concept for generation of dynamic estimates of uncertainties in practical remote sensing applications.

This PhD thesis includes three logical complementary parts. The first part of the study provides the detailed formal mathematical description and a depth analysis of the methodology employed for calculations of error estimates. This formal description is focused on discussing the calculation of the full covariance matrix, as the base step in the generations of the error estimates with the particular emphasis on the evaluation of the diagonal elements and the analysis of correlation coefficients structure. Indeed, the diagonal elements of this matrix include the variances of uncertainties that are commonly used in practice to determine the standard deviations and the error bars for the retrieved parameters. At the same time, the non-diagonal elements of the covariance matrix reveal the structure of correlation coefficients between the different parameters. The analysis of these non-diagonal elements, which will be demonstrated in this work, is very useful for identifying unobvious retrieval tendencies. This can be useful for optimizing observation schemes and retrieval setups in remote sensing applications.

The second part of the study is devoted to the evaluation of the developed error estimates. With that purpose we carry out a series of numerical experiments in which synthetic observations are inverted after perturbing the data by adding diverse random and systematic errors. The resulting deviations in the solution are compared to the theoretical error estimates produced by GRASP. The synthetic observations in those tests are generated using the aerosol models from the climatological analysis by Dubovik et al. (2002b). These models seem to represent the optical properties of main

aerosol types and they have been often used in previous sensitivity studies, for example it was used for analyzing different aspects of aerosol retrievals from sunphotometer measurements (Torres et al. (2014, 2017)) and by Lopatin et al. (2013, 2021) from sunphotometer and lidar measurements. Therefore, using these models enables us to revisit some previous results and align our analysis with known conclusions. Particularly, in this thesis, major efforts have been dedicated to analyze the cases of retrievals including measurements coming from sun-photometer only and also in combination with lidar. A deep description of the associated synthetic tests and main results are shown in chapter 4. The methodology developed for ground-based measurements has been applied in chapter 6 for a set of preliminary tests including satellite data.

The third part of the thesis is focused on the application of the developed error estimate approach in the real measurements. Specifically, we generate and validate the dynamic error estimates provided by GRASP for the different aerosol properties retrieved from real-data collected by ground-based radiometers and lidars. The main results are presented and analyzed in chapter 5. Finally, and again as preliminary results, the application of the developed error estimate approach is also presented for real data coming from satellite instruments, in particular, for data of the multi-angular polarimeter PARASOL.

1.3 Structure of the thesis

This PhD thesis is structured in seven chapters:

- This first introduction chapter provides an overview about the importance of studying the properties of atmospheric aerosol and explains the main motivations to carry out this work, outlines the objectives of the thesis and describes the layout of the manuscript.
- The second chapter discusses the microphysical and optical properties and ground-based and satellite remote sensing approaches used for aerosol characterization.
- The third chapter will provide a detailed description of theoretical formalism used in the GRASP algorithm to estimates retrieval errors. The description overviews the principles of retrieval concept used in GRASP and focuses on calculations of full error covariance matrices of the retrieved parameters as well as the characteristics that are not directly retrieved, but can be calculated based on the retrieved parameters. The chapter also introduces GRASP and describes its two main modules: the forward model and the numerical inversion.
- The fourth chapter describes the evaluation of the error estimates through different numerical tests using synthetic ground-based observations. At the beginning,

this chapter presents the concept of the numerical test and the methodology employed for evaluation of the accuracy of the estimates produced for the diagonal elements of covariance matrices. Specifically, it discusses the selection of the observations for tests, the assumptions used for the generation of the synthetic observations, the random and systematic errors, etc. Then, the chapter illustrates and discusses the results of the evaluation of error estimates for all the numerical tests which will include random and systematic errors. The effects of random and systematic errors are analyzed for all considered situations. Finally, based on the test results, an universal estimate accounting for both effects of random and systematic uncertainties is recommended to use in practice. Additionally, the chapter also includes the analysis of the error correlations structure using the estimates of the non-diagonal elements of the covariance matrix.

- The fifth chapter demonstrates the application of the developed error estimate methodology in practical applications. The GRASP code is used to retrieve aerosol properties from the synergetic co-incident real observations of different aerosols over Buenos Aires (Argentina) by lidar and sun/sky photometer. The error estimates for the different retrieved parameters are illustrated and discussed.

- The sixth chapter presents the preliminary results of using GRASP error estimates to characterize aerosol properties retrieved from satellite observations. The aerosol and surface retrievals from POLDER/PARASOL observations over Mongu and Banizoumbou AEROENT stations are specifically considered. First, the GRASP generated errors are analyzed using synthetic observations. The performance of GRASP error estimates in the retrieved aerosol and surface parameters including AOD, SSA and BRDF were evaluated for the full 2008 year of real POLDER/PARASOL observations.

- The last chapter contains the general conclusions and discusses the perspectives of the overall study.

Chapter 2

Remote Sensing and Aerosols

The study and understanding of atmospheric aerosols is of great importance due to different factors. For instance, high aerosol concentrations produce a negative impact on human health (Lelieveld et al. (2015)), on agricultural activity, as well as air traffic (Kienle et al. (1990)) and tourism. Moreover, aerosols play an important role in the atmosphere, thus one of the main reasons for aerosols interest is their influence on the Earth's climate.

This chapter provides a description of the atmospheric aerosols and some remote sensing instruments used to monitor the atmosphere. The first part of the chapter provides an overview of the types, sources, and effects of aerosols. It also includes an introduction concerning the bases of light-matter interactions (scattering, absorption) which allows to obtain information about optical and microphysical aerosol properties. The second part of this chapter reviews the main instruments used along this thesis and their physical principles: sun-photometer, lidar and satellite.

2.1 Aerosol Properties

Aerosols are the solid and liquid particles suspended in the atmosphere (Seinfeld and Pandis (2006)). They are a very complex component of the atmosphere whose influence depends not only on their loading, but also on their different physical, chemical and optical properties. Aerosols can be classified according to different criteria. For example, from their origin, they can be classified as natural (windborne dust, sea salt particles, volcanic emissions, biomass burning, etc) or anthropogenic (industrial emissions, transportation sources, etc). According to the mechanisms of their formation, aerosol particles are classified as primary, those which are emitted or injected directly into the atmosphere, e.g. from the volcanoes, oceans, arid regions; and secondary, those are formed through chemical reactions of gases or compounds present in the atmosphere.

On the other hand, the shape of aerosol particles is quite variable and may be

irregular as shown at the top of Figure 2.1; but also their size distribution presents high variability. Actually, aerosol size is also used as a property for aerosol classification, since aerosols can be found from a few nanometers to several hundreds of micrometers in the atmosphere. The aerosol size distributions are characterized by different size modes (Fig. 2.1), considering the sources and their different transformation:

- Nuclei mode, whose particle diameter ranges from a few nanometers to 0.1 μm. It includes combustion particles emitted directly into the atmosphere and particles formed in the atmosphere by gas-to-particle conversion.

- Accumulation mode, which ranges between 0.1 μm and $2 - 2.5$ μm, consists of small particles that coagulate too slowly to reach the coarse mode. They are combustion particles, smoke particles and coagulated nuclei mode particles.

- Coarse mode, whose particle diameter ranging from $2 - 2.5$ μm. These particles consist of windblown dust, large sea salt particles from sea spray and mechanically generated anthropogenic particles such as those from agriculture and surface mining. Because of their large size, the coarse particles readily settle out or impact on surface, so their lifetime in the atmosphere is short (`http://aerosol.ees.ufl.edu/atmos_aerosol/section04-1.html`)

Atmospheric aerosols play an important role in the Earth-atmosphere radiative budget, affecting the climate through three well-known mechanisms splitted into three different categories (King et al. (1999)). The first one is the direct radiative forcing, whereby aerosols scatter and absorb the radiation altering the radiative balance of the Earth-atmosphere system or, equivalently, the planetary albedo (Haywood and Boucher (2000)). The second is the indirect radiative forcing, by which aerosols modify the microphysical and hence the radiative properties and lifetime of clouds. This effect could be separated in two groups: i) Twomey effect (Twomey (1975, 1977)), in which the aerosol influences cloud formation by providing additional nuclei for droplet of ice crystals growth (Boucher (1999)); and ii) Albrecht effect (Albrecht (1989)), in which the aerosol effects change cloud lifetime and other cloud properties like liquid water content and cloud top height. Finally, the indirect effects of aerosols on heterogeneous atmospheric chemistry is characterized by aerosol particles that could modify the atmospheric temperature profile by absorbing aerosols, and then affecting the presence of clouds. Detailed description of these definitions about aerosol climate effects can be found in Haywood and Boucher (2000).

As a result of these processes and due to the large spatial and temporal variability of aerosols in the atmosphere, aerosols constitute a key component of climate change and a challenge to quantify, not only locally, but also in the global and annual mean. Its

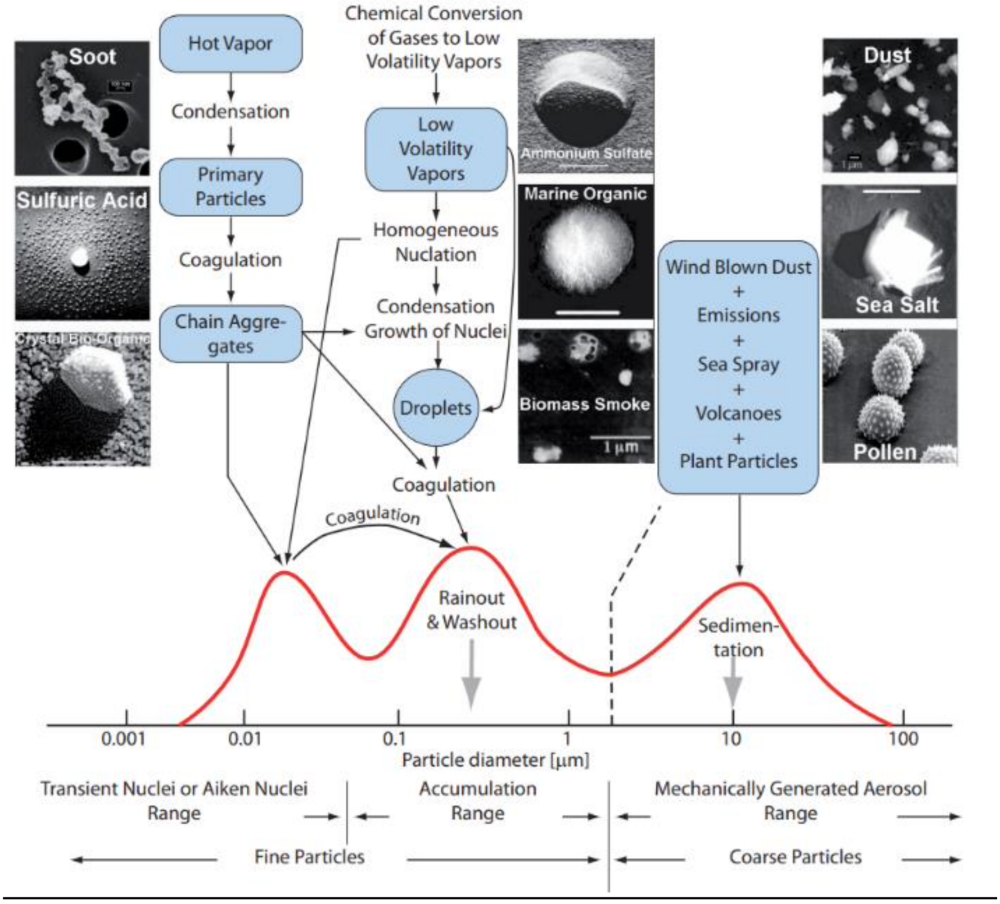


Figure 2.1: Schematic graphic of size ranges of aerosols, given in function of their particle diameter [μm], with their sources and particle formation and removal mechanisms from Zieger (2011).

effects continue to be the factor that most contributes to generating uncertainty in the climate radiative forcing according to the evaluations of the Intergovernmental Panel on Climate Change (IPCC) (Boucher et al. (2013), Masson-Delmotte et al. (2021)). Recent studies indicate that the effective radiative forcing of the aerosol is between -2.0 and $-0.4\ W/m^2$ with a likelihood of 90% (Bellouin et al. (2020)). This uncertainty is still high, and therefore, the complete understanding of climate change would require a better assessment of aerosol-cloud-radiation interactions. To this end, an improvement of global observations from space and Earth in addition to better radiative models are still needed.

Beyond the aerosol classification briefly presented in this introduction, the main aerosol properties will be described in the next subsections since they are essential to understand the aerosol role in the atmosphere. These properties can be separated in two main groups: optical properties, which characterize the interaction particle-radiation, and microphysical properties, which describe the morphology of the aerosols such as size and shape.

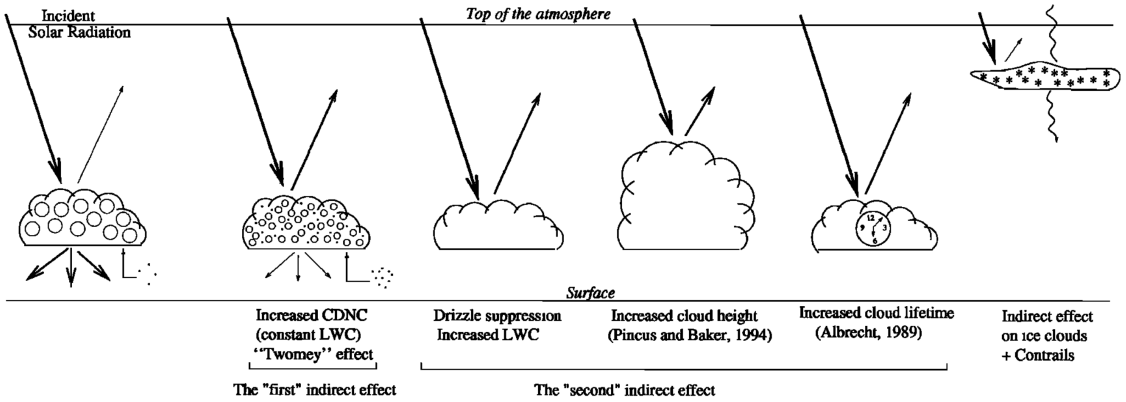


Figure 2.2: Schematic diagram with the different radiative mechanisms associated with cloud effects taken from Haywood and Boucher (2000).

2.1.1 Microphysical Properties

As mentioned above, a description of aerosol morphology is possible from their shape and size. In particular, atmospheric aerosols have shapes that are highly variable and irregular. Few aerosol examples with quite different shapes are shown in Figure 2.3. A brief description about the implementations used in GRASP algorithm will be provided in the Chapter 3.

The size of the aerosol is also highly variable from different aerosol types. In Section 2.1, we have described the different size ranges and an overview of the processes that generate particles in those ranges. It should be noted here that in normal conditions, the aerosols we found in the atmospheric column are polydisperse; that is, they present a wide range of sizes and may have quite different origins. To represent this variability as a mathematical expression (mainly for the purpose of modeling) the so-called size distribution is defined as:

$$n(r) = \frac{dN(r)}{dr} \tag{2.1}$$

where $N(r)$ represents the number of particles in the atmospheric column whose radius (or average radius for non-spherical particles) are between r and $r + dr$. In the literature, the numerical size distribution is normally found in its logarithmic representation $n(lnr) = \frac{dN(r)}{dr}$.

Early aerosol studies showed that aerosol size distribution could be approximated by the so-called Junge power function distribution (Junge (1955)). More recent analysis showed that lognormal functions (Deshler et al. (1993); Jäger and Hofmann (1991)) could well characterize most of the observed real size distributions. The number of particles $n(r)$ in a lognormal distribution is represented as:

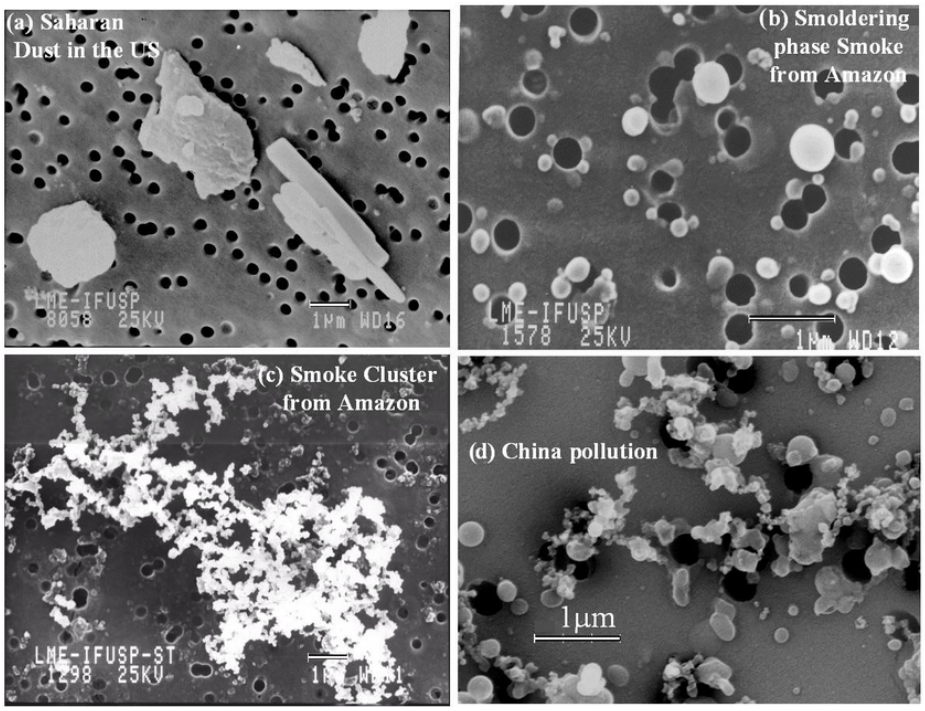


Figure 2.3: Scanning electron microscope image of ambient air aerosols (extracted from Benavent-Oltra (2019)).

$$n(r) = \frac{dN}{d\ln(r)} = \frac{N_0}{\sqrt{2\pi}} \frac{1}{\ln \sigma_0} exp\left(-\frac{(\ln(r) - \ln(r_0))^2}{2\ln^2 \sigma_0}\right) \tag{2.2}$$

where r_0 is the mode radius, σ_0 is the standard deviation of the natural logarithm of the radius and N_0 is the total number of particles.

The more recent works have shown that the usage of both the volume of the particle (instead of number) and logarithmic scale in binning of the size distribution - $dV(r)/d\ln r$ - helps to optimize the aerosol retrieval (e.g. Nakajima et al. (1996); Dubovik and King (2000); Dubovik et al. (2011)). The volume distribution is chosen since the light scattering by an ensemble of small particles depends on the particle surface or volume rather on the number concentration. At the same time, the cross sections show much smoother variability for equal relative steps $\Delta r/r$ (i.e. for equal logarithmic steps, $dr/r = d\ln(r)$) than for equal absolute steps.

The volume size distribution of a lognormal distribution can be easy derived from equation 2.2, and it takes the form:

$$v(r) = \frac{dV}{d\ln(r)} = \frac{V_0}{\sqrt{2\pi}} \frac{1}{\ln \sigma_v} exp\left(-\frac{(\ln(r) - \ln(r_v))^2}{2\ln^2 \sigma_v}\right) \tag{2.3}$$

where V_0 is the total aerosol volume per unit volume of the atmosphere, r_v is the mode radius for the volume size distribution and σ_v is the geometric standard deviation.

More often the size distribution is defined as a sum of two components, bimodal size distribution. Thus, from the mentioned above definitions it is also possible to represent the multi-mode distribution by a sum of several lognormal distributions.

2.1.2 Optical Properties

When solar radiation enters the atmosphere, there is a process of absorption and scattering due to the different atmospheric components that are suspended in it, such as aerosol particles, water vapor and other atmospheric gases in greater or lesser percentages (ozone molecules, CO2, nitrogen oxides, etc.). Then, we can distinguish two components in the radiation that reaches the earth's surface, the direct component and the diffuse component (the sum of both is called global radiation). The first follows the direction of the sun, and only undergoes the process of attenuation or extinction in that direction, and the diffuse one that comes from the scattering process of the radiation with the components of the atmosphere and arrives from all directions in space.

While absorption and scattering produced by gases are in general pretty well described (Rayleigh scattering and absorption gases bands are well described and characterized), there still remain some uncertainties concerning aerosol effects, as commented in the introduction. To reduce these uncertainties, a better characterization of aerosol optical properties would be needed since they are the main parameters which describe aerosol absorption and scattering.

Note that aerosol absorption and scattering process are not completely independent and they vary according to the optical and also microphysical characteristics of the suspended particles. Both effects are mainly characterized by the the so-called scattering efficiency factor (Q_s) and absorption efficiency factor (Q_a). In the case of spherical particles[1], both parameters are obtained directly from Mie's theory and they depend on the radius r of the particle, the refractive index m[2] and the wavelength λ of the incident beam. The sum of both gives the extinction efficiency factor:

$$Q_e(r, m, \lambda) = Q_a(r, m, \lambda) + Q_s(r, m, \lambda) \tag{2.4}$$

Then, the extinction, absorption and scattering cross section of a given particle are represented as follows:

[1]For the case of non-spherical particles, Mie's theory can not be directly applied and other techniques such as T-Matrix are use to derive the corresponding cross sections (see for instance Mishchenko et al. (2007))

[2]The complex refractive index is defined as $m = n + ik$, where, the real part n defines the speed propagation of the electromagnetic wave in the medium, and the imaginary part k is related to the aerosol absorption ability. Thus, the real refractive index can take values between 1.35 to 1.6 for the visible and near-infrared, whereas the imaginary refractive index varies from values nearly to 0 to 0.1 (Kovalev and Eichinger (2004)).

$$\begin{aligned} C_e^{(a)}(r,m,\lambda) &= \pi r^2 Q_e(r,m,\lambda) \\ C_a^{(a)}(r,m,\lambda) &= \pi r^2 Q_e(r,m,\lambda) \\ C_s^{(a)}(r,m,\lambda) &= \pi r^2 Q_s(r,m,\lambda) \end{aligned} \tag{2.5}$$

where superscript (a) denotes aerosol.

For N particles of the same refractive index with n(r) size distribution the extinction, absorption and scattering cross sections (or also named directly as aerosol extinction, absorption and scattering coefficients) are defined as follows:

$$\begin{aligned} \sigma_e^{(a)}(m,\lambda) &= \int_0^\infty \pi r^2 Q_e(r,m,\lambda) n(r) dr \\ \sigma_a^{(a)}(m,\lambda) &= \int_0^\infty \pi r^2 Q_e(r,m,\lambda) n(r) dr \\ \sigma_s^{(a)}(m,\lambda) &= \int_0^\infty \pi r^2 Q_s(r,m,\lambda) n(r) dr \end{aligned} \tag{2.6}$$

The relation between scattering and extinction coefficient is called the single scattering albedo (SSA),

$$\omega_0 = \frac{\sigma_s^{(a)}}{\sigma_e^{(a)}} = \frac{\sigma_s^{(a)}}{\sigma_a^{(a)} + \sigma_s^{(a)}} \tag{2.7}$$

The single scattering albedo is a key parameter for the estimation of the direct radiative impact of the aerosols since it is a measure of the fraction of aerosol total light extinction due to scattering and provides information about the absorption properties of the aerosols. Therefore, it is a very important parameter to quantify the impact of aerosols on climate. Purely scattering particles (e.g. sulphates, sea salt) exhibit values ~ 1, while very strong absorbers (e.g. black carbon) can have values ~ 0.2 (Dubovik et al. (2002b); Schnaiter et al. (2003)).

On the other hand, we can define the aerosol optical thickness of an atmospheric layer between the heights z_1 and z_2 as:

$$\tau_{layer}^{(a)}(\lambda) = \int_{z_1}^{z_2} \sigma_e^{(a)}(z,\lambda) dz \tag{2.8}$$

with $\sigma_e(z,\lambda)$ the value of the extinction coefficient at height z. If we refer to the entire column atmospheric, then, the total aerosol optical depth, or just aerosol optical depth (AOD), can be defined as the integration of the aerosol extinction coefficient of the layers from the ground to the top of the atmosphere (TOA):

$$\tau^{(a)}(\lambda) = \tau_a(\lambda) = \int_0^{TOA} \sigma_e^{(a)}(z,\lambda) dz \tag{2.9}$$

A parameter relating the spectral AOD to the particle size distribution, is the Angstrom exponent:

$$\alpha = \frac{\ln \frac{\tau_a(\lambda_1)}{\tau_a(\lambda_2)}}{\ln(\frac{\lambda_1}{\lambda_2})} \tag{2.10}$$

calculated typically between $440\,nm$ and $870\,nm$. It is a measure of differences of AOD at different wavelengths. It tends to be inversely dependent on particle size; higher values are generally associated with smaller aerosol particles. For example, values larger than 1.5 indicate presence of fine particles, e.g. smoke particles and sulphate, and smaller values (tends to 0) indicate the presence of coarse particles, e.g., desert dust.

The radiative effect of the rest of atmospheric components can be also characterized through their extinction, absorption and scattering cross sections. Thus, the relationship between the energy of the radiation field before and after the beam passes through the atmosphere, taking into account the conservation of energy, can be described by the total extinction coefficient ($\sigma_e^{(total)}$ or just σ_e). In this way, the variation in the intensity of the radiation field dI for the atmosphere, after a light beam passed through the medium without considering emissivity, is given by:

$$dI = -\sigma_e m_0 I dz \tag{2.11}$$

where dz is the infinitesimal thickness of the horizontal layer, I the intensity $[Wm^{-2}]$ of the light beam and m_0 is the optical mass of the incident beam which in the plane parallel approximation can be considered as $1/\cos\theta$, where θ is the zenith angle corresponding to the given path if $\theta < 75°$ (exact formulation for $\theta > 75°$ can be found in Kasten and Young (1989)).

Solving this last equation, it is observed that this interaction is represented as an exponential extinction or loss of energy as light propagates through the atmosphere according to Beer-Bouguer-Lambert law:

$$I(z,\lambda) = I_0 e^{-\int_{z_1}^{z_2} \sigma_e m_0 dz} \tag{2.12}$$

where I_0 is the intensity in $z = z_1$, I is the intensity in $z = z_2$, λ is the wavelength.

Similarly as we did it only for aerosols, we can define the total optical thickness as follows

$$\tau(\lambda) = \int_{z_1}^{z_2} \sigma_e(z,\lambda) dz \tag{2.13}$$

And then we obtain that:

$$I(z,\lambda) = I_0 e^{-\tau m_0} \tag{2.14}$$

As we has aforementioned, the optical depth is affected by the contribution of aerosols, molecules, gases. Thus, the corresponding values of aerosol optical depth under cloud free conditions can be obtained by subtracting the contribution due to Rayleigh scattering (τ_{ray}) and gas absorption contribution (τ_{abs}), where:

$$\tau_a = \tau - \tau_{ray} - \tau_{abs} \tag{2.15}$$

Finally, note the scattering process modifies the state of polarization of the radiation incident on a molecule or a particle. Thus, the amount of monochromatic radiative power scattered by an elementary volume dV of the scattering medium into a solid angle $d\Omega$ around the direction Θ is given by:

$$d\tilde{I} = \sigma_s \mathbf{F}(\Theta)\mathbf{I} dV d\Omega/4\pi \tag{2.16}$$

where, $\mathbf{F}(\Theta)$ is the 4x4 so-called normalized Stokes scattering matrix and it can be represented as follow:

$$\mathbf{F}(\Theta) = \begin{bmatrix} \mathbf{F}_{11}(\Theta) & \mathbf{F}_{12}(\Theta) & 0 & 0 \\ \mathbf{F}_{12}(\Theta) & \mathbf{F}_{22}(\Theta) & 0 & 0 \\ 0 & 0 & \mathbf{F}_{33}(\Theta) & \mathbf{F}_{34}(\Theta) \\ 0 & 0 & \mathbf{F}_{34}(\Theta) & \mathbf{F}_{44}(\Theta) \end{bmatrix} \tag{2.17}$$

the element $\mathbf{F}_{11}(\Theta) = P_{11}(\Theta)$ is the well known phase function and expresses the probability to have the scattered radiation to be distributed in any direction. This is the simplified structure for i) an ensemble of randomly oriented particles each of which has a plane of symmetry; ii) an ensemble containing an equal number of particles in random orientation and iii) a group of randomly oriented particles with an equal number of mirror particles (Lenoble et al. (2013)).

Note that $\mathbf{I}$ is the so called Stokes vector defined as:

$$\mathbf{I} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \frac{\varepsilon_0 c}{2} \begin{bmatrix} E_l E_l^* + E_r E_r^* \\ E_l E_l^* - E_r E_r^* \\ E_l E_r^* + E_r E_l^* \\ iE_l E_r^* - iE_r E_l^* \end{bmatrix} \tag{2.18}$$

where ε_0 is the dielectric constant of vacuum, c is the speed of light in the vacuum, the asterisk denote a complex conjugate value and l and r denote the parallel and

perpendicular components to a reference plane, respectively. Moreover, the first Stokes parameter I is the intensity; Q and U represent the linear polarization, and V describe the circular polarization.

2.2 Remote Sensing Instruments

Remote sensing techniques offer the capability to continuously and automatically monitor the atmosphere from the ground and space. The combination of these instruments is also possible to take full advantage of the synergies; while satellites remote sensing have demonstrated the potential of high spatial coverage and resolution, ground-based aerosol remote sensing has the benefit of a higher accuracy since it provides wide angular and spectral measurements of solar and sky radiation in key locations.

These instruments can also be classified as active or passive. The active remote sensing instruments emit the energy and then it is collected after interacting with the atmospheric particles. On the other hand, passive remote sensing instruments consist of a detection system of natural light (for instance, the Sun or the Moon) that is transmitted and scattered by aerosols in the atmosphere or reflected from the surface.

Remote sensing by ground-based sensors generally can provide accurate information for aerosol characterization (Dubovik and King (2000)) but they provide aerosol properties at local scale. In this regard, different networks have been deployed in the world with the aim of providing global coverage. They are generally based on using the same instruments and with established protocols for the measurements. Among the most known ground-based networks are AERONET (Aerosol Robotic Network; Holben et al. (1998)) and SKYNET (Sky Radiometer Network, Nakajima et al. (2020)) of ground-based photometric observations that provide aerosol column-integrated properties with high temporal resolution. On the other hand, some examples of lidar networks are MPLNET (Micro-Pulse Lidar NETwork, Welton et al. (2001)), EARLINET (European Aerosol Research Lidar Network, Pappalardo et al. (2014)) and LALINET (Latin American LIdar NETwork; Guerrero-Rascado et al. (2016); Antuña-Marrero et al. (2017)) that provide information about aerosol vertical distribution. Particularly, MPLNET and LALINET generally have also sun/sky photometers collocated to provide complementary information. Thus, these networks provide valuable information about aerosol properties which are commonly used to validate satellite observations (for example as in Chen et al. (2020)).

On the other hand, satellite remote sensing provide global monitoring of aerosol properties (Dubovik et al. (2019); Sayer et al. (2020)) showing improvements and rapid and extensive developments in the last decades. An overview of single-view instruments and those with multi-viewing capabilities will be described in a section below.

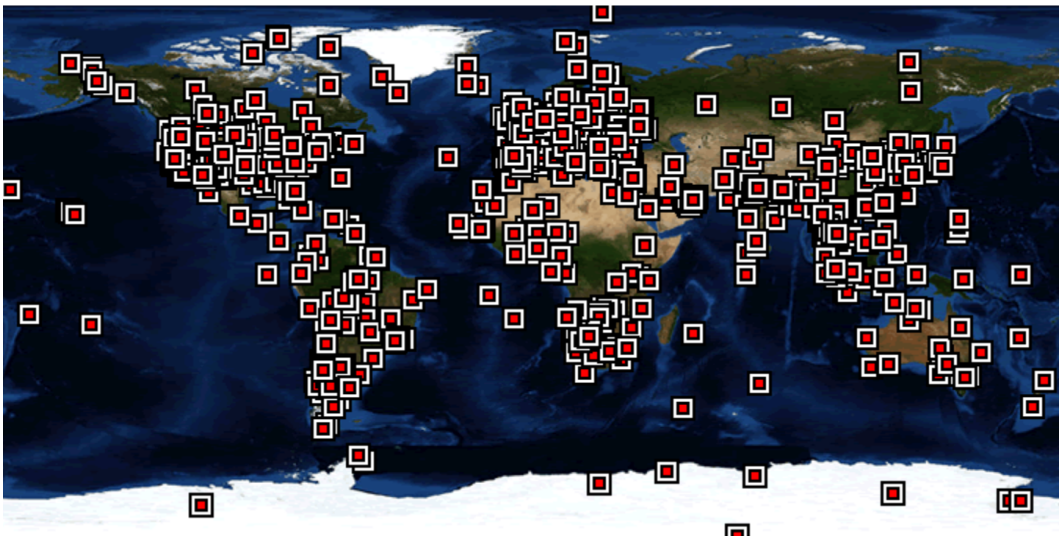


Figure 2.4: AERONET available sites in the world taken from `https://aeronet.gsfc.nasa.gov/`

The following sections provide a review about the sunphotometer, lidar and satellite observations that have been used in this work.

2.2.1 Sunphotometer/AERONET network

As mentioned above, there are different ground-based networks that characterize tropospheric aerosols. Particularly, the sun/sky photometer measurements provided by AERONET are described in this section since this thesis involves the sunphotometric observations within this network.

AERONET (AERosol RObotic NETwork, Holben et al. (1998)) is a federal network that provides measurements of columnar AOD and aerosol properties for more than 25 years in more than 600 stations distributed in the worldwide (Figure 2.4). From its beginnings in 1980, AERONET was established by NASA (Goddard Space Flight Center, GSFC) (Holben et al. (1998)), and the University of Lille (PHOtometrie pour le Traitement Operationnel de Normalisation Satellitaire - PHOTONS) (Goloub et al. (2008)), and it continues to grow thanks to the efforts and collaborations from different national agencies, institutes, universities, individual scientists, and partners.

The network relies on the standardization of its instrumentation, the specific calibration and the free available data and retrieved products (`https://aeronet.gsfc.nasa.gov/`). The standardized instrument is the Sun-sky photometer CE-318 created by Cimel Electronique (some sites include the most recent model Sun Sky Lunar CE318-T) which perform sun collimated direct beam irradiance measurements and directional sky radiance measurements at least within the spectral channels of 440, 670, 870 and 1020 nm.

Moreover, AERONET provides different data quality levels: 1.0, 1.5 and 2.0. In the previous AERONET Version 2 (operating from 2006-2019), Level 1.0 was defined as prescreened data and Level 1.5 represented near-real-time automatically cloud-screened data. After the calibration post-deployment, the data were reprocessed (assuming linear change rate in the calibration coefficients) and manually inspected, following a set of criteria (`http://aeronet.gsfc.nasa.gov/new_web/PDF/AERONETcriteria_final1.pdf`). If the data fulfill the criteria, they were raised to the Level 2.0 (quality assured data). Since 2019 the new AERONET Version 3 (Giles et al. (2019); Sinyuk et al. (2020)) is operational in the network. The main changes are: a) several modifications in the cloud screening process (Level 1.5, see differences with Version 2 in Table 2.1) and b) a fully automatic cloud screening and instrument anomaly quality controls which eliminates the need for manual quality control by an analyst and increases the timeliness of quality-assured data (`https://aeronet.gsfc.nasa.gov/cgi-bin/draw_map_display_inv_v3`).

The data are distributed through the AERONET website, with a clear data policy that must be accepted by the user before downloading data. The unique source of data ensures that the latest version of the data processing is used. It must be noted that AERONET remarks that only Level 2.0 data are quality assured for scientific research. However this data level is only available when the instrument is returned to the calibration facility after an operation period, therefore it may take months to have it available. Applications that need near-real time data may use Level 1.0 or Level 1.5 data, but those need to be handled with care.

Table 2.1: Summary of cloud-screening-related quality control changes from Version 2 to Version 3 (extracted from Giles et al. (2019)).

Algorithm/parameter	Version 2	Version 3
Very high AOD restoration	n/a	$\tau 870 > 0.5$; $\alpha 675 - 1020 > 1.2$ or $\alpha 870 - 1020 > 1.3$, restore if eliminated by cloud screening
Optical air mass range	Maximum of 5.0	Maximum of 7.0
Number of potential measurements	$N_{\text{remain}} < 3$, reject all measurements in the day	After all checks applied, reject all measurements in the day if $N_{\text{remain}} <$ MAX{3 or 10 % of N}
Triplet criterion	All wavelengths checked; AOD triplet variability > MAX{0.02 or $0.03 \cdot \tau_{\text{aerosol}}$}	AOD triplet variability > MAX{0.01 or $0.015 \cdot \tau_{\text{aerosol}}$} for 675, 870, and 1020 nm wavelengths simultaneously
Ångström exponent (AE) limitation	n/a	If $\text{AE}_{440-870\,\text{nm}} < -1.0$ or $\text{AE}_{440-870\,\text{nm}} > 3.0$, then eliminate triplet measurement.
Smoothness check	$D < 16$	For AOD 500 nm (or 440 nm) $\Delta\tau_{\text{aerosol}} > 0.01$ per minute, then remove larger τ_{aerosol} in pair. Repeat condition for each pair until points are not removed.
Solar aureole radiance curvature check (Sect. 3.2.2)	n/a	Using 1020 nm solar aureole radiances, compute the curvature (k) between 3.2 and 6.0° scattering angle (φ) at the smallest scattering angle. If $k < 2.0 \times 10^{-5}\varphi$ and if slope of curvature (M) is greater than 4.3 (empirically determined), then radiances are cloud contaminated. For sky scan measurements, all τ_{aerosol} measurements are removed within 30 min of the sky measurement. For Model T, special aureole scan measurements will remove all τ_{aerosol} within a 2 min period superseding any sky scan aureole measurements.
Stand-alone measurements	n/a	If no data exist within 1 h of a measurement, then reject it unless AE 440–870 nm > 1.0.
AOD stability check	Same as Version 3	If daily averaged AOD 500 nm (or 440 nm) has σ less than 0.015, then do not perform 3-σ check.
3-σ check	Same as Version 3	AOD 500 nm and AE 440–870 nm should be within the MEAN $\pm 3\sigma$; otherwise, the points are rejected.

n/a = not applicable.

2.2.1.1 AERONET measurements:

AERONET measurements consist of direct sun measurements and sky radiances at different wavelengths and in selected directions.

The direct solar radiation at the Earth's surface is obtained pointing the sunphotometer to the Sun. Thus, the monochromatic direct solar flux density (irradiance) at each wavelength measured is described by the Beer-Bouguer-Lambert law (as shown in the Eq. 2.14, and here integrated for the whole atmospheric column):

$$I_0(\lambda) = I_0(\lambda)e^{-m_0\tau} \qquad [Wm^2\mu m^{-1}] \tag{2.19}$$

where I_0 is the solar irradiance incident on the top of the atmosphere, m_0 is the atmospheric air mass, a function of solar zenith angle θ given by $m_0 = 1/\cos\theta$, and τ is the total optical thickness.

The direct sun measurements are performed at different wavelengths: 340, 380, 440, 500, 675, 870, 940, 1020 nm and 1640 nm which 940 nm channel is used to retrieve precipitable water. These measurements are made every $\sim$ 5 to 15 minutes. Note that 1640 nm channel is only available in the so-called extended instruments.

AERONET also provides sky radiances in different scenarios: the almucantar and principal plane geometries (Figure 2.5) (Holben et al. (1998); Kaufman et al. (2002); Olmo et al. (2008)) and newest instruments also perform hybrid scan measurements (Sinyuk et al. (2020)).

The sky radiances in the almucantar geometry are at 440, 675, 870 and 1020 nm (nominal wavelengths with 380, 500 and 1640 nm added in newer instruments). These measurements are made at optical air masses of 4, 3, 2 and 1.7 (solar zenith angle of 75, 70, 60, 54°, respectively) in the morning and in the afternoon, and once an hour in between as is explained by Sinyuk et al. (2020). This scan is performed at fixed view zenith angle equal to the SZA with a varying azimuth angle ranging from ±3.50° to ±180°, including sweeps in both directions from the Sun position. The average of the radiances in these both sides allow the elimination of contaminated data by clouds, i.e. when differences in radiances are more than 20% the data is not taken into account.

In the principal plane geometry the measurements are made at constant azimuth angle while the zenith angle varies (in the solar principal plane) and the range of scattering angles equal to the sum of the SZA and maximum view zenith angle, which is set to 75°, corresponding to a 15° elevation angle. Unlike the almucantar measurements, the principal plane geometry does not present any symmetry with respect to the solar position and, therefore, it is not easy to find a criterion to identify and detect sky inhomogeneities. Moreover, this geometry is more strongly affected by assumptions about aerosol vertical distributions than in almucantar geometry as was shown by Torres et al. (2014).

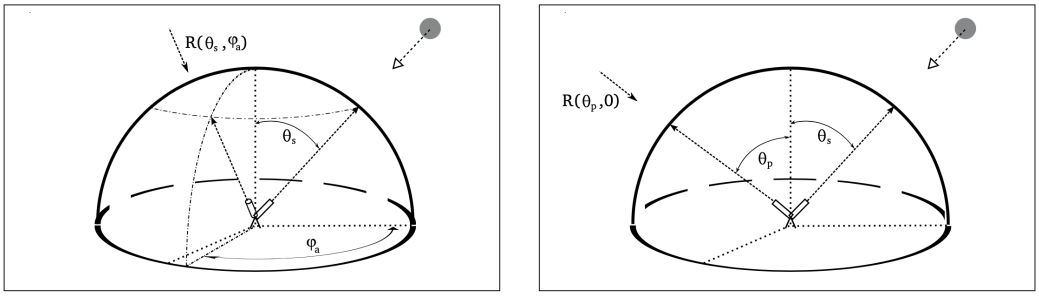


Figure 2.5: Almucantar (left) and principal plane (right) illustration geometries performed in the AERONET network for the measurements of the sky radiances (extracted from Torres et al. (2014).

An important aspect to mention is the relation between the scattering angle Θ, the solar zenith angle θ_s and the observation angles θ_a and φ_p for almucantar and principal plane, respectively. For almucantar measurements the scattering angle can be expressed as $\cos(\Theta) = \cos^2(\theta_s) + \sin^2(\theta_s)\cos(\varphi_a - \varphi_s)$ and for the principal plane is provided by $\cos(\Theta) = \cos(\theta_p \mp \theta_s)$ (Nakajima et al. (1996)). As a consequence, the maximum scattering angle that can be reached in almucantar geometries is $\Theta_M = 2\theta_s$ and in the principal plane $\Theta_M = \theta_s + 90°$. Consequently, the information contained in the radiance measurement critically depends on the geometry selected, especially for small values of the solar zenith angle (Nakajima et al. (1996); Torres et al. (2014)). In order to benefit of both geometries, the newest instruments have the capability to perform the hybrid scan measurements of directional sky radiances. It allows additional retrievals below 50° to 25° of SZA. More details can be found in Sinyuk et al. (2020).

AERONET provides from these measurements, optical and microphysical properties by the inversion algorithm (Dubovik and King (2000). Furthermore, it is important to note that some improvements are expected in AERONET data accuracy since some changes have been adapted with the new Version 3 processing algorithm (Giles et al. (2019) and Sinyuk et al. (2020)).

2.2.2 Lidar measurements

In the previous section, the ground-based sunphotometer measurementes were described which were capable of providing information of columnar aerosol properties. This section presents some general aspects of the Lidar technique and its measurements, which allows obtaining details height-resolved aerosols and cloud distributions.

Lidar is an acronym for LIght Detection And Ranging and consist in an active instrument that has the ability to provide routinary range-resolved observations of aerosols and their characteristics in the low-to-middle atmosphere from the return signal. Lidar systems operate on similar principles to those of RADAR (Radio Detection

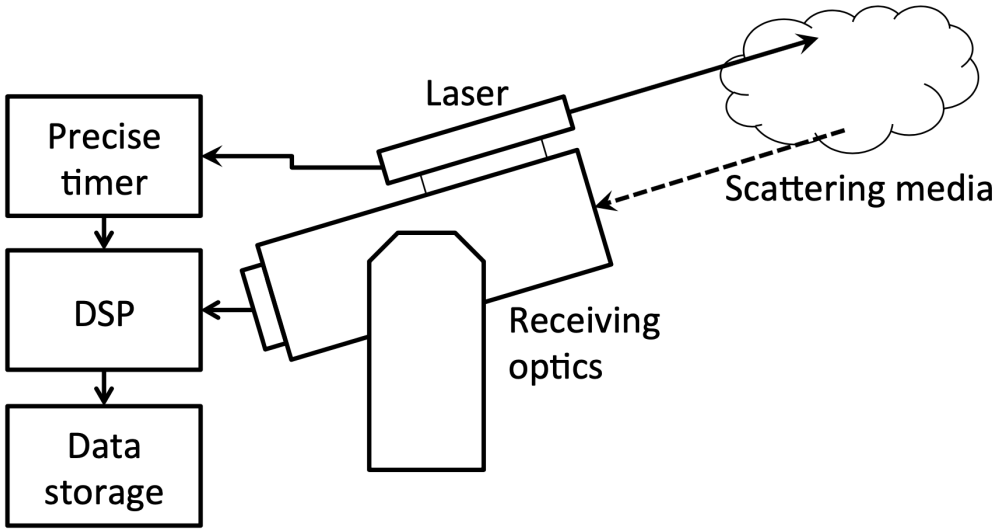


Figure 2.6: Illustration of monostatic lidar system scheme. (extracted from Lopatin (2013)).

And Ranging) but in the case of lidar, a laser pulse is emitted towards the atmosphere in the ultraviolet, visible and near or middle infrared range.

Lidar systems consist in two mainly parts: receiver and transmitter. Figure 2.6 illustrates the scheme of the lidar system. It is described from a transmission system that consists of a pulsed laser source that emits radiation into the atmosphere, a system to receive scattered radiation by the atmosphere, based on a telescope, and, finally, a spectral selector and detection system radiation received.

The emitted radiation from the beam interacts with the molecules and particles suspended in the atmosphere as explained in previous section. The radiation is partly scattered in all directions and absorbed from the atmospheric constituents. A portion of the light is scattered back toward the lidar system which is collected by a telescope and directed to a photodetector that measures the amount of backscattered light as a function of distance from the lidar. The mathematical equation that represent the scattered light collected after the interaction between the laser with the atmospheric constituents is known as lidar equation:

$$P(\lambda, h) = P_0(\lambda)\, A/h^2\, c\Delta t/2\, \eta\, O(h)\, \beta(\lambda, h)\, exp\left(-2\int_0^h \sigma_e(\lambda, h')dh'\right) \tag{2.20}$$

where P is the received backscattered power at time t, with t being the time it takes for the laser pulse to propagate at the speed of light c to a height h (with $t = 2h/c$). P_0 is the emitted laser power, A is the receiver effective area and A/h^2 is the solid angle into which photons are backscattered. Δt is the laser pulse duration and $c\Delta t/2$ is the effective pulse length, η is the efficiency system, $O(h)$ is the overlap factor that describe

the fraction of the signal covered between the transmission and reception channels, $\beta(\lambda, h) = \beta_a(\lambda, h) + \beta_g(\lambda, h)$ is the backscatter coefficient of the atmospheric layer at altitude h, and $\sigma_e(\lambda, h) = \sigma_e^{(a)}(\lambda, h) + \sigma_e^{(g)}(\lambda, h)$ is the spectral extinction coefficient, where $\sigma_e^{(a)}$, $\sigma_e^{(g)}$ and β_a, β_g are the coefficients of aerosol and molecular extinction and backscatter correspondingly. Note that equation 2.20 is valid under single-scattering approximations, i.e, when time of light interaction with the media is small as well as duration of the sounding pulse, and both these periods are smaller than time between two sequential scatter acts.

From equation 2.20, we may define the lidar constant as:

$$C = P_0 \, \eta \, c\Delta t/2 \, A \tag{2.21}$$

which describes the lidar operational capabilities, which contains the laser pulse, the qualities of the receiving optics, and any signal losses or gain over time.

An useful expression in lidar applications is the so-called range corrected signal (RCS), defined as:

$$RCS(\lambda, h) = P(\lambda, h)h^2 \tag{2.22}$$

generally used to visualize the vertical structure of the atmosphere in the measurement time.

Finally, a property widely used in the field of lidar measurements is the aerosol Lidar Ratio (LR), which is defined as:

$$LR(z) = \frac{\sigma_e^{(a)}(z)}{\beta_a(z)} = \frac{4\pi}{\omega_0 \, P_{11}(\pi)} \tag{2.23}$$

where $P_{11}(\pi)$ is the phase function at 180 degrees. LR depends on the type of aerosols and may vary with the altitude due to aerosol variability at different heights. The value of the lidar ratio may be different for the several types of aerosols (e.g. Evans (1988); Ackermann (1998); Barnaba and Gobbi (2001, 2004)). Even these values can change for the same type of aerosol according to the variation of the atmospheric properties such as, the dimension, the refractive index, shape, chemical composition and humidity content (Barnaba and Gobbi (2001)).

2.2.3 Satellite remote sensing

In addition to ground-based instruments, which provide very high-resolution information in a local scale, the study of the atmosphere and its dynamics is widely studied from remote sensing from space. In this section we introduce some generalities about satellite observations and their progress over time and a brief description of POLDER

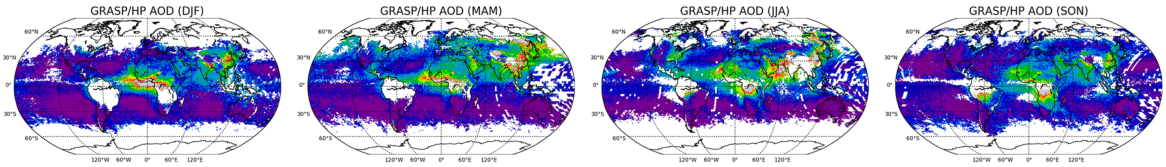


Figure 2.7: Spatial distribution of 0.1°x0.1° seasonal AOD ($550\,nm$) from PARASOL (GRASP/HP) products. DJF: December-January-February; MAM: March-April-May; JJA: June-July-August; SON: September-October-November (extracted from Chen et al. (2020)).

since they are the observations used in this work.

The use of satellites to monitor aerosols is an important tool for improving our understanding of aerosol properties as it has the advantage of providing routine measurements on a global scale. For example, Figure 2.7 shows the seasonal AOD product provided by GRASP algorithm for a global coverage from PARASOL, which was extracted from Chen et al. (2020).

Monitoring aerosols from space has been performed for over the past decades. A number of passive satellite instruments have been used to retrieve global distributions of tropospheric aerosol properties (King et al. (1999)). For example, since 1978 the AVHRR (Advanced Very High Resolution Radiometer) instrument has been operating on the NOAA POES platforms, which is one of the most used for aerosol optical thickness retrieval. TOMS (Total Ozone Mapping Spectrometer), initially launched in 1978, which has two channels sensitive to ultraviolet light, it is especially sensitive to absorbing aerosol particles, both over land and ocean (Hsu et al. (2000); Torres et al. (2002)). Moreover, GOME (Torricella et al. (1999); Carboni (2006)) was an atmospheric chemistry sensor, some of their objectives were the measurement of total column amounts, and stratospheric and tropospheric profiles of ozone on a daily basis as the investigation about the distribution of atmospheric aerosols and clouds-plus-surface spectral reflectance. Another instrument was the SeaWiFS (Sea-viewing Wide Field-of-View Sensor, Wang et al. (2000)), launched in August 1997, optimized for ocean color measurements which had 8 spectral bands from 412 to $865\,nm$. The first sensor designed for aerosol retrievals was POLDER (Polarization and Directionality of the Earth's Reflectances, Deschamps et al. (1994)) which will be further developed in more detail below.

Furthermore, on the TERRA satellite two instruments measure global aerosol concentrations and properties since 2000: MODIS (Moderate-resolution Imaging Spectroradiometer, King et al. (2003)) and MISR (Multi-angle Imaging SpectroRadiometer, Diner et al. (1998)). The first of these instruments, MODIS, consist in a 36-band spectroradiometer with moderate spatial resolution ($250 - 1000\,m$) which adopted eight channels to retrieve aerosol properties in cloud-free pixels with appropriate surface features (Martins et al. (2002); Li et al. (2005); Remer et al. (2005); Hsu et al. (2006)).

Table 2.2: Summary of major satellite measurements currently available for the aerosol products.

Properties	Sensor	Parameters	Spatial coverage	Temporal	Resolution	Reference
Loading	AVHRR	AOD	Ocean	1981-present	110km	(Husar et al., 1997; Ignatov et al., 2000; Mishchenko et al., 1999)
	TOMS		Land and ocean	1979-2001	50km	(Torres et al., 1998)
	POLDER/PARASOL			1997-present	18.5km	(Deuzé et al., 1999, 2001; Goloub et al., 1999)
	MODIS			2000-present (Terra) 2002-present (Aqua)	10km	(Hsu et al., 2004; Levy et al., 2013; Remer et al., 2002, 2005)
	MISR		Land and ocean	2000-present	17.6km	(Martonchik et al., 1998, 2002)
	OMI		Land and ocean	2005-present	13x24km	(Torres et al., 2007)
Size, shape	AVHRR	AExp	Ocean	1981-present	110km	(Higurashi et al., 1999)
	POLDER/PARASOL	FMF, AExp, non-spherical fraction	Land and ocean	1997-present	18.5km	(Deuzé et al., 2001; Herman et al., 2005)
	MODIS	FMF AExp, Effective radius	Land and ocean Ocean	2000-present (Terra) 2002-present (Aqua)	10km	(Remer et al., 2002, 2005) (Hsu et al., 2004; Levy et al., 2013)
	MISR	AExp, small, medium, large fractions, non-spherical fraction	Land and ocean	2000-present	17.6km	(Martonchik et al., 1998, 2002)
Absorption	TOMS	AI, SSA, AAOD	Land and ocean	1979-2001	50km	(Torres et al., 1998)
	OMI			2005-present	13x24km	(Torres et al., 2007)
	MISR	SSA (2-4bins)		2000-present	17.6km	(Martonchik et al., 1998, 2002)
Loading **Size, shape** **Absorption**	PARASOL (GRASP algorithm)	AOD AExp Size distribution non-spherical fraction AAOD SSA	Land and ocean	2004-2013	6km	(Dubovik et al., 2011, 2014)

Moreover, due of the simplicity of the dark ocean surface, and wide spectral range, it has the capability of retrieving AOD with a relative high accuracy of $\pm 0.03 \pm 0.05$ according to evaluation with AERONET (Chu et al. (2002); Remer et al. (2002, 2005); Levy et al. (2010)). The second one, MISR, has a wide range of along-track view angles that allow more accurate evaluation for the surface contribution to the TOA radiances. It detects the reflected light by the surface at different viewing angles along the satellite's track in a narrower spectral range ($0.44 - 0.86\,mm$). Additionally, ATSR (Along Track Scanning Radiometer) uses a mixed approach between viewing directions, as mentioned before in MISR, but also a wider spectral range ($0.55 - 1.65\,mm$) to derive the aerosol concentration and type (Veefkind et al. (2000); Grey et al. (2006)). An instrument that continues the TOMS record for total ozone and other atmospheric parameters related to ozone chemistry and climate is OMI (Ozone Monitoring Instrument, Torres et al. (2002)). It is a hyper-spectral instrument with spectral bands extending from 270 to 500 nm frequently used to characterize aerosol transport (Moulin and Chiapello (2004); Li et al. (2009a); Yu et al. (2012, 2013)). SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY, Gottwald et al. (2006)) is an spectrometer that operated in the wavelength range between 240 and 2380 nm, whose primary objective was the global measurement of various trace gases in the troposphere and stratosphere. Their large wavelength range is useful for the observation of clouds and aerosols (`https://www.sciamachy.org/`).

A summary of some of the major passive satellite measurements available for the tropospheric aerosol characterization are provided in Table 2.2.

As a result of these efforts several global climatologies are provided (Mishchenko et al. (1999); Higurashi et al. (2000); Ignatov and Stowe (2000); Torres et al. (2002); Kahn et al. (2005); Remer et al. (2005)) whose represent generally the same aerosol characteristics such as optical thickness and particle size.

It is clear that over the past decade, satellite aerosol retrievals have become increasingly sophisticated and they provide several improvements. Spaceborne instruments measure the angular dependence of radiance and polarization at multiple wavelengths from UV to the infrared (IR) at fine spatial resolution. Thus, information about particle size properties, over both ocean and land, has been provided in addition to the AOD. Furthermore, some works, such as Remer et al. (2005) and Kahn et al. (2005), show that the accuracy for AOD measurements from these sensors is about 0.05 or 20% of AOD and somewhat better over dark water. Unlike AOD, the aerosol microphysical properties (as aerosol mass types) generally present low accuracy.

2.2.3.1 POLDER/PARASOL observations

POLDER is a radiometer with wide field of view imaging that provides systematic measurements of spectral, directional and polarized properties of the solar radiation reflected by the Earth/atmosphere system (`https://www.icare.univ-lille.fr/parasol/mission/`). This instrument is on board of the PARASOL (Polarization and Anisotropy of Reflectances for Atmospheric science coupled with Observations from a Lidar) microsatellite which is part of A-Train formation.

POLDER/PARASOL observations consist of spectral information of angular distribution of both reflected total radiances, I, and polarized components, of the Stoke vector Q and U, representing the solar radiation reflected to space. These total radiance are provided in 6 windows channels (0.44, 0.49, 0.565, 0.675, 0.87 and $1.02\,\mu m$) and the polarized components in three of these channels (0.49, 0.675 and $0.87\,\mu m$). The observations are made in up to 16 viewing directions, that may cover the range of scattering angle Θ from 80 to 180 degrees. In the polarized channels, besides the total reflected radiance, I, the measurements provide the Stokes parameters Q and U referred to axes perpendicular and parallel to the local meridian plane, i.e. $Q = I_p \cos(2\alpha)$ and $U = I_p \sin(2\alpha)$ where I_p is the polarized component of reflected radiance and is the angle between the meridian plane and the polarization direction.

The set of all these observations provides a very interesting basis for a better characterization of aerosols on a global scale. Therefore, for the interpretation of this information it is necessary to use different methodologies such as look-up table based algorithms or inversion algorithms. For example, the POLDER / PARASOL retrieval algorithm based in look-up tables, use the total and polarized radiances at 670 and $865\,nm$ to retrieve total AOD (Deuzé et al. (1999)) over ocean. On the other hand,

over land (Deuzé et al. (2001)) the operational aerosol retrieval is based only on the polarized measurements at the same two wavelengths.

This methodology has provided aerosol retrievals from POLDER observations, but one of its shortcomings is that it does not use all the available information. In this regard, GRASP algorithm has been developed and adapted to use the complete set of POLDER/PARASOL observations for operational processing. It use both total radiances and linear polarization in all the spectral channels (when it is available) and it has the capability to retrieve a large number of parameters (Dubovik et al. (2011, 2014)). The retrieved products can be founded in `http://www.grasp-open.com/products/` and their evaluation and validation was provided by Chen et al. (2020).

2.3 Conclusions

In this chapter the basic theoretical concepts necessary in this thesis were described. A first part presents the different types of aerosols, their effects and the associated radiative mechanisms since they are the most variable component of the atmosphere that affect air quality and climate change. However, to understand the aerosol it is important to know not only the aerosol classifications but also their properties. This chapter has presented and described these characteristics: optical, which characterize the interaction particle-radiation, and microphysics, which describes the morphology of the aerosols such as size and shape.

Furthermore, in order to study these aerosols suspended in the atmosphere, different remote sensing techniques were presented. Ground-based measurements, like sun/sky photometer measurements and lidar which provide information in high resolution at local scale. Moreover, satellite measurements were described which provide information at global scale. All these remote sensing techniques make use of the result of the particle-radiation interaction allowing us to have valuable information to improve the understanding about the atmosphere.

Chapter 3

Inverse Modeling and Error Estimates in GRASP

One of the key challenges in implementing remote sensing is the development of the retrieval algorithms which have significantly advanced during the last decades. While a significant need led to a further advance, there are still various aspects of the retrieval algorithms to be studied. One of the most challenging and important tasks is the evaluation of the errors in the retrieval products.

Different algorithms provide aerosol properties, but in most of them the quality of the retrieval uncertainty estimates has not been routinely assessed (Sayer et al. (2020)). GRASP (Generalized Retrieved Atmosphere and Surface Properties) algorithm is one of the first algorithms to provide dynamic error estimates for the retrieved properties using several instruments based on statistical optimization approach of LSM.

Therefore, a brief description of the physical and mathematical bases of GRASP is presented in this chapter in order to get a better understanding of the algorithm. GRASP code is characterized by two main modules: forward model and numerical inversion. The chapter will mainly focus in the numerical inversion, since it provides the mathematical theory necessary for understanding error estimates.

The last part of this chapter describes the formal error propagation technique implemented in GRASP algorithm as a part of this thesis. This section begins by citing the state of the art of error estimates for aerosol properties, referencing some different methodologies developed up to now, and presenting their advantages and disadvantages.

3.1 Overview of GRASP

GRASP (Generalized Retrieval of Atmosphere and Surface Properties) algorithm is a rigorous, versatile and open-source algorithm capable of providing information of

the aerosol properties from measurements of different instruments and dynamic error estimates (Dubovik et al. (2011, 2014, 2021)). It is a flexible, generalized algorithm that relies on two independent modules: the forward model and the numerical inversion. The forward model contains the full description of the physical model including various interactions of electromagnetic solar radiations, such as aerosol scattering, surface reflectance and gaseous absorption. The multiple scattering interactions in the atmosphere are accounted by solving the vector of radiative transfer equation. Thus, GRASP forward model is capable of simulating diverse measurements in laboratory and atmosphere remote sensing including passive and active observation from space and ground. On the other hand, numerical inversion is not directly related to any physical problem and realizes formal inversion of the measurements using statistical estimation approach. Specifically, GRASP employs the Multi-term Least Square Method (LSM) that allows for a flexible utilization of multiple a priori constraints. This approach is very convenient for designing diverse remote sensing retrievals as discussed in details by Dubovik et al. (2011) and as will be described in next section.

The retrieval error estimates in GRASP are calculated by modelling propagation of measurement errors based on statistical estimation approach. In addition, the formulation used for estimating errors account for some contribution of systematic errors. These could be originated from biases in the measurement or from some modifications implemented in the algorithm for improving retrieval convergence of non-linear solutions.

In the following sections there is a description of the used forward model in GRASP, of the numerical inversion and of the overall concept and specific key implementations of the errors estimates in GRASP algorithm.

3.1.1 Forward model in GRASP

The forward model in GRASP algorithm is developed in a rather universal way and contains all the necessary elements for the simulation of the inverted remote sensing observations. It consists from several distinct modules such as multiple scattering, aerosol single scattering columnar/volume properties, aerosol vertical profile, surface reflectance and gas absorption calculations; which have been described in different works (e.g., Dubovik et al. (2011); Lopatin et al. (2013); Dubovik et al. (2021); Lopatin et al. (2021); Derimian et al. (2016); Torres et al. (2017), etc.) for diverse applications, and particularly Dubovik et al. (2021) and Lopatin et al. (2021) explain in detail. In this section, some general aspects are described about radiative transfer, modeling of aerosol properties and surface in GRASP.

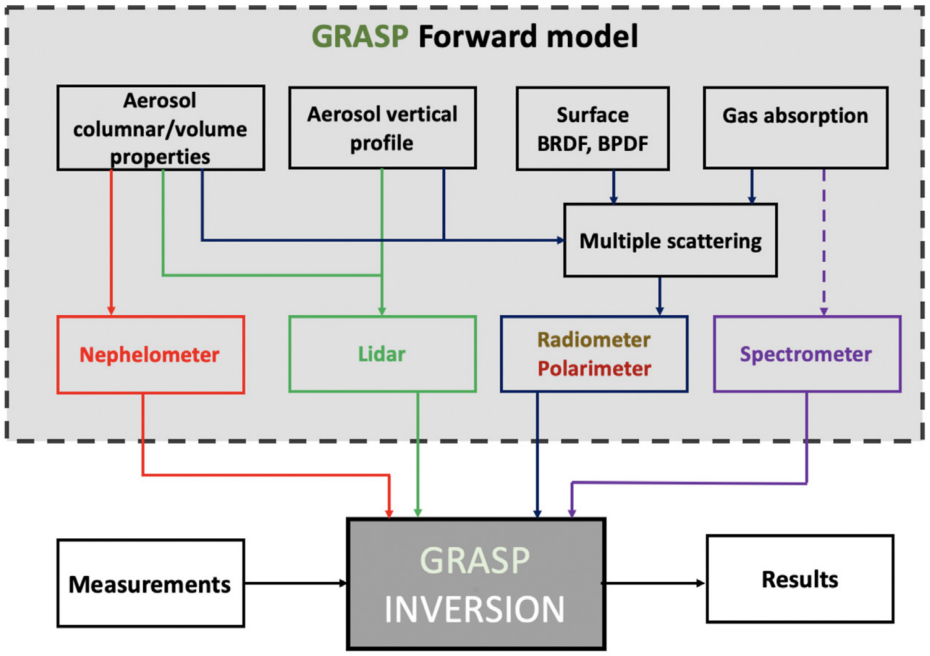


Figure 3.1: Main modules of the forward model and its connection with the numerical inversion in the GRASP algorithm (extracted from Dubovik et al. (2021)).

3.1.1.1 Radiative Transfer vector

This module is the responsible to resolve the complex problem of the interaction of radiation field with scattering-absorbing-emitting medium, where the photons are submitted to multiple scattering.

In GRASP, the multiple scattering effects are estimated by the successive order of scattering radiative transfer algorithm as is described in Lenoble et al. (2007). As was shown by Hansen (1971) the radiation properties measured by passive remote sensing exhibit negligible circular polarization of the electromagnetic field. Thus, assuming plane-parallel atmosphere the code provide full information about the radiative field including I, Q and U Stokes parameters.

This code allow the calculation of the radiance for several aerosol components which are described by defined vertical profile of spectral extinction and altitude independent phase matrix and single scattering albedo. The vertically invariant for both phase matrix and single scattering albedo are driven by: the size distribution giving the aerosol particle volume in the total atmospheric column per unit of surface area $[\mu m^3/\mu m^2]$; the complex refractive index $n(\lambda)$ and $k(\lambda)$; and the fraction of the spherical particles C_{sph}. Moreover, to account for the vertical variability of aerosol extinction is used the vertical distribution of aerosol concentration $c(h)$ as an additional normalized functional characteristic.

Details about the implementations in radiative transfer code such as the reduction of computation time without any significant loss of retrieval accuracy can be founded

in Dubovik et al. (2011) and Dubovik et al. (2021).

3.1.1.2 Aerosol Single Scattering

One of the most elaborate modules in GRASP is the aerosol single scattering. This allows modeling the optical properties of aerosols from different possibilities.

The atmospheric aerosols can be modeled as a mixture of small polydisperse particles of different shapes and composition (Dubovik et al. (2006)) for all remote sensing applications. Specifically, the optical properties of homogeneous layer of aerosol are defined by layer scattering and extinction optical thickness and by the elements of the scattering matrix $P_{ii'}(\lambda;\Theta;h)$ that can be modeled with one or several aerosol components using the microphysical properties of each component (Dubovik et al. (2021)).

For the different applications, aerosol in GRASP is usually modelled as external mixture of k aerosol components. Thus, the main equations that provide the aerosol radiative properties in a particular atmospheric layer (defined between h_{min} and h_{max}) is given by:

$$\tau_{scat/ext}(\lambda) = \sum_{k=1}^{K} \left(\int_{h_{min}}^{h_{max}} \int_{\ln\varepsilon_{min}}^{\ln\varepsilon_{max}} \int_{\ln rmin}^{\ln r_{max}} \frac{C^{(a)}_{s/e}(n_k(\lambda);k_k(\lambda);h;\varepsilon;r)}{V(r)} \frac{dV_k(h)}{dh}\frac{dN_k(\varepsilon)}{d\ln\varepsilon}\frac{dV_k(r)}{d\ln r} dh\ d\ln\varepsilon\ d\ln r \right) \tag{3.1}$$

and

$$\tau_{scat} P_{ij}(\lambda;\Theta) = \sum_{k=1}^{K} \left(\int_{h_{min}}^{h_{max}} \int_{\ln\varepsilon_{min}}^{\ln\varepsilon_{max}} \int_{\ln rmin}^{\ln r_{max}} \frac{C_{ij}(n_k(\lambda);k_k(\lambda);h;\varepsilon;r;\Theta)}{V(r)} \frac{dV_k(h)}{dh}\frac{dN_k(\varepsilon)}{d\ln\varepsilon}\frac{dV_k(r)}{d\ln r} dh\ d\ln\varepsilon\ d\ln r \right) \tag{3.2}$$

where $\tau_{scat/ext}$ is the aerosol scattering/extinction optical depth of the layer, λ denotes wavelength, Θ denotes scattering angle, ε denotes axis ratios of spheroid ($\varepsilon = a/b$, a - axis of spheroid rotational symmetry, b - axis perpendicular to the axis of spheroid rotational symmetry) and r is the radius of the equivalent sphere. $V(r)$ is the volume of particle with radius r and $C_{s/e}(n_k(\lambda);k_k(\lambda);h;\varepsilon;r)$, $C_{ij}(n_k(\lambda);k_k(\lambda);h;\varepsilon;r;\Theta)$ are cross sections of scattering/extinction and directional scattering corresponding to matrix elements $P_{ij}(\Theta)$ of an aerosol particle. Each of k_{th} aerosol components may have different size distribution $\frac{dV_k(r)}{d\ln r}$, shape distribution $\frac{dN_k(\varepsilon)}{d\ln\varepsilon}$, real refractive index $n_k(\lambda)$, imaginary refractive index $k_k(\lambda)$ and vertical profile $\frac{dV_k(h)}{dh}$.

All these characteristics that defined each aerosol component, mentioned in the equations 3.1 and 3.2, can be modeled using different approaches. We will comment these ideas in the following paragraphs

Complex refractive index: One important property obtained in different GRASP applications is the complex refractive index that can be defined from different strategies. For example, the spectral values $n_k(\lambda)$ and $k_k(\lambda)$ are retrieved for each wavelength in similar way as it is realized for AERONET (Dubovik and King (2000) and POLDER (Dubovik et al. (2011)) retrievals; another possibility is a utilization of modeling of $n(\lambda)$ and $k(\lambda)$ by assuming aerosol as a mixture of several K components mixed internally as is described by Li et al. (2019). Also, another possibility is assumed the complex refractive index and fixed it for each aerosol component.

Volume size distribution: The most general representation of size distribution is a superposition of several base functions:

$$\frac{dV_k(r)}{d\ln r} = \sum_{i=1}^{N_r} c_i^k v_i(r) \tag{3.3}$$

where $v_i(r)$ are fixed functions (so-called 'bins') and c_i^k are the weights of corresponding bins that are retrieved. These fixed functions can be represented as a superposition of triangular bins, or log-normal bins, also it can be represented as a bi-modal log-normal approximation of the size distribution or using the fixed shape of size distribution. These base functions can vary depending on the situation. For example, if we compare between AERONET retrievals or satellite retrievals, in the first case there is a rather high information content regarding the size distribution, which uses 22 bines (Dubovik and King (2000)), while in the satellite retrievals the information content of reflected radiation is lower and therefore a smaller number of bins is used (Dubovik et al. (2011)).

Vertical profile of volume concentration: The double integral concerning the volume respect the radius and the height in equations 3.1 and 3.2 should give the total volume concentration for each aerosol component. The aerosol parametrization in GRASP considers that the volume size distribution is independent on the height and, typically normalized to the total volume concentration for each aerosol component (i.e. sum of all $c_i^k v_i(r)$ defined in equation 3.3 equal to the total volume concentration for each aerosol component). Therefore, the function $\frac{dV_k(h)}{dh}$ in equation 3.1 and equation 3.2 is implicitly a normalized function:

$$\int_{h_{BOA}}^{h_{TOA}} \frac{dV_k(h)}{dh} dh = 1 \tag{3.4}$$

The function can be defined in different ways: exponential profile, Gaussian profile

or as a superposition of triangular bins. For computation purposes, it can also be discretized in layers, with the same Δ h[1], where the volume concentration in each layer is considered constant:

$$\int_{h_j}^{h_{j+1}} \frac{dV_k(h)}{dh} \sim \mathcal{C}_v^{k,j} \tag{3.5}$$

And therefore, the normalization that has been fore-mentioned in equation 3.4 takes the form:

$$\int_{h_{BOA}}^{h_{TOA}} \frac{dV_k(h)}{dh} dh \sim \Delta h \sum_{j=1}^{H} \mathcal{C}_v^{k,j} = 1 \tag{3.6}$$

where H means the total number of layers considered and superscript j refers to a certain layer. We insist in the idea that even if the concentration may change from one layer to another, the rest of aerosol properties of the component, which includes refractive indices and size distribution, remains constant.

Aerosol shape distribution: In the equations 3.1 and 3.1, the aerosols are approximated as spheroids (Mishchenko et al., 2002). As discussed by Mishchenko et al. (1997), the usage of r and ε is convenient for separating the effect of particle shape and size in analysis of aerosol mixture light scattering. Then the functions $\frac{dV(r)}{d\ln(r)}$ and $\frac{dn(\varepsilon)}{d\ln(\varepsilon)}$ denotes the volume distribution of the spheroids and the number particle shape (axis ratio) distribution accordingly. On the other hand, Dubovik et al. (2006) have demonstrated that the particle shape distribution $\left(\frac{dn(\varepsilon)}{d\ln(\varepsilon)}\right)$ for the non-spherical fraction of any tropospheric aerosol can be approximated as constant over the particle size distribution. Thus, the shape distribution is represented by two components of purely spherical particles and not spherical with an assumed shape distribution as described in detail by Dubovik et al. (2006). This assumption simplifies equation 3.1 and the aerosol extinction is calculated for the retrieval as a mixture of spherical and non-spherical fractions. Moreover, in order to perform fast and accurate calculations, and using the consideration of volume distribution in equation 3.3, and the considerations concerning the volume concentration variation with height in equation 3.6, the integrals are replaced by sums of pre-calculated kernels. The aerosol optical depth due to the k component at the layer j can be obtained as follows:

[1]The code also allows to use a logarithmic scale for the altitude. In that case, the discretization defines layers where Δ ln (h) is constant. This approach is quite interesting since we obtain a better description of lower layers when typically we have more accurate information. In order to synthesise, in this thesis we have only done the mathematical developments in natural scale and supposing that concentration are constant in layers with the same Δ h. The developments using logarithmic scale can be gained in Lopatin et al. (2013) and Dubovik et al. (2021).

$$\begin{aligned}\tau_a^{k,j}(\lambda) =&\tau_{\text{sph}}^{k,j}(\lambda) + \tau_{\text{nons}}^{k,j}(\lambda) = \sum_{i=1,\ldots,N_r} \Big(C_{\text{sph}}^k \mathbf{K}_\tau^{sph}(\lambda, k, n, r_i) \\ &+(1 - C_{\text{sph}}^k)\mathbf{K}_\tau^{nons}(\lambda, k, n, r_i)\Big)\cdot c_i^k v_i(r) \cdot \mathcal{C}_v^{k,j}\end{aligned} \tag{3.7}$$

where N_r is the number of bins used to represent the size distribution, C_{sph} is the fraction of the spherical particles and $\mathbf{K}_\tau^{sph}$ and $\mathbf{K}_\tau^{nons}$ are the kernels for spherical and non-spherical particles respectively. The total aerosol optical depth due to all the components in the whole column is obtained by the sum:

$$\tau_a = \sum_{k=1}^{K}\sum_{h=1}^{H} \tau_a^{k,j} \tag{3.8}$$

where the $\tau_a^{k,j}$ are be defined as:

$$\tau_a^{k,j} = \tau_a^k \mathcal{C}_v^{k,j} \Delta h \tag{3.9}$$

and then the optical depth of a layer would correspond to:

$$\tau_a^j = \sum_{k=1}^{K} \tau_a^k \mathcal{C}_v^{k,j} \Delta h \tag{3.10}$$

The complete information about the forward model and the detailed calculation of the kernels can be gained in Dubovik et al. (2006, 2011).

Note that so far we have intentionally done the whole description by layers instead of just limiting our developments to the estimation of aerosol properties integrated in the column. This effort has been done to show that several characteristics provided from different lidar observations can be easily modeled in GRASP. For example, the expression for the aerosol component of extinction $\sigma_{\text{a}}(\lambda)$ of a certain layer is obtained by:

$$\sigma_e^{(a)}(\lambda; h_j) = \frac{1}{\Delta h}\sum_{k=1}^{K} \tau_a^{k,j} = \sum_{k=1}^{K} \tau_a^k(\lambda)\mathcal{C}_v^{k,j} \tag{3.11}$$

where h_j refers to the average altitude of the layer j and Δh the height of the layer.

Equally the backscatter profiles $\beta_{\text{a}}(\lambda)$ can be calculated as:

$$\begin{aligned}\beta_{\mathrm{a}}(\lambda, h_j) &= \frac{1}{4\pi}\sigma_s^{(\mathrm{a})}(\lambda; h_j) P_{11}^k(\lambda; 180^\circ) \\ &= \frac{1}{4\pi\Delta h}\sum_{k=1}^{K}\tau_{sca}^{k,j}(\lambda) P_{11}^k(\lambda; 180^\circ) \\ &= \frac{1}{4\pi}\sum_{k=1}^{K}\tau_a(\lambda)^k \omega_0^k \mathcal{C}_v^{k,j} P_{11}^k(\lambda; 180^\circ)\end{aligned} \tag{3.12}$$

Moreover, the profile of the lidar ration could be obtained as:

$$S_{\mathrm{a}}(\lambda; h_j) = \frac{\sigma_e^{(a)}}{\beta_{\mathrm{a}}} = \frac{4\pi\sum\limits_{k=1}^{K}\tau_a(\lambda)^k \mathcal{C}_v^{k,j}}{\sum\limits_{k=1}^{K}\tau_a(\lambda)^k \omega_0^k \mathcal{C}_v^{k,j} P_{11}^k(\lambda; 180^\circ)} \tag{3.13}$$

3.1.1.3 Surface Reflectance

In GRASP code is also modeled the surface reflectance characteristics. A bidirectional reflectance distribution function (BRDF) and bidirectional polarization distribution function (BPDF) are used to model the effects of surface reflectance.

There are different models to calculate these characteristics. For example, the Cox and Munk model (Cox and Munk (1954)) is implemented for the reflective properties of ocean surface which estimate the Fresnel reflection on the agitated sea surface.

On the other hand, because refelctance over land can vary from one location to another, a key aspect is the correct determination of appropriate surface reflectance model and appropriate parameters. There are different models to study the effect of the directionality of land surface reflectance, for example: the semi-empirical Ross-Li model (Ross (1981); Li et al. (1992); Roujean et al. (1992); Wanner et al. (1995)), Rahman-Pinty-Vestarte (RPV) model (Rahman et al. (1993)), physically-based models for bare soil and vegetated surfaces (Litvinov et al. (2012)) as well as physically-based models for snow and ice (Kokhanovsky and Zege (2004); Kokhanovsky and Breon (2012)).

In GRASP algorithm the included BRDF and BPDF models are capable to reproduce reasonably the surface total and polarized reflectance (Maignan et al. (2004, 2009); Litvinov et al. (2011a,b)). The implemented models in GRASP have been used with different purposes, for example, for interpreting observations by MISR, MODIS, POLDER and other instruments as is shown by Justice et al. (1998), Martonchik et al. (1998), Govaerts et al. (2010) and Wagner et al. (2010).

Furthermore, for global processing of different remote sensing measurements (PARASOL, MERIS, OLCI, S5p/TROPOMI and other), GRASP has been used the combination of Ross-Li sparse BRDF model (Ross (1981); Li et al. (1992); Wanner et al. (1995)) and one parametric Maignon-Breon model (Maignan et al. (2009)). These have

allowed an optimal balance between speed, accuracy linearity and number of parameters. Nevertheless, other possible combinations of different BRDF and BPDF models are possible in GRASP algorithm and are always the subject of the studies on increasing retrieval performance.

3.2 GRASP numerical inversion

As it has been seen in Chapter 2, remote sensing measurements have been increasing over the last decades. There are several instruments and physical models capable of providing different information about, for example, the aerosol suspended in the atmosphere. Thus, numerical inversion methods provide mathematical tools that allow estimating the parameters which characterize a physical system from observed data and a priori information.

However, the theory of the inversion problem allows to estimate more than just the parameters of the model. It can be used to estimate the quality of their retrievals, to establish which parameters, or which combination of them is the best suitable (considering the information contain), as well as give the basis to provide error estimates of the retrieved parameters.

The inversion is particularly crucial and demanding for interpreting high complexity measurements where many unknowns should be derived simultaneously. Moreover, there are different important aspects of inversion optimization such as accounting for errors in the data used, inverting multi-source data with different levels of accuracy, accounting for a priori and ancillary information, estimating retrieval errors, clarifying potential of employing different mathematical inverse operations.

Thus, in the following sections are described some general aspects about numerical inversion in GRASP algorithm.

3.2.1 Introduction to numerical inversion based on statistical optimization concept

By using measurements of remote sensing instruments, commonly the number of measurements (N_f) are greater than the number of parameters to be calculated (N_a), i.e. the amount of physical information increases. This redundancy of information ($N_f > N_\mathrm{a}$) allows the use of optimization methods to find the solution to the problem, i.e. those methods that allow us to minimize uncertainties in the presence of random noise in the measurements.

The Multi-term LSM employed in GRASP searches for the solution using statistically optimized fitting under multiple a priori constraints (Dubovik (2004); Dubovik et al. (2011, 2021)). It considers both measurements and a priori data in

similar manner considering them as a data from different and independent data sources i.e.:

$$\mathbf{f}_k^* = \mathbf{f}_k(\mathbf{a}) + \Delta\mathbf{f}_k^* \tag{3.14}$$

where k denotes different data sets, $\mathbf{f}_k^*$ is a vector of the measurements, $\mathbf{f}_k(\mathbf{a})$ is a physical model, $\mathbf{a}$ is a vector of unknowns parameters and $\Delta\mathbf{f}_k^*$ is a vector of uncertainties associated to the measurements (i.e. $\Delta\mathbf{f}_k^* = \mathbf{f}_k^* - \mathbf{f}_k^{real}$). This vector of error may have two components, systematic and random according to whether they are constant between consecutive measurements, or vary randomly:

$$< \Delta\mathbf{f}_{k\,rand}^* >= \mathbf{0} \quad \text{and} < \Delta\mathbf{f}_{k\,sys}^* >= \mathbf{b}_k \tag{3.15}$$

where $\mathbf{b}_k$ is the average systematic error or so-called bias.

Thus, it is possible to apply the statistical properties of the random errors characterized by the Probability Density Function $P(\Delta\mathbf{f}_k^*)$ (PDF) to improve the solution $\hat{\mathbf{a}}$. Then, considering an adequate physical model $\mathbf{f}_{k\,real}^* = \mathbf{f}_k(\mathbf{a}^{real})$, it is possible to write:

$$\Delta\hat{\mathbf{f}}_k^* = \mathbf{f}_k^* - \mathbf{f}_k(\hat{\mathbf{a}}) \tag{3.16}$$

where $\Delta\mathbf{f}_k^* \approx \Delta\hat{\mathbf{f}}_k^*$ if $\hat{\mathbf{a}} \approx \mathbf{a}^{real}$.

If a known probability distribution is assumed, it is possible to obtain the information of the statistical properties and the best solution will be the most probable error realization, i.e. PDF maximum:

$$P(\Delta\hat{\mathbf{f}}_k^*) = P(\mathbf{f}(\hat{\mathbf{a}}) - \mathbf{f}^*) = P(\mathbf{f}(\hat{\mathbf{a}})|\mathbf{f}^*) = max \tag{3.17}$$

that is the known Method of Maximum Likelihood, which provides statistically the best solutions in many senses (Edie et al. (1971)). The approach is useful for optimum data combination, optimum use of a priori information, continuous solution space, rigorous error estimates, etc. Thus, as is explained by Dubovik et al. (2000) and Dubovik et al. (2021), focusing the features of method of maximum likelihood (MML), the solution is:

- asymptotically non-biased $< \hat{\mathbf{a}} > \rightarrow \mathbf{a}^{real}$;
- asymptotically consistent $\hat{\mathbf{a}} \rightarrow \mathbf{a}^{real}$;
- asymptotically efficient, i.e. variance of $\hat{\mathbf{a}}$ converges to the smallest possible value;
- and $\hat{\mathbf{a}}$ has asymptotically normal distribution $\hat{\mathbf{a}}_N \longrightarrow_{N\rightarrow\infty} N\left(\mathbf{a}^{real}, I_{\hat{\mathbf{a}}}^{-1}\right)$ where $I_{\hat{\mathbf{a}}}$ is a Fisher information matrix.

Equation 3.14 show the different k data set that are not correlated and may have different levels of uncertainties described by different covariance matrices $\mathbf{C}_k$. Such explicit differentiation of the input data makes the retrieval more transparent because it clearly identifies the different used data sets. Correspondingly joint probability density function (PDF) of independent data sets $\mathbf{f}_1^*, \mathbf{f}_2^*, \cdot, \mathbf{f}_K^*$ can be obtained by the simple multiplication of the PDFs of data from all K sources:

$$\begin{aligned}\mathbf{P}(\mathbf{f}(\mathbf{a})|\mathbf{f}^*) &= \mathbf{P}(\mathbf{f}_1(\mathbf{a}), \ldots, \mathbf{f}_k(\mathbf{a})|\mathbf{f}_1^*, \ldots, \mathbf{f}_k^*) \\ &= \prod_{k=1}^{K} \mathbf{P}(\mathbf{f}_k(\mathbf{a})|\mathbf{f}_k^*) \\ &\approx exp\left(-\frac{1}{2}\sum_{k=1}^{K}(\mathbf{f}_k(\mathbf{a}) - \mathbf{f}_k^*)^T\mathbf{C}_k^{-1}(\mathbf{f}_k(\mathbf{a}) - \mathbf{f}_k^*)\right)\end{aligned} \tag{3.18}$$

where $\mathbf{C}_k$ is the covariance matrix of the vector $\mathbf{f}_k^*$ and T denotes the matrix transposition.

It can be noted that Eq. 3.14 not assume any relations between forward models $\mathbf{f}_k(\mathbf{a})$, i.e. forward models $\mathbf{f}_k(\mathbf{a})$ can be the same or different. In the frame of LSM approach, i.e. under the assumptions of a normal PDF of the error $\Delta\mathbf{f}_k^*$, the solution of the Eq. 3.14 corresponds to the minimum of the following functional (i.e. maximize Eq. 3.18):

$$\mathbf{\Psi}(\mathbf{a}) = \frac{1}{2}\sum_{k=1}^{K}(\mathbf{f}_k(\mathbf{a}) - \mathbf{f}_k^*)^T\mathbf{C}_k^{-1}(\mathbf{f}_k(\mathbf{a}) - \mathbf{f}_k^*) = \min \tag{3.19}$$

For the general case, of non-linear functions $\mathbf{f}_k(\mathbf{a})$ the solution of Eq. 3.19 is sought iteratively:

$$\mathbf{a}^{p+1} = \mathbf{a}^p - \Delta\mathbf{a}^p \tag{3.20}$$

where $\Delta\mathbf{a}^p$ is the solution that can be found by solving the system of so-called normal equations, i.e.:

$$\left(\sum_{k=1}^{K}\mathbf{K}_{k,p}^T(\mathbf{C}_k)^{-1}\mathbf{K}_{k,p}\right)\Delta\mathbf{a}^p = \sum_{k=1}^{K}\mathbf{K}_{k,p}^T(\mathbf{C}_k)^{-1}\Delta\mathbf{f}_k^p \tag{3.21}$$

where $\Delta\mathbf{f}^p = \mathbf{f}(\mathbf{a}^p) - \mathbf{f}^*$, and $\mathbf{K}_p$ is Jacobean matrix at p-th iteration of the functions $\mathbf{f}_k(\mathbf{a})$ in the vicinity of $\mathbf{a}^p$ with the elements $\{\mathbf{K}_{k,p}\}_{j,i} = \left.\frac{\partial\mathbf{f}_{k,j}(\mathbf{a})}{\partial a_i}\right|_{\mathbf{a}=\mathbf{a}^p}$.

The asymptotic limit of the minimized quadratic form, for most applications, can be written as:

$$2\Psi(\mathbf{a}) = min \rightarrow \sum_{k=1}^{K} N_k - n \tag{3.22}$$

It should be noted that the LSM solution defined by Eq. 3.19 corresponds to the minimum of quadratic form $\hat{\Psi}(\mathbf{a})$ and does not depend in any way from the value of this minimum. Considering this fact, in practical application is often convenient to renormalize the minimized quadratic $\hat{\Psi}(\mathbf{a})$, in situations when only one data set is inverted it is convenient to weighting matrix $\mathbf{W} = \mathbf{C}/\varepsilon_1^2$ and minimize the quadratic form $\Psi'(\mathbf{a}) = \varepsilon_1^2\Psi(\mathbf{a})$. In such approach one does not need to know the exact value of ε_1^2. Moreover, ε_1^2 can be estimated from asymptotic LSM expectations provided by Eq. 3.22.

In frame of Multi-term approach the use of weighting matrices additionally allows for making more explicit the contribution of different data sources. Indeed, using the weighting matrices $\mathbf{W}_k$ instead of covariance matrices $\mathbf{C}_k$ the Eq. 3.21 can be written as:

$$\left(\sum_{k=1}^{K} \gamma_k \mathbf{K}_k^T (\mathbf{W}_k)^{-1} \mathbf{K}_k\right) \Delta\mathbf{a}^p = \sum_{k=1}^{K} \gamma_k \mathbf{K}_k^T (\mathbf{W}_k)^{-1} \Delta\mathbf{f}_k^p \tag{3.23}$$

In this formulation the relative contribution of the data from different data sources are scaled by the corresponding Lagrange parameters γ_i, defined as:

$$\mathbf{W}_i = \frac{1}{\varepsilon_i^2}\mathbf{C}_i \quad \text{and} \quad \gamma_i = \frac{\varepsilon_1^2}{\varepsilon_i^2} \tag{3.24}$$

where ε_i^2 is the first diagonal element of $\mathbf{C}_i$, i.e. $\varepsilon_i^2 = \mathbf{C}_{i11}$ and γ_i is the ratio of the variances of scattered radiances and variances of the corresponding data set. Evidently, that $\gamma_1 = 1$ as discussed by Dubovik and King (2000); Dubovik (2004); Dubovik et al. (2011), etc. This renormalization strategy is especially convenient on Multi-term LSM approach once some of data sets correspond to a priori information. In addition, the renormalized definition of the minimized quadratic function (or *residual*) as $\Psi'(\mathbf{a}) = \varepsilon_1^2\Psi(\mathbf{a})$, the measurement error ε_1^2 can be estimated from the residual of the fit. Indeed, once the weighting matrices used in the solution, Eq. 3.23 minimizes quadratic with the limit depending on ε_1^2:

$$2\Psi'(\mathbf{a}) = 2\,\varepsilon_1^2\,\Psi(\mathbf{a}) = min \rightarrow \varepsilon_1^2\left(\sum_{k=1}^{K} N_k - n\right) \quad \text{and} \quad \hat{\varepsilon}_1^2 \approx \frac{2\Psi'(\mathbf{a}^p)}{\sum_{k=1,\ldots,K}(N_{\mathbf{f}_i}) - N_{\mathrm{a}}} \tag{3.25}$$

The use of the weight matrix makes evident the relative contribution of the data from different data sources, it allows rather transparent interpretation of lagrange multipliers γ- parameters determining the contributions of a priori terms into solution.

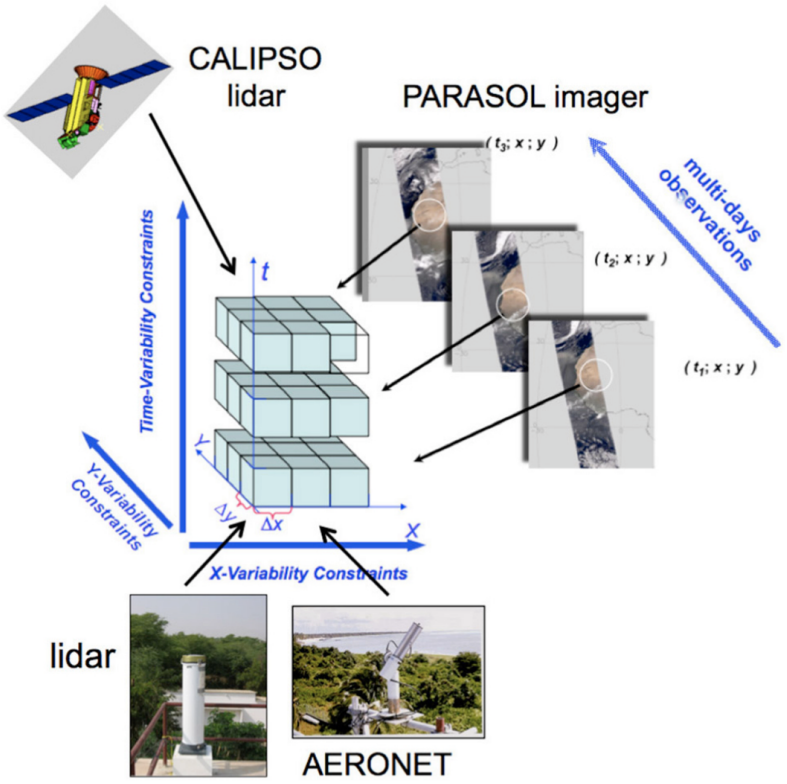


Figure 3.2: Concept of single and multi-pixel in GRASP algorithm Dubovik et al. (2014).

3.2.2 A priori constraints in Multi-term LSM approach and in GRASP algorithm

As discussed in detail by Dubovik et al. (2021) the Multi-term LSM concept has been proposed as methodologically convenient approach for integrating different types of a priori constraints in remote sensing applications (Dubovik (2004); Dubovik and King (2000); Dubovik et al. (1995, 2000, 2008, 2011)). In the Multi-term LSM a priori estimates are considered to be 'equivalent' to the measurements, i.e. characterized by their PDF and treated equivalently to the actual measurements. In these regards, Eqs. (3.14-3.21) do not show any distinction between different $\mathbf{f}_k(\mathbf{a})$. At the same time, in practice there are always two different types of data sets: measurements and a priori constraint on the unknowns $\mathbf{a}$.

In the GRASP algorithm there are two types of a priori limitations: the single pixel and multi pixel constraints (Fig. 3.2). The single pixel, is a conventional approach that considers each single pixel is inverted completely independently. On the other hand, multi-pixel retrieval is the approach that considers a group of pixels are inverted simultaneously and extra a priori constraints on the inter-pixel variability of the retrieved parameters can be applied.

3.2.2.1 Single-pixel constraints used in GRASP

In the practice there are always two different types of data sets: measurements and a priori constraint on the unknowns $\mathbf{a}$. Therefore, the vector of the measurement $(\mathbf{f}^*)^T = (\mathbf{f}_1^*, \mathbf{f}_2^*, \cdot, \mathbf{f}_k^*)^T$ can be written as:

$$(\mathbf{f}^*)^T = (\mathbf{f}_1^*, \mathbf{f}_2^*, \cdot, \mathbf{f}_k^*, \mathbf{f}_1^{\mathrm{a}}, \mathbf{f}_2^{\mathrm{a}}, \cdot, \mathbf{f}_k^{\mathrm{a}})^T \tag{3.26}$$

where $\mathbf{f}_i^* = \mathbf{f}_i^*(\mathbf{a})$ represent directly measured characteristics and $\mathbf{f}_i^{\mathrm{a}} = \mathbf{f}_i^{\mathrm{a}}(\mathbf{a})$ represent a priori known characteristics of unknowns $\mathbf{a}$. Correspondingly the right side of Eq. 3.18 can be formally split in two groups:

$$P(\mathbf{f}(\mathbf{a})|\mathbf{f}^*) = \prod_{k=1}^{K} P(\mathbf{f}_k(\mathbf{a})|\mathbf{f}_k^*) \prod_{n=1}^{N} P(\mathbf{f}_n^{\mathrm{a}}(\mathbf{a})|\mathbf{f}_n^{\mathrm{a}}) \tag{3.27}$$

Therefore, the Eq. 3.23 can also be formally arranged to identify the contribution of measurements and a priori terms:

$$\begin{aligned}\left(\sum_{k=1}^{K} \gamma_k \mathbf{K}_{k,p}^T (\mathbf{W}_k)^{-1} \mathbf{K}_{k,p} + \sum_{n=1}^{N} \gamma_{\mathrm{a},n} \mathbf{K}_{\mathrm{a},n,p}^T (\mathbf{W}_{\mathrm{a},n})^{-1} \mathbf{K}_{\mathrm{a},n,p}\right) \Delta\mathbf{a}^p = \\ = \sum_{k=1}^{K} \gamma_k \mathbf{K}_{k,p}^T (\mathbf{W}_k)^{-1} \Delta\mathbf{f}_k^{*,p} + \sum_{n=1}^{N} \gamma_{\mathrm{a},n} \mathbf{K}_{\mathrm{a},n,p}^T (\mathbf{W}_{\mathrm{a},n})^{-1} \Delta\mathbf{f}_n^{\mathrm{a}^*,p}\end{aligned} \tag{3.28}$$

where two groups of the terms in left and right parts of the equation represent the contributions of the set of K measured characteristics $\mathbf{f}_k(\mathbf{a})$ and the set of N a priori $\mathbf{f}_n^{\mathrm{a}}(\mathbf{a})$ characteristics, and the Lagrange parameters are defined as:

$$\gamma_k = \frac{\varepsilon_{k=1}^2}{\varepsilon_k^2} \quad \text{and} \gamma_{\mathrm{a},n} = \frac{\varepsilon_{n=1}^2}{\varepsilon_{a,n}^2} \tag{3.29}$$

As discussed by Dubovik (2004) and Dubovik et al. (2021) the Multi-term approach is a simple rearranging the base LSM formulation, while the resulting Eq. 3.23 provides a solid basis for unifying many known formulas of constrained inversion in a single formalism and practically convenient and efficient for developing remote sensing algorithms using diverse complimentary observations and a priori constrains.

While, the Multi-term LSM concept allows flexible utilizations of nearly arbitrary a priori constraints, GRASP algorithm is fully adapted for using the most popular and physically transparent a priori constraints such as direct a priori estimates of unknowns $\mathbf{a}$ and, smoothness constraints in situations, when the unknown vector $\mathbf{a}$ or any group of unknowns included in this vector, represent continuous smooth function. For example, if vector $\mathbf{a}$ represents aerosol size distribution, that is known to be rather smooth, the

system given by Eq. 3.14 can be explicitly written as follows:

$$\begin{cases} \mathbf{f}^{*}_{k=1} = \mathbf{f}^{*}_{k=1}(\mathbf{a}) + \Delta\mathbf{f}^{*}_{k=1} \\ \mathbf{f}^{\mathrm{a},*}_{n=1} = \mathbf{f}^{\mathrm{a},*}_{n=1}(\mathbf{a}) + \Delta\mathbf{f}^{\mathrm{a},*}_{n=1} \\ \mathbf{f}^{\mathrm{a},*}_{n=2} = \mathbf{f}^{\mathrm{a},*}_{n=2}(\mathbf{a}) + \Delta\mathbf{f}^{\mathrm{a},*}_{n=2} \end{cases} = \begin{cases} \mathbf{f}^{*}_{1} = \mathbf{f}^{*}_{k=1}(\mathbf{a}) + \Delta\mathbf{f}^{*}_{1} \\ \mathbf{f}^{\mathrm{a},*}_{1} = \mathbf{f}^{\mathrm{a},*}_{1}(\mathbf{a}) + \Delta\mathbf{f}^{\mathrm{a},*}_{1} \\ \mathbf{f}^{\mathrm{a},*}_{2} = \mathbf{f}^{\mathrm{a},*}_{2}(\mathbf{a}) + \Delta\mathbf{f}^{\mathrm{a},*}_{2} \end{cases} = \begin{cases} \mathbf{f}^{*}_{1} = \mathbf{f}^{*}_{k=1}(\mathbf{a}) + \Delta\mathbf{f}^{*}_{1} \\ \mathbf{a}^{*} = \mathbf{a} + \Delta_{\mathbf{a}^{*}} \\ \mathbf{0}^{*} = \mathbf{G}^{\mathbf{a}}_{m} + \Delta_{\mathbf{g}^{*}} \end{cases} \tag{3.30}$$

The a priori constraints defined by the second line $\mathbf{a}^{*} = \mathbf{a} + \Delta_{\mathbf{a}^{*}}$ represents the most common of constraints of solution by direct a priori estimates of unknowns $\mathbf{a}^{*}$, where $\Delta_{\mathbf{a}^{*}}$ are the uncertainties of the estimates $\mathbf{a}^{*}$ and are generally considered to be unbiased random errors within the covariance matrix $\mathbf{C}_{\mathbf{a}^{*}}$. These constraints can be easily included in Eq. 3.28 by defining: $\mathbf{K}_{\mathrm{a}} = \mathbf{1}$- unity matrix; i.e. $\mathbf{K}^{T}_{\mathrm{a}}\mathbf{W}^{-1}_{\mathrm{a}}\mathbf{K}_{\mathrm{a}} = \mathbf{W}^{-1}_{\mathrm{a}}$ and $\mathbf{K}^{T}_{\mathrm{a}}\mathbf{W}^{-1}_{\mathrm{a}}\mathbf{f}^{\mathrm{a},*}_{1} = \mathbf{W}^{-1}_{\mathrm{a}}\mathbf{a}$. Utilization of a priori estimates $\mathbf{a}^{*}$ was introduced in the pioneering studies by Twomey (1963) and later evolved and discussed in detail in the Rodgers (2000) textbook on inversion. The third line represents another common type of a priori constraint known as smoothness constraints that limit the variability of retrieved functions by using a priori knowledge about limitations on derivatives of those functions. For example, a priori knowledge limits high frequency variations of continuous functions $v(x)$, such as the aerosol size distribution. In GRASP, the smoothness constraints are related to a priori known limited values of the derivatives, i.e. with their m-th derivative deviations from zero:

$$\frac{\partial^{m} v(x)}{\partial x^{m}} \approx 0 \tag{3.31}$$

For the vector of unknowns $\mathbf{a} = (\mathrm{a}_1, \mathrm{a}_2, \cdot, \mathrm{a}_n)^{T}$ that contains discrete elements describing the continuous function $v(x)$, the knowledge on the smoothness of function $v(x)$ can be defined using a vector-matrix linear system (e.g. see Dubovik et al. (2021)): $\mathbf{0}^{*} = \mathbf{G}^{\mathbf{a}}_{m} + \Delta_{\mathbf{g}^{*}}$, where $\mathbf{G}_{m}$ is the Jacobean matrix of the matrix of the m-th derivatives. In practice, these are often approximated by matrices of the m-th finite difference estimated in point $\mathbf{a}$. The errors $\Delta_{\mathbf{g}^{*}}$ reflect the uncertainty in the knowledge of the deviations of $y(x)$ from the assumed constant ($m = 1$), straight line ($m = 2$), parabola ($m = 3$), and so on. Under assumption that the $\Delta_{\mathbf{g}^{*}}$ have a normal distribution, with the unbiased covariance matrix $\mathbf{C}_{g}$, these constraints can be easily included in Eq. 3.28 by defining: $\mathbf{K}_{\mathrm{a},2} = \mathbf{G}_{m}$ and $\mathbf{f}^{\mathrm{a},*}_{2} = \mathbf{0}^{*}$, i.e. $\mathbf{K}^{T}_{\mathrm{a}}\mathbf{W}^{-1}_{\mathrm{a}}\mathbf{K}_{\mathrm{a}} = \mathbf{G}^{T}_{m}\mathbf{W}^{-1}_{\Delta g}\mathbf{G}^{T}_{m}$ and $\mathbf{K}^{T}_{2}\mathbf{W}^{-1}_{2}\Delta\mathbf{f}^{\mathrm{a},*}_{2} = \mathbf{G}^{T}_{m}\mathbf{W}^{-1}_{\Delta g}(\mathbf{a}^{p} - \mathbf{0}^{*}) = \mathbf{G}^{T}_{m}\mathbf{W}^{-1}_{\Delta g}\mathbf{a}^{p}$. Utilization of such smoothness constraints was suggested by one of the first formulations of constrained inversion by Phillips (1962) and was also considered in article by Tikhonov (1963) and Tikhonov's later studies.

Thus, for a case where only a direct a priori estimates and smoothness constraints

are used, Eq. 3.28 can be explicitly written via weighting matrices as:

$$\left(\mathbf{K}^T\mathbf{W}_f^{-1}\mathbf{K} + \gamma_{\mathrm{a}}\mathbf{W}_{\mathrm{a}}^{-1} + \gamma_g\Omega_m\right)\Delta\mathbf{a}^p = \mathbf{K}^T\mathbf{W}_f^{-1}\Delta\mathbf{f}^p + \gamma_{\mathrm{a}}\mathbf{W}_{\mathrm{a}}^{-1}(\mathbf{a}^p - \mathbf{a}^*) + \gamma_g\mathbf{\Omega}_m\mathbf{a}^p \quad (3.32)$$

where $\mathbf{\Omega}_m$ denotes the smoothness matrix defined as:

$$\mathbf{G}_m^T\mathbf{W}_{\Delta g}^{-1}\mathbf{G}_m^T = \mathbf{\Omega}_m \quad (3.33)$$

the explicit formulation of $\mathbf{\Omega}_m$ can be found in the paper by Dubovik et al. (2011). The Eq. 3.30 generalizes the commonly used base equations of constrained inversion by Phillips (1962), Twomey (1975, 1977), Tikhonov (1963) and Rodgers (1976, 1990, 2000). It should be noted that Eq. 3.30 is written for the simplest situation when the vector $\mathbf{a}$ represents only one continues function $v(x)$, while in many GRASP applications the vector of unknowns includes several components $\mathbf{a}^T = (\mathbf{a}_{sd}^T, \mathbf{a}_{n(\lambda)}^T, \mathbf{a}_{k(\lambda)}^T, \mathbf{a}_h^T, ...)^T$, where each component is relevant to continues functions representing such physical characteristics as aerosol particle size distribution ($\mathbf{a}_{sd}$), spectral dependence of real ($\mathbf{a}_{n(\lambda)}$) and complex ($\mathbf{a}_{k(\lambda)}$) parts of refractive index, vertical distribution ($\mathbf{a}_h$), etc. Each of those characteristics is continues function and therefore in retrieval the smoothness constraints can be applied on each of corresponding component of the vector of unknowns. Evidently, direct a priori constraints can be applied to each single element of the vector $\mathbf{a}$, while from practical view point separating outlining the contribution of a priori estimates for each component, e.g. $(\mathbf{a}^*)^T = ((\mathbf{a}_{sd}^*)^T, (\mathbf{a}_{n(\lambda)}^*)^T, (\mathbf{a}_{k(\lambda)}^*)^T, (\mathbf{a}_h^*)^T, ...)^T$. Similarly, the inverted measurements may come from different sources and therefore have different levels of accuracy and different weighting matrices. As a result, in practice, all the first, second and third terms in Eq. 3.30 may have many quite different components, and therefore actual formulation of the solution can be significantly more complex. Some of explicit equations can be found in the paper by Dubovik et al. (2011).

The realization of the inversion in GRASP, in principle, is based on general Eq. 3.23, while for used convenience there is a logical separation as indicated in Eq. 3.28 into actual measurements and a priori constraints. For each measurement data set $\mathbf{f}_k^*$ two types of errors can be set: relative or absolute and the magnitude of the errors is defined by the standard deviation and a weighting matrix $\mathbf{W}_i$. The standard deviation is used inside of the code to calculate corresponding Lagrange parameters γ_i. The weighting matrix $\mathbf{W}_i$ is assumed as the unity matrix by default, while it can also be set diagonal with different values at the diagonal, as well, in more general way with non-zero non-diagonal values too. For applying the a priori constraints, as discussed above, there are two main possibilities: using direct a priori constraints or applying

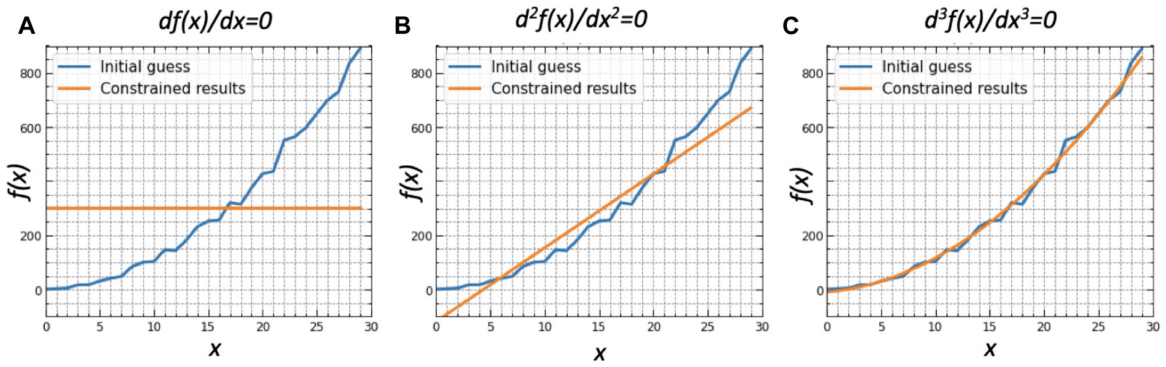


Figure 3.3: Effect of the smoothness constraints limiting the derivatives of different order ($m = 1$, $m = 2$ and $m = 3$, respectively) in the solution (Dubovik et al. (2021))

smoothness constraints for the parameters that define continues functions.

The direct a priori estimates a_i^* for each of value a_i in the vector of unknows $\mathbf{a} = (a_1, a_2, ..., a_n)^T$ can be provided with the corresponding Lagrange parameters γ_{a_i}. There is also a possibility to assume a vector $\mathbf{a}^*$ of a priori estimates for all the retrieved parameters or for selected groups (e.g. parameters describing size distribution) with common Lagrange parameter γ_a. In this case weighting matrix $\mathbf{W}_a$ is also provided that is assumed as the unity matrix by default, or can be set diagonal with different values at the diagonal or in more general way with non-zero non-diagonal values.

The smoothness a priori constraints can be applied for each group of parameters describing a continuous function (e.g., $\mathbf{a}_{sd}^T, \mathbf{a}_{n(\lambda)}^T, \mathbf{a}_{k(\lambda)}^T, \mathbf{a}_h^T, ...$, etc) by defining the order m of limited derivatives ($m = 0$ is a constant; $m = 1$ is a straight line; $m = 2$ is a parabola, etc.) and the strength of the applied a priori smoothness constraints is defined by Lagrange parameters γ_n. The smoothness matrix $\mathbf{\Omega}_m$ is defined as in Eq. 3.31 where weighting matrix $\mathbf{W}_{\Delta g}$ is unity matrix by default and be set diagonal with different values on the diagonal in case if the retrieved continues function has different level of variability for different ordinates. Figure 3.3 illustrates the effect of different order of the smoothness constraints described above.

3.2.2.2 Multi-pixel constraints used in GRASP

The multi-pixel approach is mostly used for satellite inversions since the information content in the reflected radiation from a single pixel is generally insufficient for a unique characterisation of all retrieved parameters. In this context, the use of the multi-pixel approach comes to improves the stability of the retrieval. Moreover, the recent work by Lopatin et al. (2021) demonstrate also the application for simultaneous inversion of collocated but not fully coincident backscatter profiles registered by advanced lidar systems or radiosondes and radiation measured by the Sun photometer.

The implementation of multi-pixel fitting allows to use it as one aditional contraint.

This multi-pixel constraints limit variability for unknowns in different groups of similar parameters when several such groups of unknowns are retrieved simultaneously from coordinated but not fully co-incident or not fully co-located observations. For example, a priori constraints about known limited inter-pixel variability of retrieved parameters can be realized by using a priori knowledge about limitations on derivatives on time or spatial variability of parameters retrieved from observations in different pixels. It can be provided from the neighboring pixels to the observed pixel in the satellite observations. In principle, the variability of each physical parameter a_i can be considered as a value of continues function $\mathrm{a}_i = \mathrm{a}_i(x, y, z, ..., t)$. Therefore, the limitation on the variability of every parameter in time and space that can be used as additional constraints. Similar to Eq. 3.31, inter-pixel variability constraints are related with limited values of the derivatives, i.e. with their m-th derivatives deviations from zero. At present, the time- and spatial - variation of each parameter in GRASP can be limited using the following a priori assumptions:

$$\begin{aligned} &\frac{\partial^m \mathrm{a}_i(x,y,z,t,...)}{\partial x^m} \approx 0, \frac{\partial^m \mathrm{a}_i(x,y,z,t,...)}{\partial y^m} \approx 0, \\ &\frac{\partial^m \mathrm{a}_i(x,y,z,t,...)}{\partial z^m} \approx 0 \quad \text{and} \quad \frac{\partial^m \mathrm{a}_i(x,y,z,t,...)}{\partial t^m} \approx 0 \end{aligned} \tag{3.34}$$

that can be presented in matrix form similarly to Eq. 3.23 written for a single-pixel case as:

$$\begin{cases} \mathbf{f}_2^{\mathrm{a},*} = \mathbf{f}_2^{\mathrm{a},*}(\mathbf{a}) + \Delta\mathbf{f}_2^{\mathrm{a},*} \\ \mathbf{f}_3^{\mathrm{a},*} = \mathbf{f}_3^{\mathrm{a},*}(\mathbf{a}) + \Delta\mathbf{f}_3^{\mathrm{a},*} \\ \mathbf{f}_4^{\mathrm{a},*} = \mathbf{f}_4^{\mathrm{a},*}(\mathbf{a}) + \Delta\mathbf{f}_4^{\mathrm{a},*} \\ \mathbf{f}_5^{\mathrm{a},*} = \mathbf{f}_5^{\mathrm{a},*}(\mathbf{a}) + \Delta\mathbf{f}_5^{\mathrm{a},*} \end{cases} \rightarrow \begin{cases} \mathbf{0}_x^* = \mathbf{G}_{x,m_x}\mathbf{a} + \Delta_x^* \\ \mathbf{0}_y^* = \mathbf{G}_{y,m_y}\mathbf{a} + \Delta_y^* \\ \mathbf{0}_z^* = \mathbf{G}_{z,m_z}\mathbf{a} + \Delta_z^* \\ \mathbf{0}_t^* = \mathbf{G}_{t,m_t}\mathbf{a} + \Delta_t^* \end{cases} \tag{3.35}$$

Then, for the solution for multi-pixel fitting equivalent of Eq. 3.32, written for a single-pixel case, can be presented as:

$$\begin{pmatrix} \mathbf{A}_{1,p} & \mathbf{0} & ... & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{2,p} & ... & \mathbf{0} \\ & & & \\ ... & ... & ... & ... \\ \mathbf{0} & \mathbf{0} & ... & \mathbf{A}_{N,p} \end{pmatrix} + \boldsymbol{\Omega}_{inter} \Delta\mathbf{a}^p = \begin{pmatrix} \Psi_1(\mathbf{a}_1^p) \\ \Psi_2(\mathbf{a}_2^p) \\ ... \\ \Psi_N(\mathbf{a}_N^p) \end{pmatrix} + \boldsymbol{\Omega}_{inter}\mathbf{a}^p \tag{3.36}$$

where $\mathbf{A}_{i,p}$ and $\Psi_i(\mathbf{a}_i^p)$ refer to the left and right parts of Eq. 3.23 defined for i-th single pixel, so that Eq. 3.23 can be denoted compactly as $\mathbf{A}_{i,p}\Delta\mathbf{a}_i^p = \Psi_i(\mathbf{a}_i^p)$.

The matrix $\mathbf{\Omega}_{inter}$ of multi-pixel constraints is defined via smoothness matrices for the spatial and temporal variability of each retrieved parameter as:

$$\mathbf{\Omega}_{inter} = \gamma_x \mathbf{\Omega}_{x,m_x} + \gamma_y \mathbf{\Omega}_{y,m_y} + \gamma_z \mathbf{\Omega}_{z,m_z} + \gamma_t \Omega_{t,m_t} \tag{3.37}$$

The detailed description of multi-pixel constraints and their application is provided in the paper by Dubovik et al. (2011). Specifically, Dubovik et al. (2011) provide explicit expressions for $\mathbf{\Omega}_{x,m_x}$, $\mathbf{\Omega}_{y,m_y}$ and $\mathbf{\Omega}_{t,m_t}$, as well as corresponding Lagrange parameters γ_x, γ_y and γ_t. Note, the equations are fully implemented in GRASP code with the exception of $\mathbf{\Omega}_{z,m_z}$ and γ_z in Eqs. 3.34-3.35 that are not yet implemented in the same way since in the practice there are almost no fully independent vertical observations.

3.2.3 Non-linear inversion in GRASP and used Levenberg-Marquardt optimization

Since most of atmospheric remote sensing applications are strongly non-linear, the Levenberg-Marquardt optimization (Press et al. (1992); Ortega and Rheinboldt (1970)) is realized to optimize convergence of GRASP solutions. Specifically, as described by Dubovik et al. (2021) in GRASP it is assumed that the correction of the solution at p-th iteration $\Delta\mathbf{a}^p$ should be limited, especially at the initial iterations when the linearization error is the largest. For such cases, in GRASP, for the determination of $\Delta\mathbf{a}^p$ in the iterative procedure an additional constrained on the correction $\Delta\mathbf{a}^p$ is added at each iteration:

$$\Delta\mathbf{a}^{p,*} = \mathbf{0}^* + \Delta\mathbf{a} \tag{3.38}$$

Correspondingly, using this additional requirement, an additional term will be introduced in Eq. 3.23:

$$\left(\sum_{k=1}^{K} \gamma_k \mathbf{K}_k^T (\mathbf{W}_k)^{-1} \mathbf{K}_k + \mathbf{D}_{\Delta\mathbf{a}}^p\right) \Delta\mathbf{a}^p = \sum_{k=1}^{K} \gamma_k \mathbf{K}_k^T (\mathbf{W}_k)^{-1} \Delta\mathbf{f}_k^p \tag{3.39}$$

where matrix $\mathbf{D}_{\Delta\mathbf{a}}$ is diagonal matrix with the elements:

$$\mathbf{D}_{\Delta\mathbf{a}ii} = \gamma_{\Delta\mathbf{a}_i} = \frac{\varepsilon_1^2}{\varepsilon_{\Delta\mathbf{a}_i}^2} \tag{3.40}$$

The variance $\varepsilon_{\Delta\mathbf{a}_i}^2$ can be determined, for example, assuming that whole known range of each parameter $\mathbf{a}_i$ variability should be covered by $3\varepsilon_{\Delta\mathbf{a}_i}^2$, i.e. $\mathbf{a}_{i,max} - \mathbf{a}_{i,min} \approx 3\varepsilon_{\Delta\mathbf{a}_i}$.

Also, following common Levenberg-Marquardt procedure the impact of the correc-

tion $\Delta\mathbf{a}^p$ is always scaled by a factor t_p in Eq. 3.20 as follows:

$$\mathbf{a}^{p+1} = \mathbf{a}^p - t_p \Delta\mathbf{a}^p \tag{3.41}$$

where t_p is in the range $0 < t_p \leq 1$. It is selected empirically to provide convergence, by decreasing $t_p = t_p/2$ until decrease of the residual $\Psi'(\mathbf{a}^p) \leq \Psi'(\mathbf{a}^{p-1})$ is achieved (see Dubovik et al. (2011)).

Thus, in case of non-linear $\mathbf{f}_k(\mathbf{a})$ and/or $\mathbf{f}_i^{\mathrm{a}}(\mathbf{a})$ the inversion in GRASP includes Levenberg-Marquardt like optimizations and implemented in the frame of Eqs. 3.20 and 3.21. While this optimization certainly helps to achieve successful convergence of the solution in practice, it also should be considered as one of possible sources of uncertainties, as pointed by Dubovik et al. (2021) and will be discussed below.

3.3 Error estimates in GRASP

The least-squares fitting procedures, described in the last sections, are commonly used in data analysis and they are extensively discussed in the literature. Particularly, GRASP algorithm has a rather general and rigorous numerical inversion that allows transparency and flexibility to retrieve unknown parameters from several instruments and synergies. However, the proper assessment of errors resulting from such fits has received relatively little attention.

The error analysis consists in the sensitivity of the retrieval to all of the sources of error, including noise in the measurement, error in the non-retrieved parameters and in the retrieval parameters, and the effect of modelling the true physics of the measurement by some forward model, if needed.

This chapter contains the formal mathematical descriptions of the implemented error estimates as part of this thesis, in the GRASP algorithm. It starts introducing the state of the art of error estimates in the frame of aerosol properties briefly describing different methodologies employed in several works and the differences with the GRASP approach. Then this chapter is focused in the formalized equations of the errors estimates of the retrieved and derived parameters from the use of the statistical optimization approach of multi-term LSM. Moreover, in this chapter is presented the advantages of obtaining a full covariance matrix of the retrieved parameters focusing in the correlation matrix. Some important aspects of this matrix and their interpretation are described since the analysis of the full covariance matrix will help to identify characteristics of retrievals that can be optimised, and it is in any case an essential part of the documentation of any retrieved data set. It can be argued that a retrieval method without an error analysis and characterisation is of little value.

Table 3.1: AOD and extinction data sets providing prognostic (predictive) uncertainty estimates and their associated key references for uncertainty estimate calculation. When it is applicable, algorithm names are given first with instrument names in parentheses. (Extracted from Sayer et al. (2020)).

Data set	Key references for uncertainty	Note
ADV/ASV (ATSR2, AATSR)	Kolmonen et al. (2016); Kolmonen and Sogacheva (2018)	Jacobians at retrieval solution
AerGOM (GOMOS)	Vanhellemont et al. (2016)	Maximum likelihood with smoothness constraints
BAR (MODIS)	Lipponen et al. (2018)	Maximum likelihood, retrieves whole granule at once
CALIPSO	Young et al. (2013, 2018)	Propagation of contributions through lidar equation
CATS	Young et al. (2013)	Propagation of contributions through lidar equation
CISAR (CHRIS, SEVIRI)	Govaerts and Luffarelli (2018); Luffarelli and Govaerts (2019)	Optimal estimation with smoothness constraints
DB AAC (MODIS, SeaWiFS, VIIRS)	Sayer et al. (2016); Sayer et al. (2019b)	Maximum likelihood
DB land (MODIS)	Sayer et al. (2013)	Empirical expression from AERONET validation results
GOCI	Choi et al. (2018)	Empirical expression from AERONET validation results
GRASP (MERIS, POLDER)	Dubovik et al. (2011)	Maximum likelihood with smoothness constraints
IMARS (IASI)	Klüser et al. (2011)	Propagated measurement and forward model terms
JAXA (AHI)	Yoshida et al. (2018)	Optimal estimation
LDA (SEVIRI)	Govaerts et al. (2010)	Optimal estimation
LMD (AIRS, IASI)	Pierangelo et al. (2004); Capelle et al. (2014)	Parametric from sensitivity studies and validation
MAIAC (MODIS)	Lyapustin et al. (2018)	Propagated from uncertainty on surface reflectance
MAPIR (IASI)	Vandenbussche et al. (2013)	Optimal estimation
MIPAS	Günther et al. (2018)	Maximum likelihood with smoothness constraints
MISR dark water	Witek et al. (2018b)	Width of cost function distribution vs. AOD
MISR heterogeneous land	Martonchik et al. (1998, 2009)	Standard deviation of well-fitting aerosol models
MODACA (MODIS)	Meyer et al. (2015)	Maximum likelihood
NOAA EDR (VIIRS)	Huang et al. (2016)	Empirical expression from AERONET validation results
OMPS LP	Loughman et al. (2018)	Confidence envelope based on aerosol signal strength
ORAC (ATSR2, AATSR, SEVIRI, SLSTR)	Thomas et al. (2009, 2010, 2017)	Optimal estimation
PMAp	EUMETSAT (2015)	Standard deviation of aerosol models
OSIRIS	Bourassa et al. (2007); Rieger et al. (2019)	Optimal estimation
SAGE	McCormick et al. (2002)	Propagated measurement plus interfering species error
SU (ATSR2, AATSR, MERIS+AATSR)	North et al. (2010, 2017); Bevan et al. (2012)	Second derivative of error function
ULB (IASI)	Clarisse et al. (2013)	Propagated measurement and forward model terms

3.3.1 State of the art of error estimates for aerosol properties

The study of error estimates is a great challenge in recent times. As mentioned above, over the years, the installation of different remote sensing instruments has advanced, such as the development of algorithms for retrieving different aerosol and atmospheric properties. However, the estimation of the uncertainties of the retrieved properties still remains as one of the most important open questions to be resolved.

There are different ways to estimate the error of the retrieved and derived properties using several observations. For example, some authors do it from validation with data-sets from other instruments (for example, retrieval of AOD from satellite measurements validated with AOD directly measured from ground based sun-photometers) or perturbing some input parameters and analysing the retrieved properties. These studies are useful to identify general tendencies of bias or loss of sensitivity under different conditions and to evaluate possible ways to improve them (Sayer et al. (2020)). Nevertheless, there are also algorithms that consist of formal error propagation techniques or more empirical methods. For example, some of these algorithms used from satellite data are presented in Table 3.1 extracted from Sayer et al. (2020).

In recent years, the inclusion of different formal techniques for uncertainty estimates within the satellite AOD data sets has increased. However, more studies are still necessary since they are conditioned by several factors. Some of them consist to have an appropriated forward model, capable of providing unbiased estimates of the

observations; suitable covariance matrices; verify that the retrievals converge around a global minimum and not a local minimum, etc.

In parallel with the recent advances in the satellite remote sensing techniques ground-based observations, the algorithms associated (for the retrievals of atmospheric properties) and the error estimates analysis (of the retrieved properties) have experimented an important growing in the last years.

As was mentioned in previous chapter, one of the most visible data sets of ground-based radiometric observations is provided by AERONET (AErosol RObotic NETwork, Holben et al. (1998)), network of more than 500 (according to Sinyuk et al. (2020)) operational sites distributed over the world. Different studies were performed the AERONET retrieval errors. First, Dubovik et al. (2000) has provided a rather comprehensive analysis of retrieval uncertainties caused by both random measurements errors and systematic errors originated from potential biases in the measurements and imperfections in the modeling aerosol properties. This analysis was revisited by Torres et al. (2014) studies that overall confirmed most of the uncertainty tendencies revealed by Dubovik et al. (2000). Recently, Sinyuk et al. (2020) published a concept for aerosol retrieval error estimates that have been adapted in Version 3 aerosol operational product of AERONET (Giles et al. (2019)). In this approach, the uncertainties are estimated using the spread of the retrieved parameters generated by 27 distinct combinations of retrievals obtained by perturbations in the input data (AOD, sky radiances, solar spectral irradiances and surface reflectances). A scheme of this implemented approach is shown in Figure 3.4.

Somehow a similar concept for error estimates was earlier employed in the LiRIC (Lidar and Radiometer Inversion Code) approach for the synergy processing of co-located lidar and AERONET radiometric observations (Chaikovsky et al. (2016)). LiRIC provided some uncertainties obtained by a series of retrievals perturbing the input data.

There are numerous advantages and disadvantages between the methods described so far. For example, one advantage of the method used in AERONET is that is logistically clear; but on the contrary, its implementation is awkward since each perturbed input parameter is separately inverted. On the other hand, several factors affect retrievals and they are difficult to model theoretically, so we are interested in error propagation models. Those techniques are based on the statistical optimization approach (e.g., Edie et al. (1971); Fourgeaud and Fuchs (1967); Rodgers (2000), etc) which provide asymptotically comprehensive estimates for random retrieval errors. At the same time, it should be noted that both the result of perturbation tests and statistical estimates of propagated error rely on the employed forward model. In this sense, the inaccuracies related to the limitations of the selected forward model may not be fully represented. Some additional evaluations and considerations are always desirable

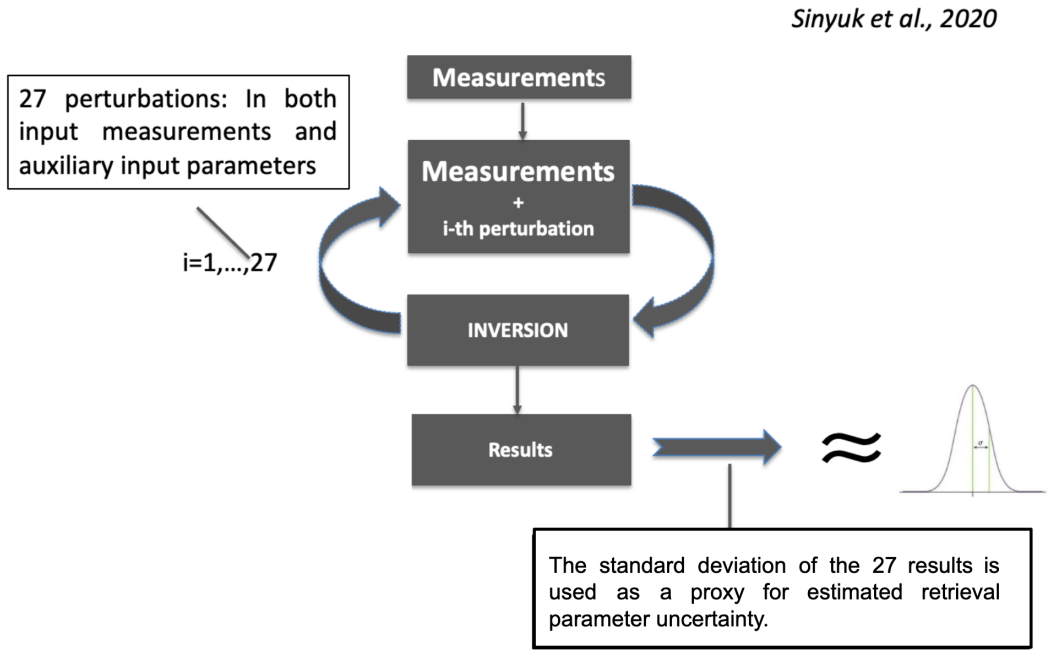


Figure 3.4: Scheme of error estimates approach implemented in the new AERONET version 3.

for assessing the adequacy of the chosen forward model and its potential limitations.

Therefore, this thesis focuses on the study of the errors obtained by the GRASP algorithm, whose technique and formalized equations will be described in the next sections. This algorithm is based on the rather general and rigorous forward modeling and inversion approaches, and in addition to the retrieved parameters, it provides dynamic error estimates. Specifically, the great advantage is the capability of generating full covariance matrices that include both random and systematic components which allows a complete study of the error estimates.

3.3.2 Theoretical developments in GRASP

As part of the numerical inversion, GRASP algorithm provides dynamic error estimates for directly retrieved parameters and for derived parameters. After understanding the numerical inversion methods and the employed equations in the last sections, an important step to obtain the error estimates in the fitting procedure is the understanding of the formalized equations of error estimates under the assumptions of statistical optimization approach. It is based on fundamental principles where the retrieval is based on consideration of probability density function (PDF) of both measurements and retrieved values. Thus, this strategy presents the advantage of allowing the easy characterization of uncertainties since the idea of the optimal retrieval is mainly based on getting a PDF as narrow as possible which can be used almost straightforwardly to

characterize errors. Therefore, the key in this study consists in obtaining the covariance matrix of the retrieved and derived parameters which will allow subsequent studies of the error estimates.

In the Section 3.2 we assumed the normal distribution of $\Delta\mathbf{f}^*$. Consequently, the error in the retrieved parameters $\Delta\hat{\mathbf{a}}$ also will be normally distributed and will have two components, random and systematic:

$$\Delta\hat{\mathbf{a}} = \Delta\hat{\mathbf{a}}_{rand} + \Delta\hat{\mathbf{a}}_{sys} \tag{3.42}$$

Their variance is defined as following:

$$< (\Delta\hat{\mathbf{a}})^2 >=< (\Delta\hat{\mathbf{a}}_{rand})^2 > + < (\Delta\hat{\mathbf{a}}_{sys})^2 > \tag{3.43}$$

where $\hat{\mathbf{a}}$ is the logarithm of the retrieved parameter.

These variances correspond to the diagonal elements of the full covariance matrix, $\mathbf{C}_{\hat{\mathbf{a}}}$ provided by GRASP algorithm. Thus, from covariance definition and using Eq. 3.42, $\mathbf{C}_{\hat{\mathbf{a}}}$ can be calculated as:

$$\begin{aligned}
\mathbf{C}_{\hat{\mathbf{a}}} &= < \Delta\hat{\mathbf{a}}\Delta\hat{\mathbf{a}}^T > \\
&= < (\Delta\hat{\mathbf{a}}_{ran} + \Delta\hat{\mathbf{a}}_{sys})(\Delta\hat{\mathbf{a}}_{ran} + \Delta\hat{\mathbf{a}}_{sys})^T > \\
&= < (\Delta\hat{\mathbf{a}}_{ran} + \Delta\hat{\mathbf{a}}_{sys})(\Delta\hat{\mathbf{a}}_{ran}^T + \Delta\hat{\mathbf{a}}_{sys}^T) > \\
&= < (\Delta\hat{\mathbf{a}}_{ran}\Delta\hat{\mathbf{a}}_{ran}^T) + (\Delta\hat{\mathbf{a}}_{sys}^T\Delta\hat{\mathbf{a}}_{ran}) + (\Delta\hat{\mathbf{a}}_{ran}^T\Delta\hat{\mathbf{a}}_{sys}) + (\Delta\hat{\mathbf{a}}_{sys}\Delta\hat{\mathbf{a}}_{sys}^T) > \\
&= < \Delta\hat{\mathbf{a}}_{ran}\Delta\hat{\mathbf{a}}_{ran}^T > + < \Delta\hat{\mathbf{a}}_{sys}^T\Delta\hat{\mathbf{a}}_{ran} > + < \Delta\hat{\mathbf{a}}_{ran}^T\Delta\hat{\mathbf{a}}_{sys} > + < \Delta\hat{\mathbf{a}}_{sys}\Delta\hat{\mathbf{a}}_{sys}^T > \\
&= \mathbf{C}_{\Delta\hat{\mathbf{a}}_{ran}} + \hat{\mathbf{a}}_{sys}\hat{\mathbf{a}}_{sys}^T
\end{aligned} \tag{3.44}$$

Equation 3.44 clearly shows the covariance matrix of the errors have also two components, random and systematic. In the case of multi-term LSM and applying the definition of covariance for $\Delta\hat{\mathbf{a}}_{ran}$ (Dubovik (2004)), the covariance matrix of random errors ($\Delta\hat{\mathbf{a}}_{ran}$) can be estimated in the linear approximation as:

$$\begin{aligned}
\mathbf{C}_{\Delta\hat{\mathbf{a}}_{ran}} &= \left(\sum_{k=1}^{K}\gamma_k\mathbf{K}_k^T\mathbf{W}_k^{-1}\mathbf{K}_k\right)^{-1}\hat{\varepsilon}_0^2 \\
&= \mathbf{A}_p^{-1}\hat{\varepsilon}_0^2
\end{aligned} \tag{3.45}$$

where $\mathbf{A}_p$ is a Fisher matrix and the variance $\hat{\varepsilon}_0^2$ is estimated from the general residual fitting as:

$$\hat{\varepsilon}_0^2 \sim \frac{\mathbf{\Psi}(\hat{\mathbf{a}}^p)}{(N_{meas} - N_{\mathbf{a}} - N_{\mathbf{aprior}})} \tag{3.46}$$

where N_{meas} and N_{aprior} are the respective numbers of measurements and a priori data, and N_{a} is the number of retrieved parameters. This equation has an important characteristic: it is obtained from Cramer-Rao inequality since the solution to minimizing the equation 3.23 has the smallest errors, it means the optimality of LSM estimates. Thus, if all assumptions about noise in both the measurements and the a priori terms are correct, the minimum value of the above quadratic form can be theoretically estimated from χ^2 distribution, whose mean is given by:

$$\begin{aligned}\langle(2\Psi(\hat{\mathbf{a}}))_{min}\rangle &= \left\langle 2\sum_{i=1,\ldots,3}\Psi(\hat{\mathbf{a}})\right\rangle = \left\langle \left(\hat{\mathbf{f}}-\mathbf{f}^*\right)^T \mathbf{W}^{-1}\left(\hat{\mathbf{f}}-\mathbf{f}^*\right) + \gamma_\Delta \hat{\mathbf{a}}^T\mathbf{\Omega}_m\hat{\mathbf{a}} + \gamma_{\mathbf{a}^*}\hat{\mathbf{a}}^T\mathbf{W}_{\mathbf{a}^*}^{-1}\hat{\mathbf{a}}\right\rangle \\ &= \left(\sum_{i=1,\ldots,3} N_{\mathbf{f}_i} - N_{\mathbf{a}}\right)\varepsilon_0^2 = (N_{\mathbf{f}^*} + N_{\Delta^*} + N_{\mathbf{a}^*} - N_{\mathbf{a}})\varepsilon_0^2 = (N_{\mathbf{f}^*} + N_{\mathbf{a}} - m)\varepsilon_0^2\end{aligned} \tag{3.47}$$

where $N_{\Delta^*} = N_{\mathrm{a}} - m$ and $N_{\mathrm{a}^*} = N_{\mathrm{a}}$. Then, from this equation can be estimated ε_0^2 from a minimum value of the residual $\Psi(\hat{\mathbf{a}})$. Therefore, this expression can be used to verify the consistency of the retrieval. The estimation of Eq. 3.46 should be close to the assumed. A significant increase of the estimated value over expected can be considered as an indication of unaccounted biases and/or inadequate assumptions about random errors in measurements or a priori data sets.

A clear tendency is observed from Eq. 3.45, the smaller random errors occur when higher contribution of the a priori. It means, the use of more a priori constraints leads to less random errors of the retrieval. However, the incorrect use of these a priori constraints could contribute to the introduction of some systematic uncertainties, i.e., bias. Therefore, for an adequate evaluation of retrieval uncertainties is important to estimate the contribution of bias, especially when multiple a priori constraints are used.

On the other hand, the systematic component of the error $\Delta\hat{\mathbf{a}}_{sys}$ is defined as the correction calculated at the last iteration (as in the Eq. 3.23) in the process to invert the Eq. 3.21:

$$\Delta\hat{\mathbf{a}}_{sys} = \Delta\hat{\mathbf{a}}_p \tag{3.48}$$

and is estimated as following:

$$\hat{\mathbf{a}}_{sys} = \left(\sum_{k=1}^{K}\gamma_k\left(\mathbf{K}_k^T\,\mathbf{W}_k^{-1}\,\mathbf{K}_k\right)\right)^{-1}\left(\sum_{k=1}^{K}\gamma_k\left(\mathbf{K}_k^T\,\mathbf{W}_k^{-1}\,\mathbf{b}_k^*\right)\right) \tag{3.49}$$

where $\mathbf{b}_k^*$ represent the bias vector in the $k-th$ data set $\mathbf{f}_k$.

The estimation of not only random retrieval error but also error retrieval bias $\Delta\mathbf{a}_{sys}$ is important for the adequate evaluation of retrieval uncertainty, especially in the case when multiple a priori constraints are used. For example, for the case of the retrieval given by Eq. 3.30 $\mathbf{C}_{\Delta\hat{\mathbf{a}}_{ran}}$ is expressed as:

$$\mathbf{C}_{\Delta\hat{\mathbf{a}}_{ran}} \approx \left(\mathbf{K}^T\mathbf{W}^{-1}\mathbf{K} + \gamma_{\mathbf{a}}\mathbf{W}_{\mathbf{a}}^{-1} + \gamma_g\mathbf{\Omega}_m\right)^{-1}\hat{\varepsilon}_0^2 \tag{3.50}$$

A rather obvious tendency can be seen from the analysis of this equation: the higher the contributions of the second and the third terms the smaller the random errors are, i.e. the stronger a priori constraints are used the lower the random errors of the retrieval. However, in practice a priori constraints can be unintentionally inadequate and therefore introduce some systematic uncertainties, i.e. biases. In principle, there is no guaranteed approach for detecting those biases unless comprehensive analysis and validation of the retrievals have been done. Nonetheless, some biases can manifest themselves via misfit of measurements $\Delta\mathbf{f}_k^{sys} = \mathbf{f}_k(\mathbf{a}^{solution}) - \mathbf{f}_k^*$ or misfit of a priori constraints. For example, for Eq. 3.30 the bias can be introduced by a priori estimate $\mathbf{a}_{sys}^* = \mathbf{a}^{solution} - \mathbf{a}^*$ or unsmooth features in the retrieved solution: $\mathbf{a}_{sys}^{smooth} = \Omega_m\mathbf{a}^{solution} \neq \mathbf{0}$. Correspondingly, the bias for single-pixel retrieval is estimated as:

$$\hat{\mathbf{a}}_{sys} \approx \left(\mathbf{K}^T\mathbf{W}^{-1}\mathbf{K} + \gamma_{\mathbf{a}}\mathbf{W}_{\mathbf{a}}^{-1} + \gamma_g\mathbf{\Omega}_{\mathbf{m}}\right)^{-1}\left(\mathbf{K}^T\mathbf{W}^{-1}\Delta\mathbf{f}^{sys} + \gamma_{\mathbf{a}}\mathbf{W}_{\mathbf{a}}^{-1}\mathbf{a}_{sys}^* + \gamma_g\mathbf{\Omega}_{\mathbf{m}}\mathbf{a}_{sys}^{smooth}\right) \tag{3.51}$$

In this equation the contribution of a priori estimates to bias is probably the most significant in many applications since it is never possible to have fully accurate a priori values (widely used in optimum estimation approaches) for constraining. In a similar way, the a priori biases are estimated in the case when multi-pixel a priori constraints are used.

The Levenberg-Marquardt optimization of the convergence, discussed in Section 3.2.3 may also introduce a bias. Indeed, this optimization makes the iterations converge from given initial guess to fit the data even if the basic linear system is singular. Therefore, once Levenberg-Marquardt optimization is used there is an evident dependence on the initial guess that can bias the solution. In order to take this into account the Eqs. 3.45 and 3.49 are modified as the following:

$$\mathbf{C}_{\Delta\hat{\mathbf{a}}_{ran}} \approx \left(\sum_{k=1}^{K}\gamma_k\mathbf{K}_k^T\mathbf{W}_K^{-1}\mathbf{K}_k + \mathbf{D}_{\Delta\mathbf{a}}^p\right)^{-1}\hat{\varepsilon}_0^2 \tag{3.52}$$

and

$$\hat{\mathbf{a}}_{sys} = \left(\sum_{k=1}^{K} \gamma_k \mathbf{K}_k^T \mathbf{W}_K^{-1} \mathbf{K}_k + \mathbf{D}_{\Delta\mathbf{a}}^p \right)^{-1} \left(\sum_{k=1}^{K} \gamma_k \mathbf{K}_k^T \mathbf{W}_K^{-1} \mathbf{b}_k^* + \mathbf{D}_{\Delta\mathbf{a}}^p (\mathbf{a}^{solution} - \mathbf{a}^{p=0}) \right) \tag{3.53}$$

These equations allow to obtain the error estimates for the retrieved parameters. That is, for example when the configuration is from sun photometer and lidar measurements, the expected retrieved parameters are $dV/dln(r)$, real and imaginary part of refractive index, sphericity fraction and aerosol volume concentration vertically distributed.

Also, in the practice the users may not need directly the retrieved parameters $\hat{\mathbf{a}}$ but their functions $m(\hat{\mathbf{a}})$ that can be calculated from the retrieved parameters. For example, GRASP retrieves parameters of aerosol microphysics (particle sizes, refractive indices, etc.) but users need aerosol optical depth, AOD. For such situation, GRASP provides a set of such diverse indirect characteristics with the possibilities of providing the uncertainties calculated as:

$$\begin{aligned} \mathbf{C}_{\Delta\hat{m}} &\approx \mathbf{M} \left(\mathbf{C}_{\Delta\hat{\mathbf{a}}_{ran}} + \hat{\mathbf{a}}_{sys} \hat{\mathbf{a}}_{sys}^T \right) \mathbf{M}^T \\ &= \mathbf{M}\mathbf{C}_{\Delta\hat{\mathbf{a}}_{ran}} \mathbf{M}^T + \mathbf{M}\hat{\mathbf{a}}_{sys} (\mathbf{M}\hat{\mathbf{a}}_{sys})^T \\ &= \mathbf{C}_{\Delta\hat{m}_{ran}} + \hat{\mathbf{m}}_{sys} \hat{\mathbf{m}}_{sys}^T \end{aligned} \tag{3.54}$$

where **M**- is the is the matrix of first derivatives $\mathbf{M}_{ji} = \left. \frac{\partial m_j}{\partial a_i} \right|_{\mathbf{a}^{solution}}$

Finally, the effect of biases in the measurements on the solution bias $\hat{\mathbf{a}}_{sys}$ is accounted for in Eq. 3.51 based on the assumption that the presence of biases is manifested in the non-zero misfits $\Delta\mathbf{f}_k^{sys}$. Indeed, in many cases when systematic errors are present in the inverted measurements or the accurate fit of inverted data can't be achieved (e.g. see illustrations provided by numerical sensitivity tests for AERONET retrievals by Dubovik et al. (2000)). At the same time, there are many situations when biases in the measurements may not significantly affect the residual (Eq. 3.25) and the misfits $\Delta\mathbf{f}_k^{sys}$. For example, the retrievals of aerosol SSA from AERONET ground-based measurements are highly sensitive to the calibration biases in the direct Sun measurements, while the fitting of these direct measurements is always quite accurate (see discussion by Dubovik et al. (2000)). The effects of such measurement biases can be estimated by implementing proxy numerical tests applied to the measurements perturbed by possible biases. For example, the recent approach for evaluation retrieval errors of AERONET operational products is estimated using a series of ∼27 numerical proxy inversion tests with the sets of perturbations in both input measurements and auxiliary input parameters (Sinyuk et al. (2020)). Similar strategy can be used for eval-

uation of potential effects of undetected biases. Specifically, the bias term $(\hat{\mathbf{a}}_{sys})(\hat{\mathbf{a}}_{sys})^T$ in Eq. 3.44 can be estimated as:

$$(\hat{\mathbf{a}}_{sys})(\hat{\mathbf{a}}_{sys})^T \rightarrow \left\langle (\hat{\mathbf{a}}_{sys})(\hat{\mathbf{a}}_{sys})^T \right\rangle_{\text{bias proxy set}} \tag{3.55}$$

where the values of the retrieval biases are estimated as an average effect from a preselected set of possible biases in measurements and auxiliary inputs. Therefore, if we assume positive and negative bias in the equation for systematic component the contribution to Eq. 3.51 can be written as follow:

$$\hat{\mathbf{a}}_{sys} \approx \left(\mathbf{K}^T\mathbf{W}^{-1}\mathbf{K} + \gamma_{\mathbf{a}}\mathbf{W}_{\mathbf{a}}^{-1} + \gamma_g\mathbf{\Omega}_{\mathbf{m}}\right)^{-1}\left(\mathbf{K}^T\mathbf{W}^{-1}\mathbf{b}_{\mathbf{f}_{sys}} + \gamma_{\mathbf{a}}\mathbf{W}_{\mathbf{a}}^{-1}\mathbf{a}_{sys}^{*} + \gamma_g\mathbf{\Omega}_{\mathbf{m}}\mathbf{a}_{sys}^{smooth}\right) \tag{3.56}$$

where the vectors $\mathbf{b}_{\mathbf{f}_{sys}}$ represent the new bias related to the measurement.

In addition, in this work we also study the structure of the covariance matrix for different aerosols and configurations. Apparently, such matrix provides interesting information about the error estimates (focusing in the diagonal elements) and the relation between the retrieval parameters (from the covariance values, i.e. non-diagonal elements). The representation of the covariance matrix for the parameters has the following structure:

$$Cov(\mathbf{a}) = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \sigma_1\sigma_3\rho_{13} & \cdots \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{23} & \cdots \\ \sigma_3\sigma_1\rho_{31} & \sigma_3\sigma_2\rho_{32} & \sigma_3^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.57}$$

where in the diagonal are the variance of each element and the non-diagonal elements represent the covariance of each retrieved element a_i with the others. The variances, i.e. diagonal elements are always used for estimating retrieval errors and providing the error bars. The non-diagonal elements are rarely considered, while they provide the very interesting and non-obvious information about error correlations.

In order to study the error correlation structure of the error, the following correlation matrix will be considered in this work, that can be obtained from the covariance matrix (Eq. 3.57):

$$Corr(\mathbf{a}) = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \cdots \\ \rho_{21} & 1 & \rho_{23} & \cdots \\ \rho_{31} & \rho_{32} & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.58}$$

where each diagonal element corresponds to the correlation with itself which is equal to 1 and the non-diagonal elements are the correlations related to each parameter that can vary between -1 and 1.

3.3.3 Interpretation of correlation matrix

As we have seen in Section 3.3.2 GRASP has the capability to provide the full covariance matrix of the retrieved and derived parameters. These matrices are a source of information and provide interesting information about the error estimates, since we can obtain from the elements of the diagonal, i.e. variances, estimation of errors; but also from the non-diagonals elements, important information can be obtained since they contain the correlation coefficients between the retrieved parameters.

In this section the correlation matrix is described. As already mentioned, one of its main advantages is that we can identify from its different tendencies and they can help to improve the configurations of inversions.

From the normalization of the covariance matrix (Eq.3.57) can be obtained the correlation matrix, which represents the correlation between the retrieved parameters, as following:

$$Corr(\mathbf{a}) = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \cdots \\ \rho_{21} & 1 & \rho_{23} & \cdots \\ \rho_{31} & \rho_{32} & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.59}$$

where each diagonal element corresponds to the correlation with itself which is equal to 1 and the non-diagonal elements are the correlations related to each parameter that can vary between -1 and 1.

The latter elements provide extra information since it presents the correlation between the different retrieved parameters. The values close to zero indicate that the retrieved parameters are not correlated between them which is associated with independent and stable solutions. On the other hand, values close to ± 1 indicate strong dependency between the retrieved parameters. Moreover, these cases may induce large errors on the parameters through correlations effects (Eq.3.57). In the cases of solutions with correlated parameters, different combinations of them can be found quite close in a local (or absolute) minimum which may difficult the retrieval and increase the errors associated. Finally, it should be noted that a priori constraints can induce correlations between the parameters. For instance, imposed smoothness constraints in size distributions will create correlations between adjacent bins. These type correlations, which are not originated by the nature of the measurements, can be positively considered and

make the retrieval more robust. Thus, the analysis of these non-diagonal elements are very useful for identifying unobvious retrieval tendencies. Therefore, it can be a useful methodology for optimizing observation schemes and retrieval setups.

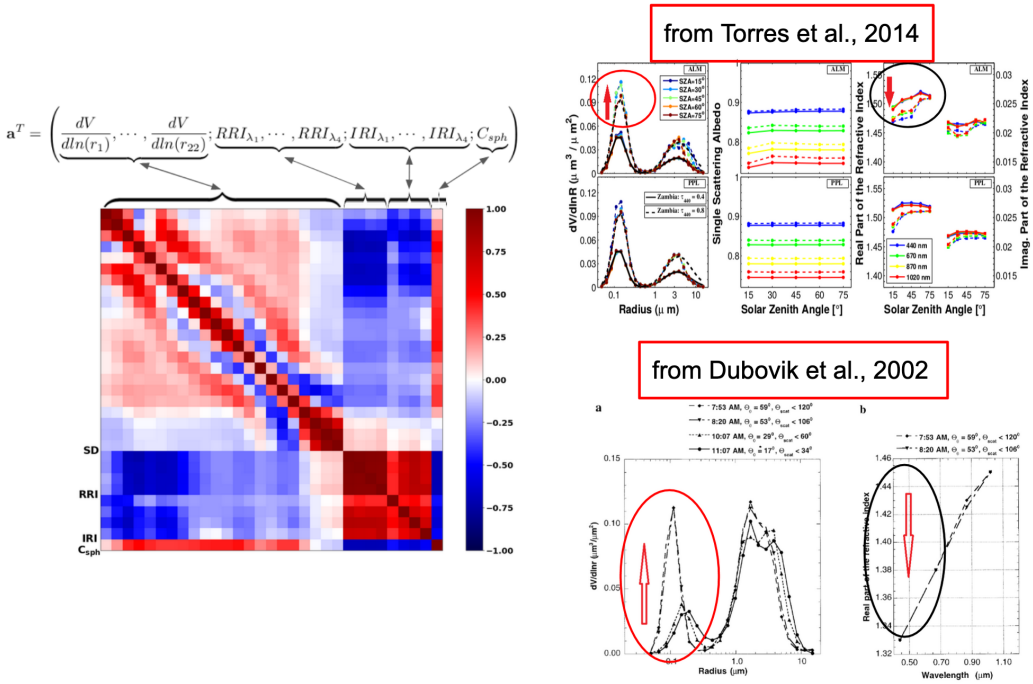


Figure 3.5: On the left, is the illustration of correlation matrix in the simulated case of biomass burning, and on the right side two examples from Torres et al. (2014) and Dubovik et al. (2002b), show the correlation between small particles and RRI at shorter wavelengths.

In the next chapter, several correlation matrices will be analyzed and described for different configurations. At this stage, we would like just to show one first example in order to illustrate the power of this tool. The example deals with the quite known correlation between the small particles and the real refractive index for some retrieval configurations, and more specifically, sun-photometer applications when the scattering information is reduced. The issue was previously shown by Dubovik et al. (2002b) and Torres et al. (2014) and it is summarized on the right part of Figure 3.5. Both authors shows that for an almucantar geometry when the solar zenith angle is reduced (at middle part of the day) and consequently the maximum scattering angle with available measurements, the concentration of fine particles and the real refractive index artificially increases and decreases, respectively. The effects start to be visible when $\Theta_{max} < 90°$ and they clearly indicate an anticorrelation between these two parameters. In those studies, the analysis of the correlation matrices could have foreseen this anticorrelation.

This analysis is done here in the left part of Figure 3.5, where the correlation matrix for a sun-photometer retrieval with similar characteristics of the previous works (fine mode predominate and moderate absorbing aerosol) is represented as a colormap.

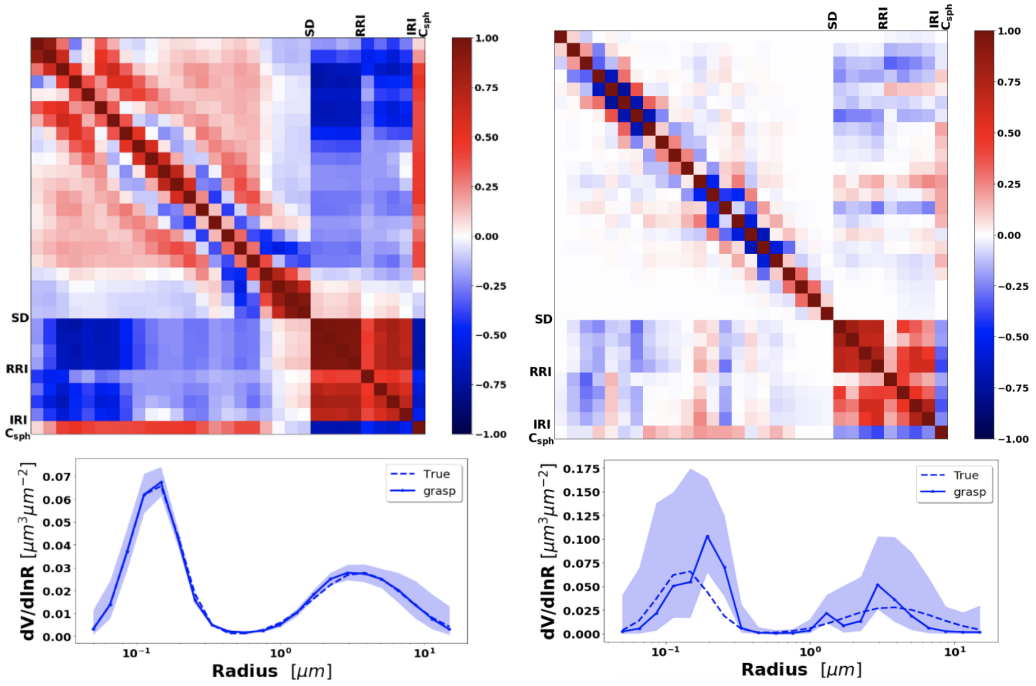


Figure 3.6: On the top, are illustrated the correlation matrix in a simulated case of biomass burning applying low values of a priori constraints for size distribution. The bottom panels show their respective size distributions. On the right side is shown the example using smoothness constraints for size distribution 1000 times lower than on the left side.

The density of the colors indicates values of the correlation coefficients changing from dense red with value equal to 1 to blue -1. Low correlations, values close to zero, are represented with white colors. The figure shows that the volume concentrations of fine particles are strongly negative correlated with RI, especially at shorter wavelengths.

Taking advantage of this fist example, another interesting effect that was aforementioned can be shown. Thus, it is possible to see that, in general, size distribution concentration show rather low correlation values between them except for neighbours pixels. The use of the typical smoothness constraint rises this correlation for the parameters of the size distribution in the vicinity of the diagonal (adjacent bins in Fig. 3.5). Thus, an important question can be formulated: how do this correlation matrix and also their retrieved parameters change if the values of the smothness constraints decrease?

Left part of figure 3.6 shows the retrieved size distribution and correlation matrix for the example presented in figure 3.5, which has been obtained using the typical smoothness constraints for size distribution. In the right part of the figure, both the retrieved size distribution and the corresponding correlation matrix can be seen when lower assumptions of smoothness constraints are used. The first thing that we can get from Figure 3.6 is that the correlations of the matrices change with the variation of the smoothness constraints. It is easy to see that the figure on the right shows lower

correlations between the adjacent points in the vicinity of the diagonal of the matrix for the size distribution part. However, the solution becomes unstable (big different between true and solution obtained by the code), and moreover, the retrieval errors become larger due to uncertainty of the inverse operator as can be seen in the shadowed area presented in bottom panels.

In an extreme case, if all the constraints are removed, the values of the covariance matrix tend to infinity. Thus, in this way, it is not possible to have solutions. The case of an inversion without applying constraints becomes unrealistic because we will not be able to obtain an appropriate solution. Although in principle, the lower the correlations, the better for the solutions, in this case the problem is the determinant of the inverse operator tends to zero. Thus, we will be in the situation of an ill-posed problem and due to the relation of this operator with the covariance matrix, errors are also affected. The relation is shown by the following equation:

$$\mathbf{C}_{\mathbf{a}ii} \sim \left\{ \left(\mathbf{K}^T \mathbf{C}^{-1} \mathbf{K} \right)^{-1} \right\}_{ii} \to \infty \quad \text{for} \quad \det(\mathbf{K}^T \mathbf{C}^{-1} \mathbf{K}) \to 0 \tag{3.60}$$

3.4 Conclusions

Chapter 3 has provided the mathematical formulation of the error estimation implemented in the GRASP algorithm, which is the main subject of study in this thesis. For a better understanding of these equations, the chapter begins by presenting this algorithm.

GRASP is a flexible, generalized algorithm that relies on two independent modules: the forward model and the numerical inversion.

The forward model is developed in a rather universal way to simulate all the inverted remote sensing observations. It consists from several distinct modules such as multiple scattering, aerosol single scattering columnar/volume properties, aerosol vertical profile, surface reflectance and gas absorption calculation.

The aerosol in GRASP retrieval can be represented as rather sophisticated mixture of one or several aerosol components that can differ by particle size distribution, shape distribution, vertical profile and complex index of refraction. All these characteristics can be either assumed and fixed or retrieved. The complexity of modeling each of these characteristics depends on the information content of observations used in each specific GRASP application.

On the other hand, this chapter presents the numerical inversion method in GRASP. It is an important part since it provides the mathematical tools that allow estimating the parameters which characterize a physical system from observed data and a priori information.

GRASP numerical inversion is based in the rigorous statistical optimization approach. It allows the use of multiple a priori constraints in the frame of multi-term of LSM. Moreover, since most of atmospheric remote sensing applications are strongly non-linear, GRASP use the Levenberg-Marquardt method to optimize the convergence of the solutions.

In addition to the retrieved parameters GRASP provides dynamics error estimates. In this chapter was provided the description of the overall concept and specific key implementations of the errors estimation in GRASP. The retrieval error estimates in GRASP are calculated by modeling propagation of measurement errors based on statistical estimation approach. In addition, the formulation used for estimating errors account for some contribution of systematic errors that could be originated from biases in the measurement or some modifications implemented in the algorithm for improving retrieval convergence of non-linear solutions. Furthermore, it can generate full covariances matrix that include both random and systematic components. The diagonal elements of the covariance matrix provide the error bars of the retrieved and derived parameters. However, the full covariance matrix has an important source of information since the non-diagonal elements present the correlation coefficients between the

different retrieved parameters. The analysis of these non-diagonal elements is very useful for identify un-obvious tendencies and contribute to improve the observations schemes and/or retrievals setups.

Chapter 4

Error estimates for synthetic ground-based observations

The present chapter focuses on the validations of the errors through numerical tests. Synthetic data, which will include random and systematic errors, will be inverted and the retrievals will be analyzed. It will allow us to assess the error estimates produced by GRASP algorithm.

Thus, the previous Chapter 3 presented the complete formulation used by GRASP algorithm to provide the errors in the retrieved parameters. However, the practical evaluation of developed error formalism and possible tuning is desirable for comprehensive evaluating of the approach and gaining full confidence in the practical efficiency of the approach. In this regard, one can probably state that the error estimates always tend to be less accurate than the retrievals themselves. Indeed, in remote sensing, the retrieval relies on formalism of electromagnetic light interaction theory that is fundamentally very accurate and well established while the factors contributing to the uncertainties can be very diverse, not fully formalized and often not even fully understood. For example, the forward model is non-linear, while the error propagations are usually (and in this work specifically) estimated in linear approximations. As commented in the previous chapter, the retrieval can be affected by not fully predicted biases in the measurements or by imperceptible biases in aerosols or surface models (in our applications). Therefore, the second important step for establishing error estimates is their evaluation and validation.

This chapter presents the results of the error data evaluation from synthetic ground-based measurements. First, it starts describing the design of the numerical experiment which includes: a) the retrieval scenarios used, b) the description of the overall experiment, c) the assumed atmospheric properties which cover different interesting situations and d) the methodology of error analysis (how the validation is carried out). Finally, the results for the different scenarios are presented in the following sections.

4.1 Aerosol retrieval approaches considered

This section describes the two widely known, and probably the most popular, retrieval scenarios used to derive detailed aerosol optical properties on which this work is focused:

- Retrieval of columnar properties of aerosol from the measurements by ground-based sun/sky-scanning radiometer alone;
- Simultaneous retrieval of both columnar aerosol properties and their vertical distribution from the combined observations by Sun/sky-scanning radiometer and multi-wavelength lidar.

4.1.1 Aerosol retrieval from Sun/sky radiometer alone

The detailed aerosol properties in the total atmospheric column provided by AERONET inversion of Sun/sky-scanning radiometers has been widely recognized as rather unique reliable data. For example, AERONET retrievals provided first reliable data about aerosol spectral absorption and other detailed aerosol optical characteristics (e.g., see Dubovik et al. (2002a); Giles et al. (2012), etc.). These detailed data are of vital importance for evaluating the impact of aerosol on such important aspects as a climate change and diverse pollution effects, and can be reliable estimated nearly uniquely from remote sensing observations (Kaufman et al. (2002)). Therefore, this retrieved aerosol information has been proven to be very useful for assessment of climate change dynamics in IPCC reports (Boucher et al. (2013); Masson-Delmotte et al. (2021)) and other high profiles analyses.

As was already mentioned the evaluations of the accuracy of retrieved aerosol parameters was mainly relied on extensive sensitivity studies by Dubovik et al. (2000). The results were used for providing quality assurance criteria and expected accuracy estimation (see Dubovik et al. (2002a); Holben et al. (2006)). Sinyuk et al. (2020) recently presented the approach to estimate retrieval uncertainties used in AERONET Version 3 data.

All the tests and analyses in the present study include the spectral observations by the ground-based Sun/sky-scanning radiometers. The set of the aerosol parameters retrieved by GRASP in this configuration are the same as the ones obtained by AERONET standard operational processing and they are shown in Table 4.1. Even though GRASP algorithm can provide the error calculations for any radiance geometry, we have focused on retrieval in Solar almucantar only, since for many years has been the only available retrieval of AERONET network. The assessments concerning the recently incorporated hybrid observational geometry (Sinyuk et al. (2020)) may be analysed in future works

Thus, the present study uses the direct sun measurements and sky radiances both at 4 different wavelengths 440 nm, 675 nm, 870 nm and 1020 nm for the inversion tests. These sky radiances measurements are measured in the Solar almucantar (fixed view zenith angle equal to the solar zenith angle, SZA) with a varying azimuth angle ranging from ±3.5 degrees to ±180 degrees (Table 4.1).

4.1.2 Aerosol retrieval from a combination of Sun/sky radiometer and lidar data

The inversion of co-located observation by Sun/sky radiometer and lidar is another popular retrieval approach in aerosol community. Indeed, radiometer direct Sun and multi-angular polarimetric observations of diffuse Sun radiation transmitted through the atmosphere have significant sensitivity to the atmospheric aerosol amount, its particles size, shape and morphology; however, they have practically no sensitivity to the vertical variability of aerosols. The lidar observations on the other hand provide the information about vertical distribution of aerosol while their sensitivity to other aerosol properties is more limited compare to radiometer observations. Therefore, the information from collocated photometric measurements and lidar systems is complementary and always desirable for enhanced characterization of aerosol properties. As was mentioned before, GRASP retrieval has been successfully adapted by Lopatin et al. (2013) for processing such combined observations which was used in numerous studies.

Moreover, GRASP has been adapted to integrated all type of lidar measurements, including Raman channels and polarized channels. Therefore, the use of GRASP concerning the synergy between lidar and photometer includes a vast quantity of different approaches since the lidars of the GRASP users have quite different channels and possibilities. The information contain of each inversion would be related to the type of lidar used, and the analyses of the error estimates would be different for every single application. Particularly in this thesis, we have considered as input data the synergy of sun/sky radiometer and lidar measurements in solar almucantar at four wavelengths combined with the correlative range corrected signal (RCS) values, at 355 nm, 532 nm and 1064 nm. This is the former combination used in the analysis of Lopatin et al. (2013) and still one of the most used today within GRASP users. On the other hand, the lidar signal provided in GRASP as input data is normalized at 60 log-spaced bins at different heights, as in Lopatin et al. (2013, 2021), giving a minimum and maximum heights. The main reason of this election is because all lidars provide observations within a certain distance range, which varies from instrument to instrument and it is limited by emitter/receiver field of view overlap in the lower part as well as by the signal-to-noise ratio in the upper part. In addition to columnar aerosol properties provided from radiometer only inversion, the combined GRASP retrieval derives the

vertical profile of aerosol concentration. Moreover, the retrieval seems to be sensitive to separate optical properties of fine and coarse mode (Lopatin et al., 2013) and aerosol can be considered as an external mixture of fine and coarse aerosol components. Thus, all retrieved parameters are provided for both aerosol modes as shown in the Table 4.1.

Table 4.1 summarizes the used input data in both schemes and the retrieved properties, which are provided also with their error estimates, in each of the cases.

Table 4.1: Summary of general input data and the set of parameters retrieved by GRASP algorithm used in this work for two configurations: Sun/sky radiometer only and Sun/sky radiometer plus lidar.

Sun/sky radiometer Only	**Sun/sky radiometer plus lidar**
Input:	*Input:*
	Sun/sky-radiometer data:
- AOD*	- AOD*
- Calibrated radiances*	- Calibrated radiances*
	Lidar data:
	- Range corrected profiles (RCS**) normalized
* at 440 nm, 675 nm, 870 nm and 1020 nm	at 60 log-spaced bins at different heights
	** at 355 nm, 532 nm and 1064 nm
Retrieved aerosol properties in the total atmospheric column:	*Retrieved aerosol properties (Column-integrated and vertical distributions)*:
- $dV(r_i)/d\ln r_i$	- $dV(r_i)/d\ln r_i$ (in total atmospheric column)
- C_{sph}	- C_{sph} (in total atmospheric column)
- $n(\lambda_i)$	- $n(\lambda_i)$ (in total atmospheric column)
- $k(\lambda_i)$	- $k(\lambda_i)$ (in total atmospheric column)
	- $C_V(h)$ (vertical distribution)
Retrievals provided for total, fine and coarse modes.	*Retrievals provided for total, fine and coarse modes.*

* Azimuth angles, for sky radiances in the almucantar geometry, relative to sun (in degrees): 3.0, 3.5, 4.0, 5.0, 6.0, 7.0, 8.0, 10.0, 12.0, 14.0, 16.0, 18.0, 20.0, 25.0, 30.0, 35.0, 40.0, 45.0, 50.0, 60.0, 70.0, 80.0, 90.0, 100.0, 110.0, 120.0, 140.0, 160.0, 180.

4.2 Structure of different error parameters analysis

As was already mentioned, GRASP has the capability to provide the full covariance matrix of the retrieval errors and this study is aimed to evaluate and illustrate the efficiency of these estimated covariance matrices. At the same time, retrieval error evaluations in most practical applications rely on consideration of mainly diagonal elements of the covariance matrices while non-diagonal elements of covariance matrices are much less used. Indeed, in spite of the fact that non-diagonal elements of covariance matrices provide valuable and interesting information about retrieval errors correlations, these non-diagonal elements are not often available in practice and the analysis of error correlations requires more sophisticated considerations compare to straightforward analysis diagonal elements only, and therefore it is less popular. Considering these aspects, in this thesis, as a first step, a more detailed and extensive analysis of the error variances is carried out, as a second step, the usefulness of obtained non-diagonal elements is illustrated.

The performance of GRASP error variances estimated provided by Eqs. 3.52-3.53 is studied using a series of numerical tests. Figure 4.1 illustrates the general scheme of these tests organization.

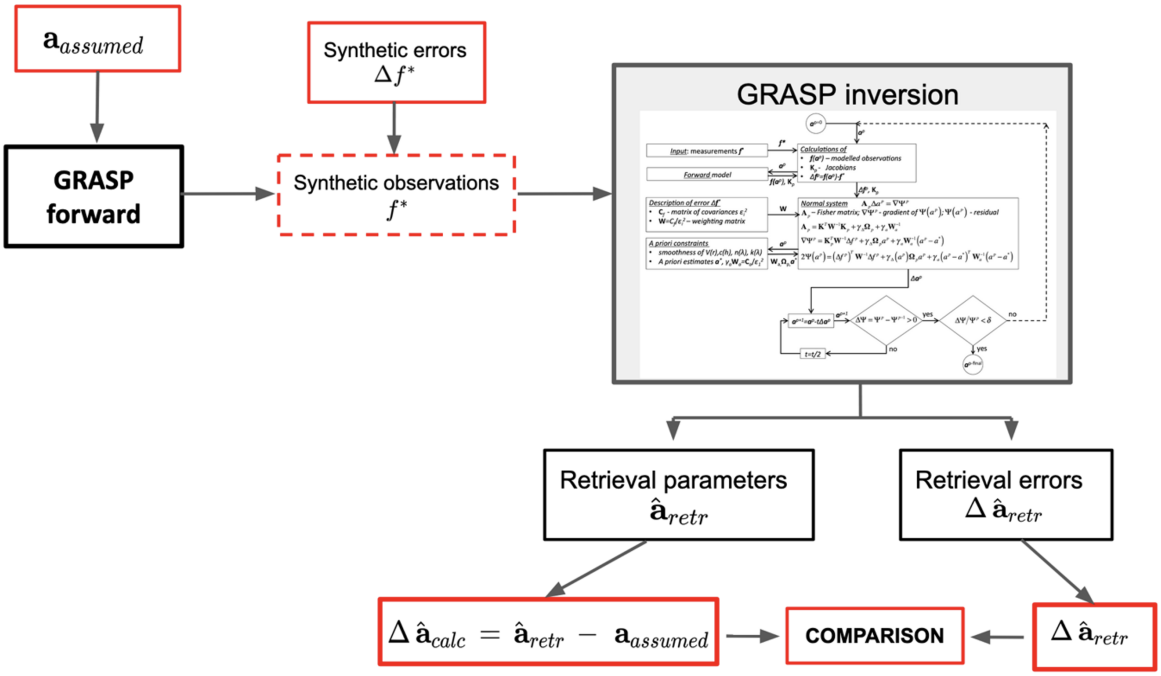


Figure 4.1: General scheme for the validation of the error estimates.

First, as showed in Fig. 4.1, the parameters $\mathbf{a}_{assumed}$ for assumed detailed aerosol properties ($dV(r_i)/d\ln r_i$, $n(\lambda_i)$, $k(\lambda_i)$, C_{sph} and $C_V(h)$ in the case of lidar) are used to obtain the synthetic observations using the GRASP forward model. These synthetic

observations include the spectral AOD, sky radiances and range corrected signals (RCS) of lidar. These data are used then in the inversion tests where the aerosol parameters and their errors are estimated from these synthetic observations using GRASP algorithm. In order to study the effects of the different uncertainties both random and systematic errors are added to the synthetic measurements before the inversion, then the retrieved parameters $\hat{\mathbf{a}}_{retr}$ are compares with $\mathbf{a}_{assumed}$ (see Fig. 4.1)). Therefore, from the retrieved parameters, the retrieval errors provided by GRASP algorithm and 'actual' retrieval errors can be compared. Thus, these actual errors are calculated comparing $\mathbf{a}_{assumed}$ and $\hat{\mathbf{a}}_{retr}$ as follow:

$$\begin{cases} \Delta\hat{\mathbf{a}}_{abs} = \hat{\mathbf{a}}_{retr} - \mathbf{a}_{assumed} \\ \Delta\hat{\mathbf{a}}_{rel} = \dfrac{\Delta\hat{\mathbf{a}}_{abs}}{\mathbf{a}_{assumed}} \cdot 100\% \end{cases} \tag{4.1}$$

where $\hat{\mathbf{a}}_{retr}$ is the retrieved parameter by GRASP algorithm and $\mathbf{a}_{assumed}$ is the parameter assumed in the input data for generation of the synthetic observation. Equation 4.1 is used for each retrieved parameter including the size distribution value at each size bin, the values of complex refractive index at each wavelength, the values of aerosol vertical profile at each altitude and the values of spherical particle fraction. We also implemented the evaluations of the errors for aerosol SSA, and other parameters, that are not part of the directly retrieved parameter while a function of the retrieved parameters and it is estimated based on $\hat{\mathbf{a}}_{retr}$. Thus, the retrieval error variances estimated by GRASP can be compared with the calculated actual retrieval errors. It should be noted here that we have always verified that the errors of the retrieval realized by GRASP from the 'error free' synthetic data (i.e. with no error specifically added) are negligibly small.

GRASP generated variances of the retrieval errors are evaluated in the presence of random errors and analyzed using a series on the numerical tests conducted for statistically representative set of random error realizations. These results are then summarized for the whole series of the tests by figures and tables. The tests with added systematic errors are discussed for most of separate systematic error type while some overall summaries are also provided.

As was mention before, in addition to the standard deviation the non-diagonal elements of covariance matrices provide additional important inside about retrieval quality. This additional information mainly relates with non-zero correlation coefficients. Therefore, in order to illustrate the correlation structure, in this work is also analyzed the correlation matrix that contains the covariance matrix elements normalized by the respective variances as shown by Eq. 3.59. This study attempts to provide several demonstrations of how the structure of the correlation matrix may help to understand several interesting observations in existing retrieval experience.

4.3 Aerosols models and realizations used in the tests

The synthetic tests were performed for several preselected realizations of aerosol in the atmosphere. These realizations were selected based on extensive experience with aerosol retrieval from Sun/sky radiometer data and their combination with co-located lidar data. It is expected that the selected aerosol realization scenarios are representative of the majority of distinct actual observations of atmospheric aerosols.

Two main observational scenarios are considered:

- single aerosol, such as biomass burning (BB), urban and dust for different aerosol loads $\tau(440) = 0.3$, 0.6 and 0.9; and
- the mixture of dust with BB and with urban (BB-Dust and Urban-Dust). For each mixture, we have selected nine different scenarios that correspond to three different aerosol loads, $\tau(440) = 0.2$, 0.5 and 1.0, where the different cases of partition between fine and coarse mode were: $\tau_f/\tau_c = 4.0$, $\tau_f/\tau_c = 1$ and $\tau_f/\tau_c = 0.25$.

The single aerosol and aerosol mixture observational scenarios are used in generation of synthetic tests with sun/sky photometer-only observations. By considering both single and two aerosol types, in this work is evaluated how the accuracy of the retrieved evolve once larger number of parameters are derived from the same information content. In a contrast, the retrieval based on the synergy between lidar and sun/sky photometer is aimed for the retrieval of the properties of two fine and coarse mode aerosol components, therefore the numerical tests for this type of the retrieval rely on mixed aerosol observation scenario. At the same time, the error estimation is also checked in the case when the joint radiometer and lidar observation of single aerosol are analyzed. The properties of each aerosol type were modeled using the climatology of aerosol retrievals from AERONET observations described by Dubovik et al. (2002b) and Torres et al. (2017). The dynamic climatological model from Mongu (Zambia) was used for BB aerosol, model from GSFC (Maryland, USA) for urban aerosol and model from Solar Village (Saudi Arabia) for dust aerosol. The real refractive index (RRI) and imaginary refractive index (IRI) for $\lambda = 355$ nm, 532 nm and 1064 nm (lidar measurements) were obtained by the extrapolation of the values from Dubovik et al. (2002b) as was suggested by Torres et al. (2017). All the scenarios were simulated assuming a solar zenith angle (SZA) equal to 75 degrees.

The retrieval settings were used similar to those that conventionally used in retrieval of aerosol from AERONET Sun/sky-radiometer observations by Dubovik and King (2000) and from combined observations by Sun/sky-radiometer and lidar by

Lopatin et al. (2013, 2021). Specifically, in the retrievals from Sun/sky-radiometer only observations the size distribution (SD) was simulated using 22 logarithmically equidistant size bins between 0.05 to 15 μm. In retrieval of aerosol mixture from combined observations by Sun/sky-radiometer the size distribution is modelled using 10 logarithmically equidistant bins between 0.05 and 0.58 μm for the fine mode and 15 logarithmically equidistant bins between 0.33 and 15 μm for the coarse mode. Similar approach was employed in the retrieval from Sun/sky-radiometer only observations in attempt when bi-component aerosol model was retrieved.

4.4 Test results

Several tests were realized to evaluate the error estimates reliability and usefulness in the presence of both random and systematic uncertainties for aerosol retrievals from the observations of Sun/sky-radiometer alone and in a combination with lidar. In this section are presented the results for two scenarios: (i) as simpler case when only one type of aerosol present and (ii) more complex case then when two distinct types of aerosol present at the same time. Moreover, the correlation matrices for both scenarios are estimated and their usefulness for understanding retrieval error tendencies optimizing retrieval approach are illustrated in this section.

4.4.1 Random error analysis

In series of these tests, to all inverted the synthetic measurements, we added random noise with standard deviation of $\varepsilon_{\Delta\tau}(\lambda) = 0.01$ for AOD, $\varepsilon_{\frac{\Delta I}{I}}(\lambda) = 5\%$ for radiances in order to model realistic uncertainties of AERONET observations (Holben et al. (1998); Eck et al. (1999); Dubovik et al. (2000); Sinyuk et al. (2020)), and $\varepsilon_{355} = 0.2$, $\varepsilon_{532} = 0.15$ and $\varepsilon_{1064} = 0.1$ for lidar attenuation measurements that vary with the altitude as explained by Lopatin et al. (2013, 2021).

4.4.1.1 Retrieval of single aerosol component from radiometer measurements

This section describes the evaluation of the error estimates assuming presence of only one type of aerosol: BB, urban or dust. As we mentioned before, the retrieval aerosol properties under assumption of the presence of single aerosol type composed by homogeneous particles is well-established approach for deriving detailed aerosol properties from ground-based observations by Sun/sky-radiometer that is adapted in operational AERONET retrievals by Dubovik and King (2000). The detailed error analysis of AERONET inversion aerosol product was provided by Dubovik et al. (2000),

Torres et al. (2017) and by the recent study of Sinyuk et al. (2020) that described the uncertainty approach adapted in for AERONET Version 3 retrieval products.

Figures 4.2 - 4.4 illustrate the error variances estimated by GRASP for all retrieved aerosol parameters in the selected synthetic tests for observation of BB, urban and dust with different aerosol loads: $\tau(440) = 0.3$, 0.6 and 0.9. The displayed error bars standard deviation are calculated from the diagonal elements of the covariance matrix (Eq. 3.57). Some tendencies can be seen from these illustrations. For example, the errors of SD in the extremes (for the largest and smallest particles) are the biggest. This is an expected tendency since these particles have typically a lower contribution to the measured signal (radiances and aerosol optical depths) compared to the particles of intermediate radius.

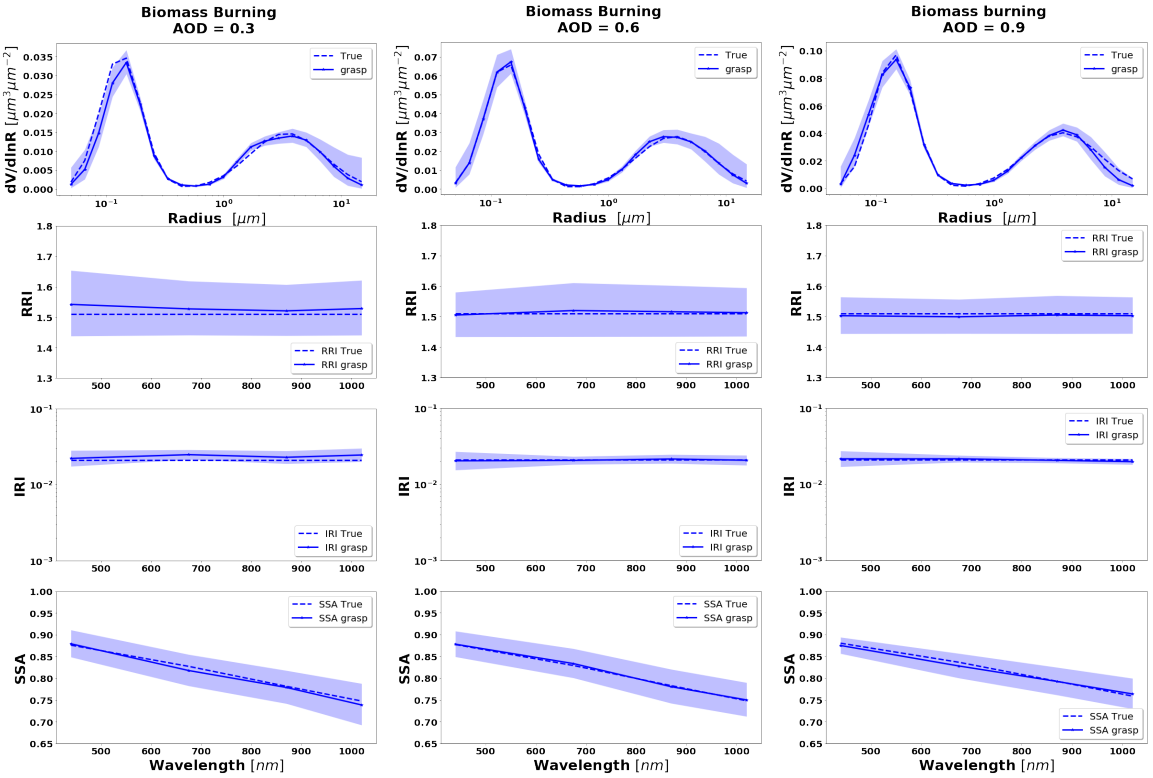


Figure 4.2: Aerosol properties retrieved from simulated sun/sky photometer data with random noise added for BB aerosol for $\tau(440) = 0.3$, 0.6 and 0.9 (left to right). The dashed lines are the simulated properties (SD, RRI, IRI and SSA), the solid lines are the retrieved parameters. The shaded area indicates error estimated by GRASP algorithm.

The retrievals improve and the error decrease when the aerosol load increases, specially for IRI and SSA (absorption information). For BB and urban, SSA error increases with the wavelength. On the other hand, SSA error decreases with the wavelength for dust. This is an expected behavior since the scattering efficiency is more pronounced at short wavelengths for small particles while it is somewhat increasing

with wavelength for large particles. Furthermore, as shown in Fig. 4.2a, the observed underestimation in the SD fine mode seems to be related with an overestimation in RRI.

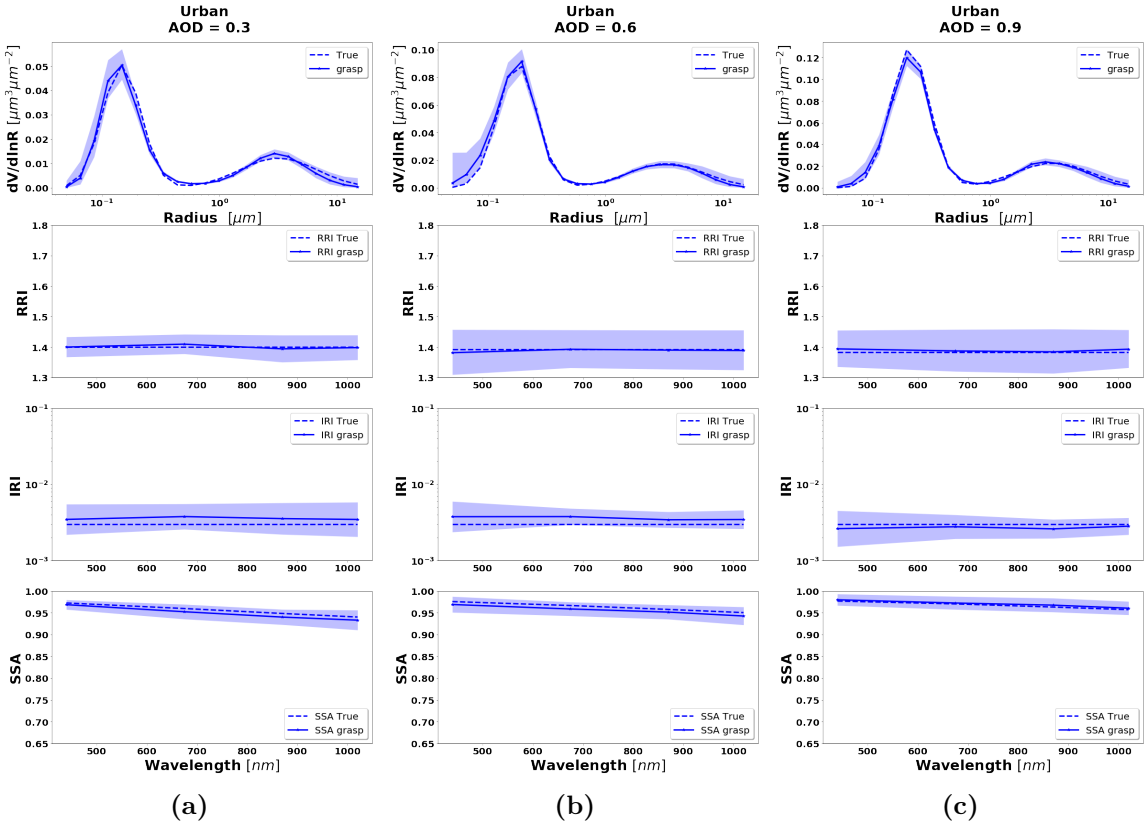


Figure 4.3: Aerosol properties retrieved from simulated sun/sky photometer data with random noise added for urban aerosol for $\tau(440) = 0.3$, 0.6 and 0.9 (left to right). The dashed lines are the simulated properties (SD, RRI, IRI and SSA), the solid lines are the retrieved parameters. The shaded area indicates error estimated by GRASP algorithm.

To evaluate the error estimates in presence of random noise, a set of the simulations for 300 different realizations of noise modelled using random numbers generator has been analyzed in this work. The results of such numerical tests conducted with statistically representative set of random errors which are summarized and illustrated using boxplots of the errors as demonstrated in Fig. 4.5 for SSA(675) values. In the upper part of the figure, the box represents 50% of the data with the whiskers representing $5th$ and $95th$ percentiles of the data, the solid line in the boxplot representing the median and the points are the mean values.

Figure 4.6 shows the distributions of the error estimates provided by GRASP (in red) and the calculated errors (in blue) for the cases when $\tau(440) = 0.6$ and the following convergence criteria are satisfied: $\Delta\tau \leq 0.01$ and $\Delta I/I \leq 5\%$. It can be seen that overall the error estimates provided by GRASP show a capture quite well the 'actual' error tendencies with some overestimation of their values. Thus, the retrieved

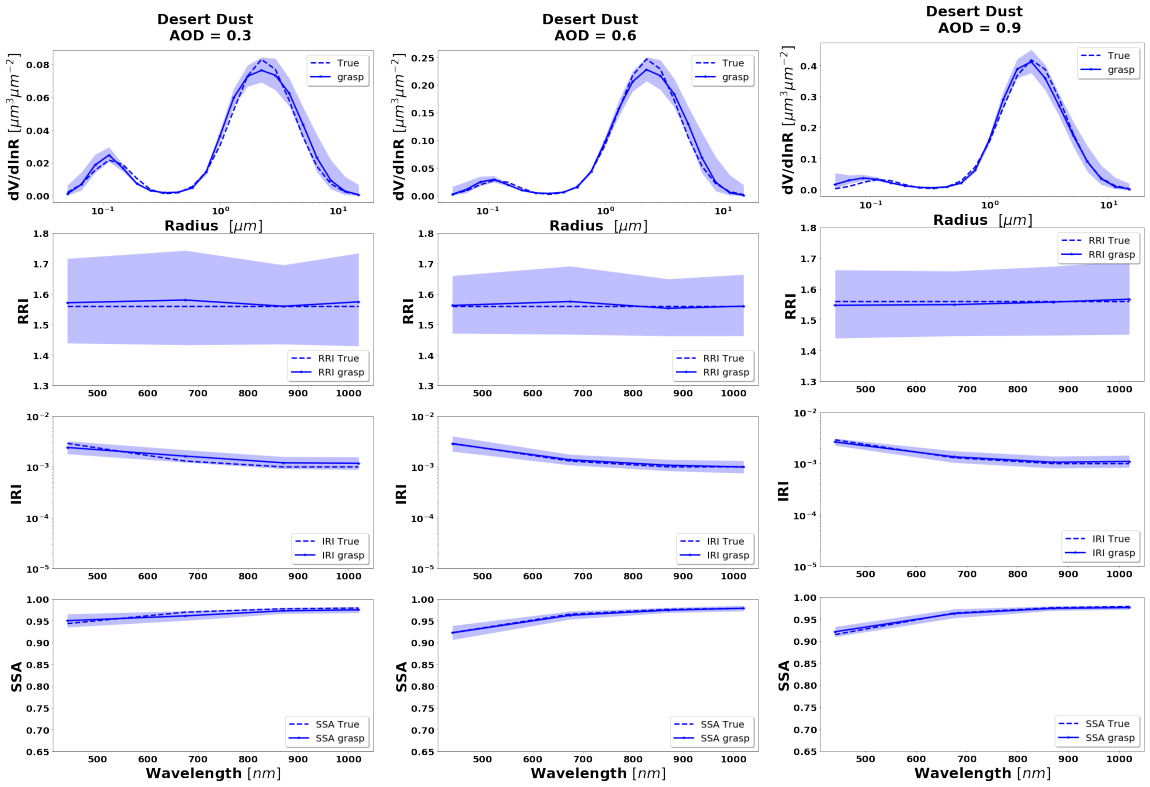


Figure 4.4: Aerosol properties retrieved from simulated sun/sky photometer data with random noise added for dust aerosol for $\tau(440) = 0.3$, 0.6 and 0.9 (left to right). The dashed lines are the simulated properties (SD, RRI, IRI and SSA), the solid lines are the retrieved parameters. The shaded area indicates error estimated by GRASP algorithm.

errors can be considered as upper estimates of actual errors. This observed general overestimation can be, at least partially, explained by the fact that the error estimates by Eqs. 3.52 and 3.53 rely on linear approximation. In this respect it is known from practice that non-linear effects often lead to some saturation while that cannot be captured by linear estimates. Some interesting tendencies can be appreciated in the obtained illustrations. For example, errors in SSA increase with the wavelength for BB and urban, and decreases for dust.

On the other hand, the RRI errors to be similar at the different wavelengths. This is likely related to the fact that spectrally RRI retrievals rely on rather strong smoothness constraints on spectral variability of RRI (e.g. Dubovik and King (2000)). In contrast to the RRI, a large variability of the calculated errors is observed in the distribution of the errors for the IRI. Indeed, in order to capture possible real spectral variability of IRI as that of dust (e.g. see Dubovik et al. (2002b)) the IRI is retrieved under milder smoothness constraints on spectral variability (see Dubovik and King (2000)). Some of the fore-mentioned and other tendencies in the retrieval errors will be further discussed and evaluated in the section which deals with error correlation matrices

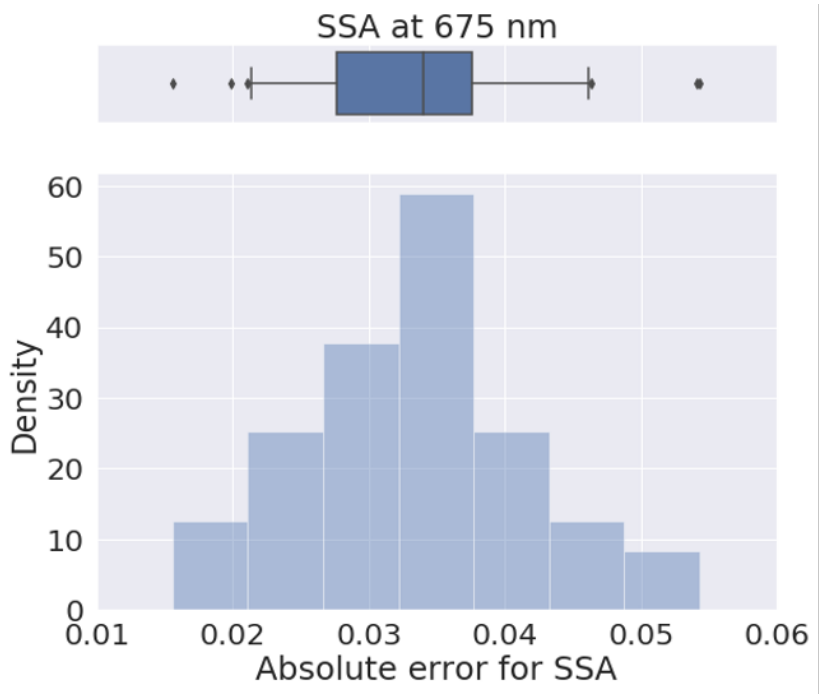


Figure 4.5: The comparison of the variance SSA(675) values estimated by GRASP algorithm with actual errors obtained for extensive tests with randomly added modeled errors. Upper panel: the box represents 50% of the data with the whiskers representing $5th$ and $95th$ percentiles of the data and the solid line in the boxplot representing the median value.

(Section 4.4.3.1).

Table 4.2 summarizes the evaluation of the error estimates represented in the boxplots. It provides the mean values for each parameter (RRI, IRI and SSA) at different wavelengths. These values correspond to the situation with a solar zenith angle equal to 75 degrees. The obtained estimates compare reasonably with the corresponding values provided by the Table 4 of the paper by Dubovik et al. (2006). Specifically, RRI error at $440\,nm$ provided by GRASP for BB is 0.079 (0.04), where the values in parenthesis are from Dubovik et al. (2000) and for urban it is 0.056. The IRI error at $440\,nm$ is 24% (30%) for BB, 54.1% for urban and 24.4% (50%) for dust. The values of the SSA errors 0.028 (0.03) for BB, 0.013 for urban and 0.014 (0.03) for dust. At the same time, RRI error at $440\,nm$ provided by GRASP for dust it is 0.201 (0.04) is quite different, though Dubovik et al. (2000) considered only spherical particles. Moreover, the error estimates for SSA are consistent with the U27 estimates provided by Sinyuk et al. (2020). For example, at $440\,nm$ for AOD = 0.6 the corresponding value for GSFC is 0.017 while grasp provides values of 0.013; for Mongu is 0.023 while grasp provides error of 0.028.

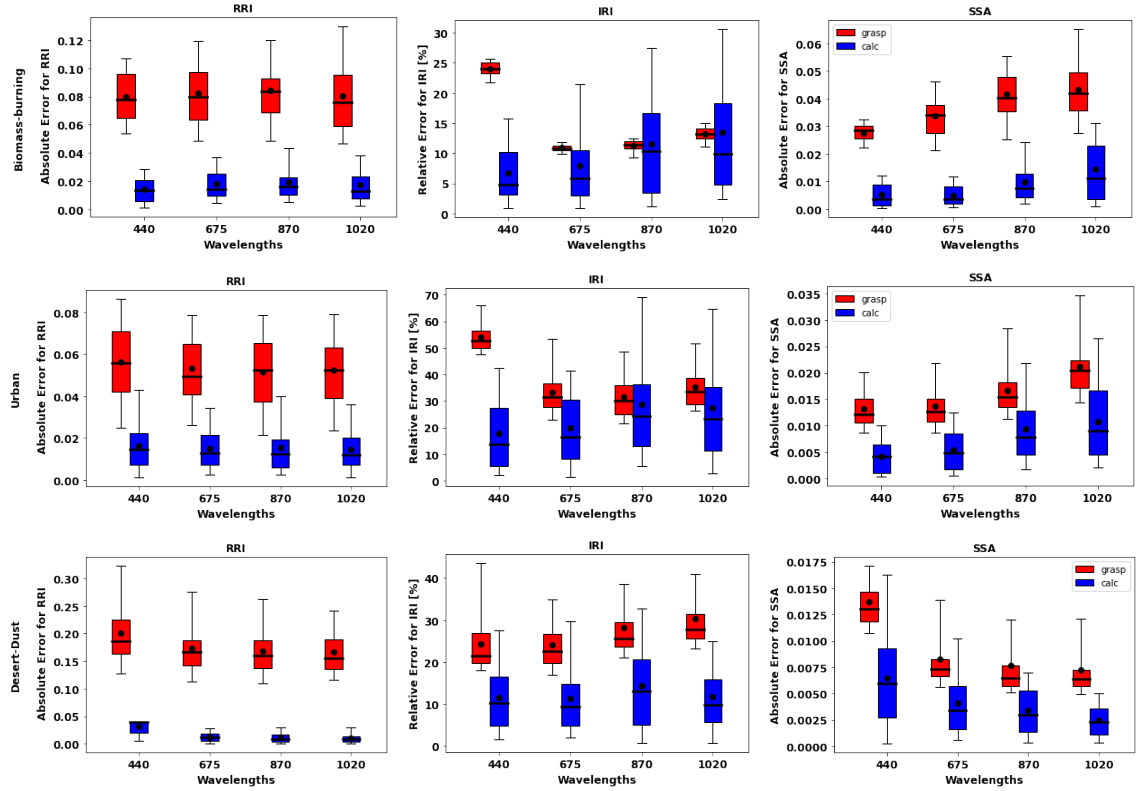


Figure 4.6: The comparison of estimated and actual error distributions for spectrally dependent aerosol parameters retrieved from sun/sky photometer simulated measurements (a case with $\tau(440) = 0.6$). The distributions were obtained using 300 realizations of added random errors. The median values of the errors are shown by a line in the boxplot along with the $25 - 75th$ percentiles indicate by a box and $5 - 95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

4.4.1.2 Retrieval of mixed aerosol properties from measurements of radiometer only

As it was already mentioned, most conventional aerosol retrievals from ground-based radiometer measurements (e.g., Dubovik and King (2000); Nakajima et al. (2020)) assumes that aerosol is represented by homogeneous polydisperse particles with the size independent refractive index. At the same time, this condition is not always correct in reality. Moreover, it is likely somewhat incorrect in majority of the cases. Dubovik et al. (2000) showed in that cases the retrieval assuming homogeneous particles would provide effective index of refraction that allows to reproduce the scattering properties of mixed aerosol rather adequately. Nonetheless, the assumption of homogeneous particles is often questioned and revisited (Xu et al. (2015)), therefore considerations of aerosol inhomogeneity is included in present studies also. In this regard, while the retrieval of the multi-component aerosol is not a part of the standard AERONET inversion, GRASP algorithm allows the retrieval of several aerosol components from diverse

Table 4.2: Errors provided by GRASP for the RRI, IRI and SSA are represented by the mean values of each boxplot for their respective wavelength. Absolute errors are given for RRI and SSA, and relative errors for IRI. Mean values of actual errors are provided in parenthesis.

	BB			**Urban**			**Dust**		
	RRI	**IRI**[%]	**SSA**	**RRI**	**IRI**[%]	**SSA**	**RRI**	**IRI**[%]	**SSA**
440	0.079	24.0	0.028	0.056	54.1	0.013	0.201	24.4	0.014
	(0.014)	(6.66)	(0.005)	(0.016)	(17.6)	(0.004)	(0.03)	(11.52)	(0.006)
675	0.082	10.9	0.033	0.053	33.3	0.014	0.17	24.2	0.008
	(0.018)	(7.8)	(0.005)	(0.015)	(19.8)	(0.005)	(0.013)	(11.36)	(0.004)
870	0.084	11.29	0.042	0.051	31.55	0.017	0.17	28.2	0.008
	(0.019)	(11.61)	(0.009)	(0.015)	(28.57)	(0.009)	(0.011)	(14.3)	(0.003)
1020	0.081	13.17	0.043	0.052	35.1	0.021	0.167	30.4	0.007
	(0.017)	(13.52)	(0.014)	(0.015)	(27.3)	(0.011)	(0.011)	(11.7)	(0.002)

remote sensing observations including the case of aerosol retrieval from radiometer measurements only.

At the same time, since the retrieval of multi-component aerosol from radiometer only is not often used and not employed for operational retrievals, the tests in this section are limited only to several illustrations and no statistical evaluation is performed. The illustrations are produced for the observations of a mixture of Urban-Dust and BB-Dust (see Section 4.3) for three cases of total $\tau(440) = 0.2$, 0.5 and 1.0. In particular, we illustrate the case for $\tau(440) = 1.0$, with $\tau_f = 0.8$ and $\tau_c = 0.2$, $\tau_f = \tau_c = 0.5$ and $\tau_f = 0.2$ and $\tau_c = 0.8$, anticipating more potential for adequate retrieval of multi-component aerosol since the effect of AOD errors decreases for higher AOD. The analysis is focused on possibilities of the differentiation between the properties of fine and coarse aerosol mode parameters such as complex refractive indices, size distributions, single scattering albedo.

Figures 4.7 and 4.8 illustrate results of bi-component retrievals and their error estimates from observations of Sun/sky-radiometer of mixed aerosol. In the same figures, we also show a zoomed plot for the effective RRI and IRI and the total SSA with their errors. Several retrieval tendencies are evident from the figures. For example, in the presence of one mode dominated in optical thickness, the retrievals and error estimates of dominating component are more accurate. For example, in Fig. 4.7c for $\tau_f = 0.2$ and $\tau_c = 0.8$, the retrievals of the coarse mode properties are more accurate. An opposite behavior can be seen in Fig. 4.7a for $\tau_f = 0.8$ and $\tau_c = 0.2$ when the predominance is in the fine mode. The clear trend can be observed in spectral dynamic of the error values for SSA: the error increases with the wavelengths in the fine mode

and decreases for the coarse mode.

The most obvious difficulties in the separation of modes are evident when the properties of each mode are not very different. For example, the such situation can be seen for IRI of Urban-Dust mixture (Fig. 4.7) and for RRI of BB-Dust mixture (Fig. 4.8). In such situations the error variances of each parameter are large and likely correlated (more details provided in discussion of covariance matrices). However, it is very important to note that while the discrimination of some parameters of each component separately is not evident, most of the total and effective properties (zoomed plots) can be estimated rather accurately.

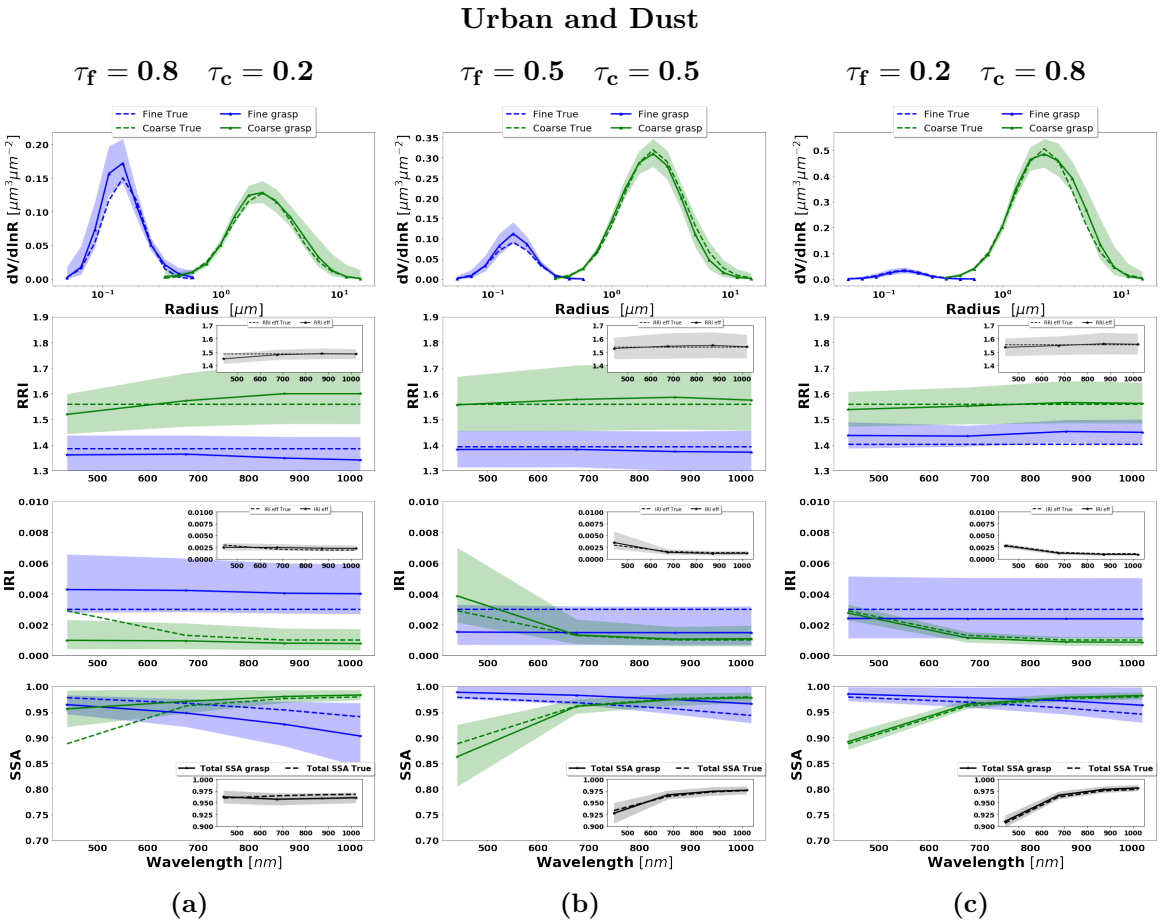


Figure 4.7: Aerosol properties retrieved from simulated sun/sky photometer data with random noise added for a mixture of Urban-Dust aerosols. The dashed lines indicate the simulated properties (SD, RRI, IRI and SSA), the solid lines are the retrieved parameters. The shaded areas indicate error estimated by GRASP algorithm. The zoomed plots represent the effective refractive index and total SSA.

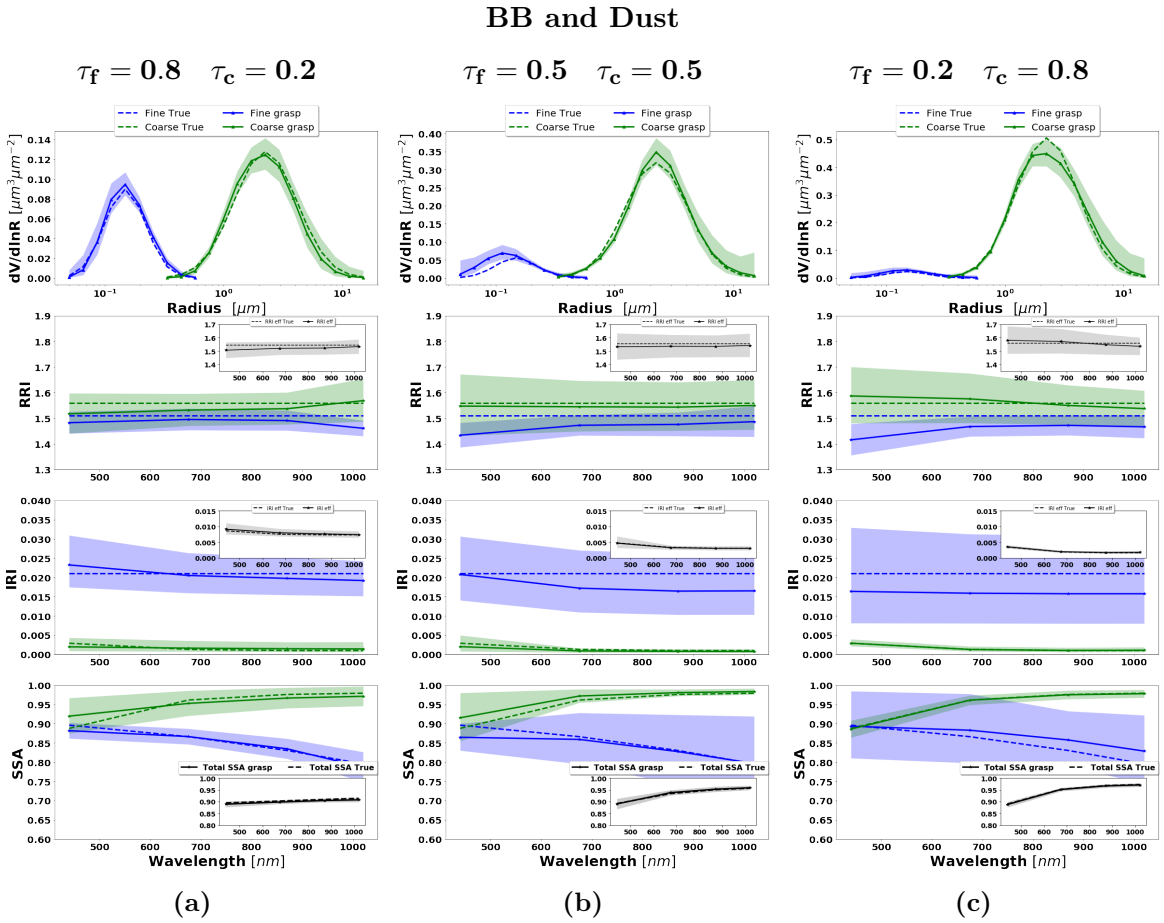


Figure 4.8: Aerosol properties retrieved from simulated sun/sky photometer data with random noise added for a mixture of BB-Dust aerosols. The dashed lines indicate the simulated properties (SD, RRI, IRI and SSA), the solid lines are the retrieved parameters. The shaded areas indicate error estimated by GRASP algorithm. The zoomed plots represent the effective refractive index and total SSA.

4.4.1.3 Retrieval of mixed aerosol properties from measurements of radiometer in combination with lidar

The GRASP aerosol retrieval from were combined Sun/sky radiometer and lidar observations were always designed for retrieval of bi-component aerosol (Lopatin et al. (2013) and the approach is employed for operation processing in frame of ACTRIS activities. Therefore, the evaluation of the random error effect in the aerosol retrieval from radiometer and lidar observations of aerosol mixtures include both analysis of the selected illustrations and the statistically representative series of numerical tests with random errors. The considered synthetic data include synthetic observations produced for the same examples of aerosols mixture (Urban-Dust and BB-Dust) as used in the Section 4.4.1.2.

Figures 4.9 to 4.12 illustrate the retrievals and their error estimates obtained for aerosol properties of both fine and coarse aerosol modes. The good agreement of

actually retrieved parameters (solid lines) with the assumed values (dashed lines) can be seen for all cases. From comparison of Figs. 4.7-4.8 with Figs. 4.9-4.10, it is easy to see that the retrieval error estimate is lower when lidar data also used. The improvements (compared to the results from radiometer only retrievals) are especially evident in the separation of the retrieved aerosols properties, especially when the contribution of the aerosol load is lower (this will be shown in Section 4.4.2 when also bias is assumed). At the same time, the total properties are accurately estimated in both retrieval scenarios.

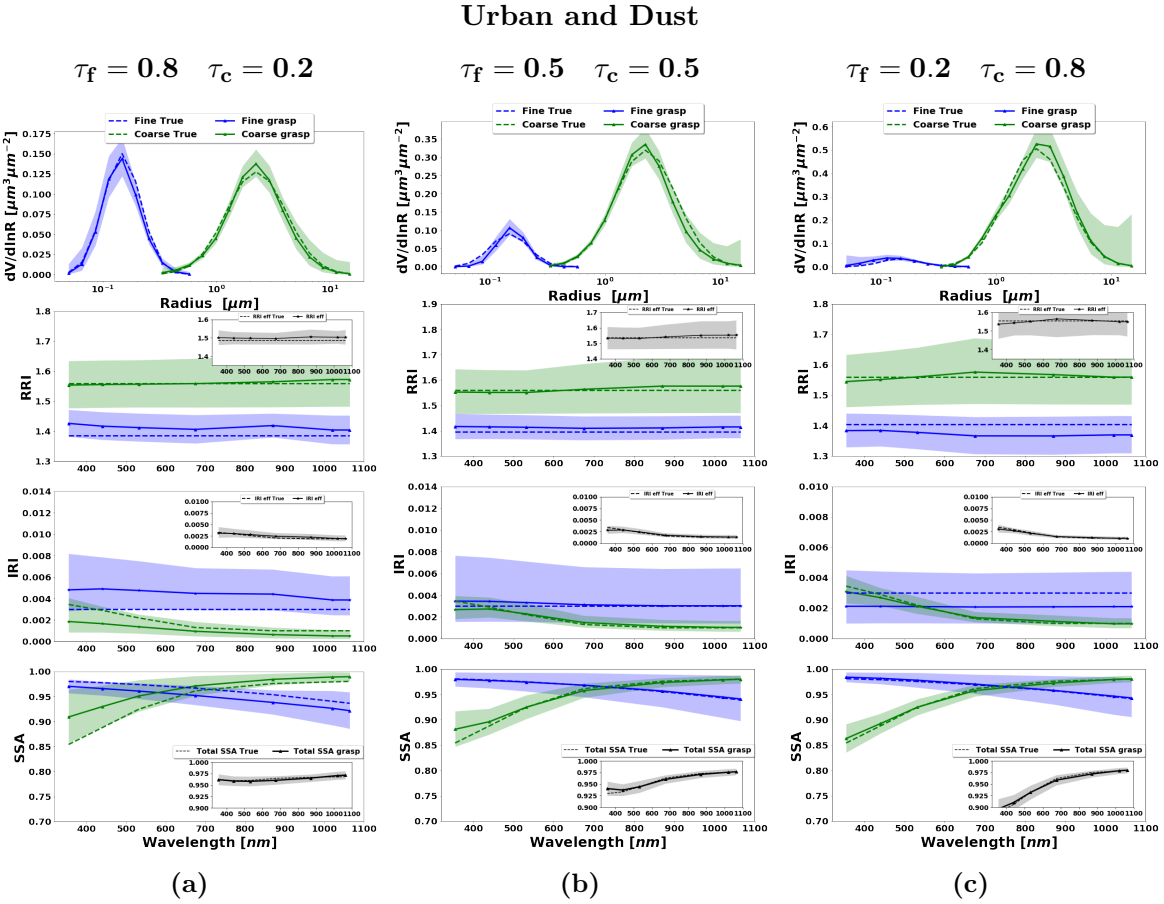


Figure 4.9: Aerosol properties retrieved from simulated sun/sky photometer and lidar data with random noise added for a mixture of Urban-Dust aerosols. The dashed lines indicate the simulated properties (SD, RRI, IRI and SSA), the solid lines are the retrieved parameters. The shaded areas indicate error estimated by GRASP algorithm. The zoomed plots represent the effective refractive index and total SSA.

Figure 4.11 shows the lidar ratio of fine mode, coarse mode and total aerosol for the three cases aforementioned. In general, good agreements of retrieved and assumed values are obtained, specially for the total LR. However, there are some discrepancies at short wavelengths for fine mode lidar ratios.

Figure 4.12 illustrates the retrieval of aerosol vertical profile for each case. The agreement between retrieved and assumed values of vertical profiles is good mainly for the coarse mode at the altitudes where it has maximum and dominates. At the

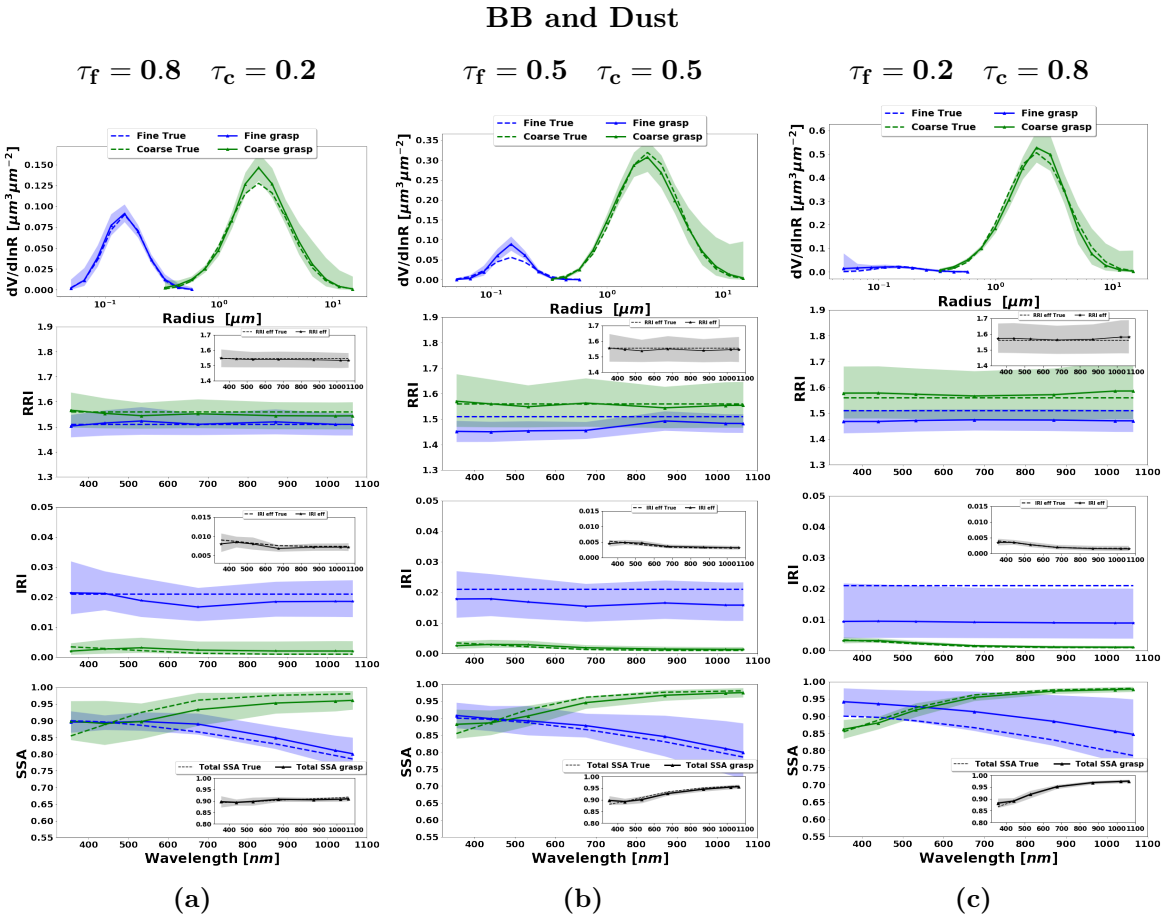


Figure 4.10: Aerosol properties retrieved from simulated sun/sky photometer and lidar data with random noise added for a mixture of BB-Dust aerosols. The dashed lines indicate the simulated properties (SD, RRI, IRI and SSA), the solid lines are the retrieved parameters. The shaded areas indicate error estimated by GRASP algorithm. The zoomed plots represent the effective refractive index and total SSA.

altitudes where there is a superposition of aerosol layers with comparable presence of both aerosols the retrieval struggles to discriminate the contribution of both modes and a clear overestimation of fine mode (and a consequently underestimation of coarse mode) can be seen.

In order to evaluate the error estimates in presence of random errors, a set of simulations with adding 300 realization of random noise values is analyzed. Figures 4.13 to 4.20 show the comparisons of all retrieved aerosol parameters separately for fine and coarse aerosol modes. In addition, the retrieval of total SSA and LR are shown. The case for total $\tau(440) = 1.0$ is shown more extensively, as in previous Section, due to the interest the retrieval in the situation with higher aerosol loads. The main result that can be gained from illustrations is that GRASP error estimates are typically higher than actual errors; this same result was obtained for the retrieval of only photometer data.

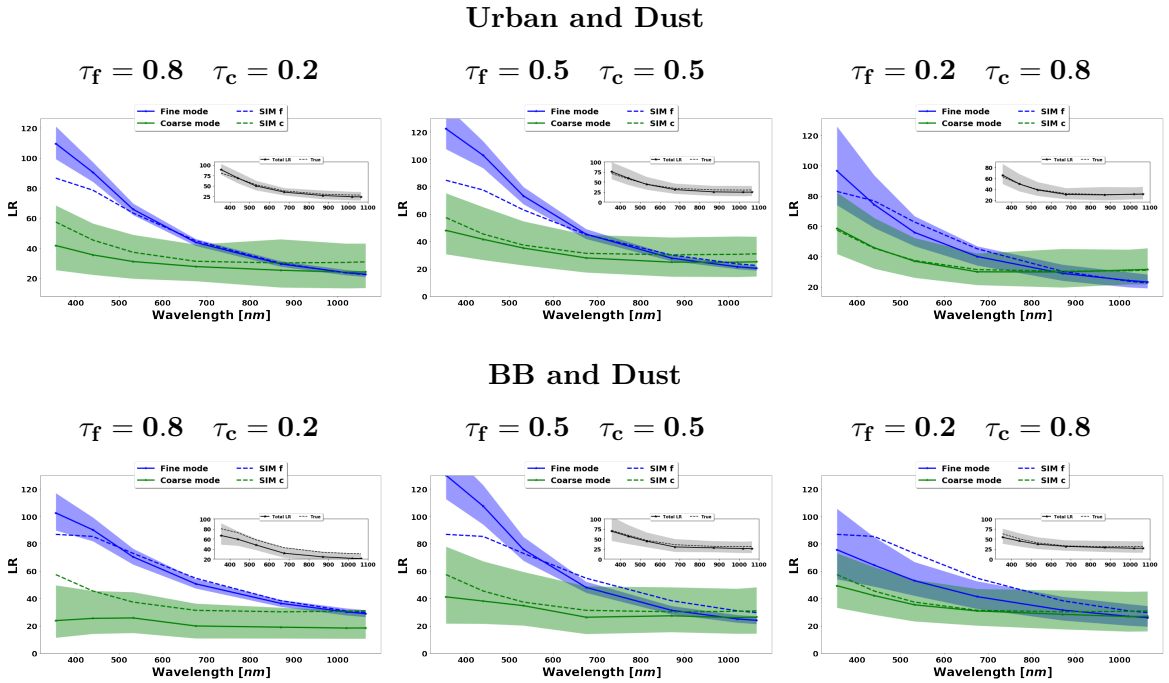


Figure 4.11: The aerosol lidar ratio (LR) retrieved from simulated sun/sky photometer and lidar data with random noise added for a mixture Urban-Dust aerosols (above) and BB-Dust (below). The dashed lines indicate the simulated properties (SD, RRI, IRI and SSA), the solid lines are the retrieved parameters. The shaded areas indicate error estimated by GRASP algorithm. The zoomed plots represent the results for total LR retrievals.

Figures 4.13 and 4.14 illustrate the comparisons of distributions of the GRASP error estimates and actual errors for RRI and IRI of fine and coarse aerosol modes for situations when mixtures of Urban-Dust and BB-Dust are observed. It can be seen that the accuracy of the refractive index retrievals for each mode depends strongly on the contribution of the mode to the signal, as was observed by Lopatin et al. (2013). For example, if we analyse the performance of the fine mode, the higher the contribution of fine optical thickness, the better the accuracy in the retrievals of fine mode aerosol parameters.

Figures 4.15 and 4.16 show the situation for error distribution for SSA of fine mode, coarse mode and total. Similarly as observed in earlier tests, for the retrieval of fine mode parameter, the errors increase with the wavelength while for the retrieval of the coarse mode parameters the errors decrease with the wavelengths. Also, the results show the error in the total SSA are rather small even if the SSA of fine and coarse modes are quite high.

The error evaluation for LR is represented in the Figs. 4.17 and 4.18. In most of the cases we see good agreements between the error estimations and actual error. The only exception is the errors of LR of fine mode at short wavelengths where the actual errors are higher than the errors provided by GRASP. This tendency seems to

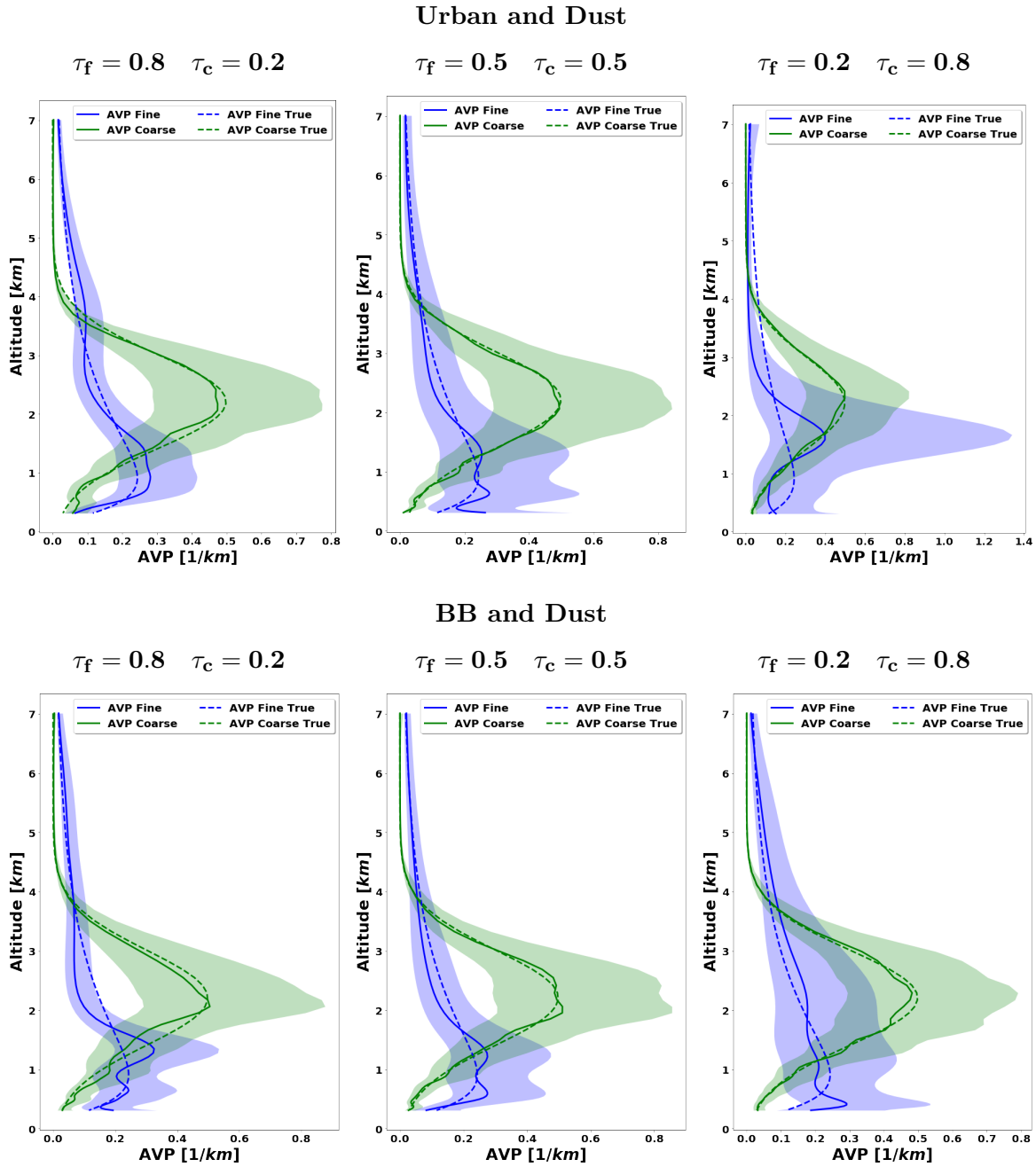


Figure 4.12: The aerosol AVP retrieved from simulated sun/sky photometer and lidar data with random noise added for a mixture of Urban-Dust aerosols (above) and BB-Dust (below). The dashed lines indicate the simulated properties (AVP), the solid lines are the retrieved parameters. The shaded areas indicate error estimated by GRASP algorithm.

be anticorrelated with the results found for the coarse mode LR error estimates, where the GRASP error estimates are notably higher than actual values.

The results illustrated by the figures are summarized in Tables 4.3 to 4.5. These tables show the mean values of the GRASP error estimates for the cases when the total $\tau(440) = 1.0$ and $\tau_f = \tau_c = 0.5$, i.e. when there is no predominance of either

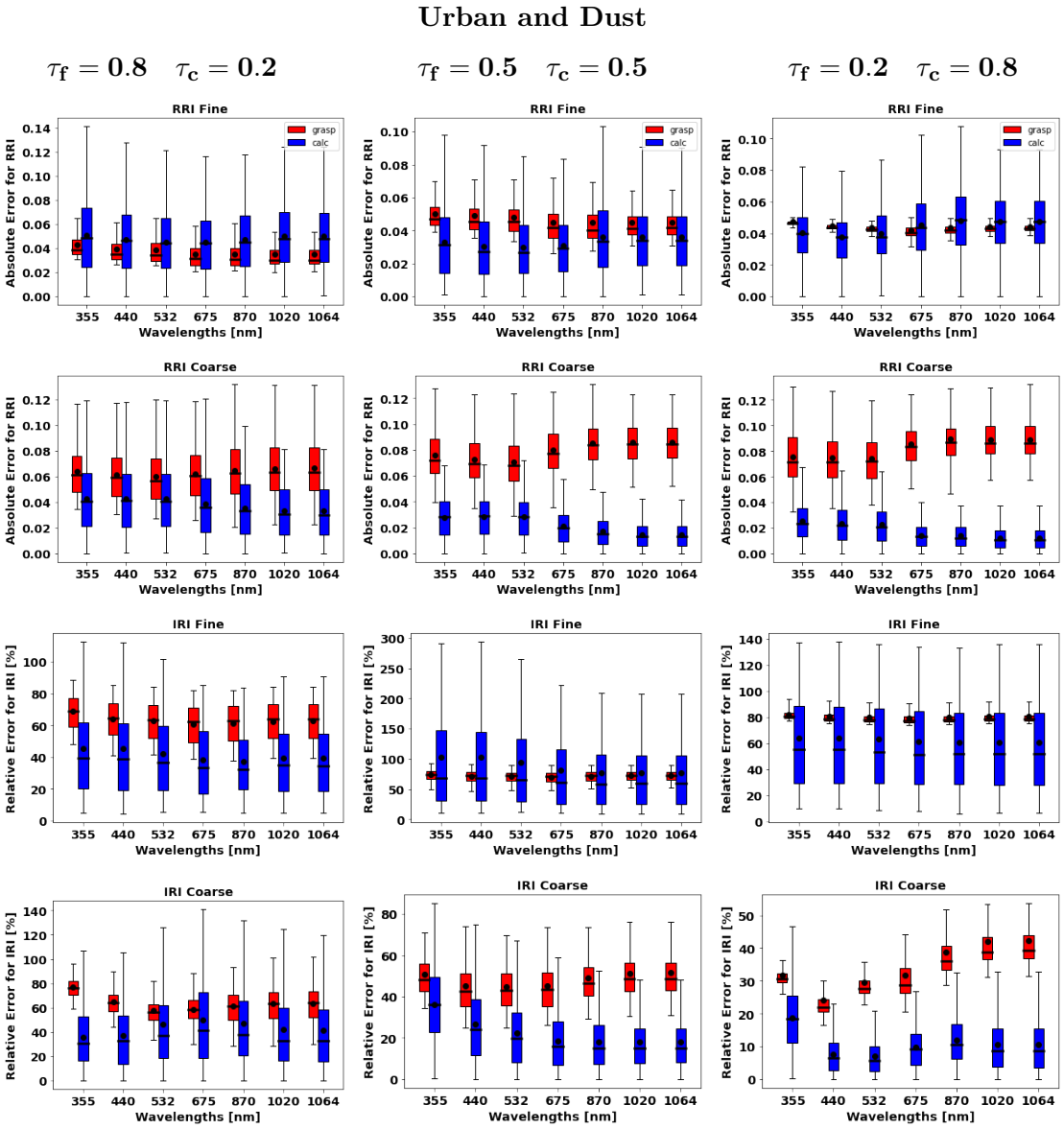


Figure 4.13: The comparison of estimated and actual error distributions for spectrally dependent aerosol parameters retrieved from measurements by sun/sky photometer simulated and lidar for a mixture of Urban-Dust aerosol. The distributions were obtained using 300 realizations of added random errors. The median values of the errors are shown by a line in the boxplot along with the $25-75th$ percentiles indicate by a box and $5-95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

modes. The values are provided for the aerosol parameter considered at different wavelengths both for simulation of Urban-Dust and BB-Dust observations calculated for a case of the SZA= 75 degrees. For retrieval error of fine mode in the case of urban aerosol parameters, the median values for RRI are around 0.05, and the values do not present much variability with the wavelength. For retrieval of IRI, the mean values of the GRASP retrieval errors are at the level of around 73%, showing a pronounced

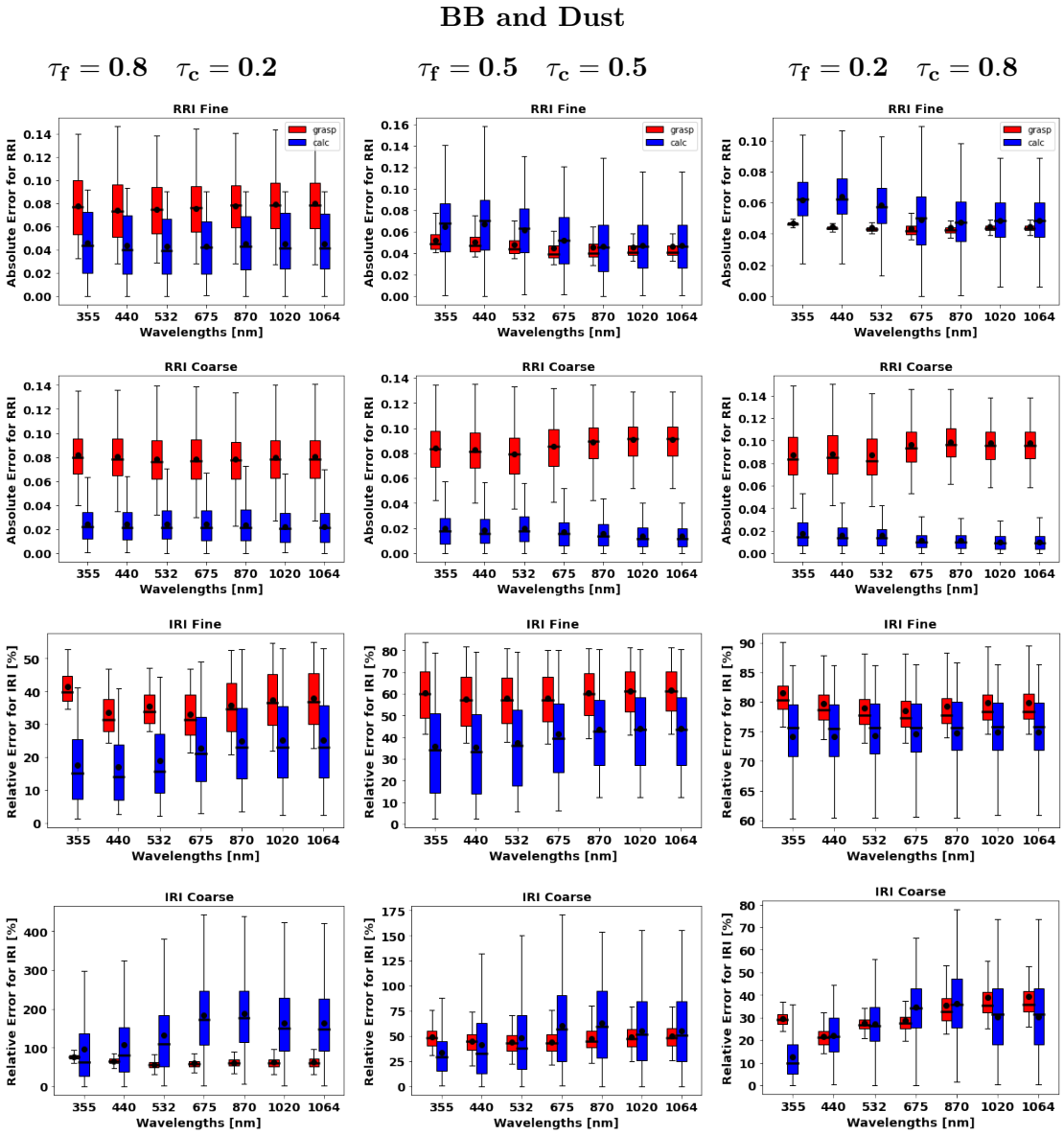


Figure 4.14: The comparison of estimated and actual error distributions for spectrally dependent aerosol parameters retrieved from measurements by sun/sky photometer simulated and lidar for a mixture of BB-Dust aerosol. The distributions were obtained using 300 realizations of added random errors. The median values of the errors are shown by a line in the boxplot along with the $25-75th$ percentiles indicate by a box and $5-95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

underestimation respect to the actual error, at short wavelengths. With respect to SSA errors provided by GRASP a clear tendency is observed: the error increases with the wavelengths from 0.024 to 0.061. Finally, the mean values of LR errors provided by GRASP decrease with the wavelength between 15% to 10%, with notable underestimations respect to the actual errors at short wavelengths. In the case of the retrieval of fine mode BB parameters, mean values for RRI errors provided by GRASP

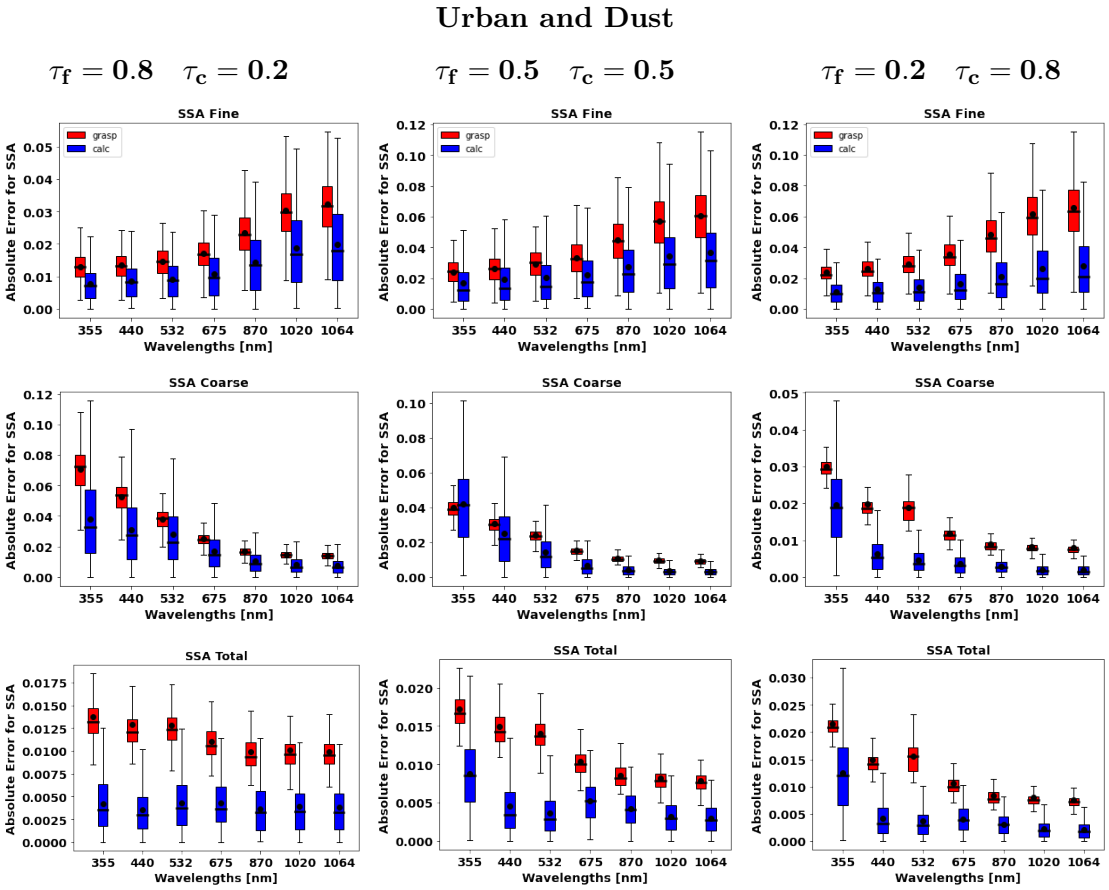


Figure 4.15: The comparison of estimated and actual error distributions for aerosol SSA retrieved from measurements by sun/sky photometer simulated and lidar for a mixture of Urban-Dust aerosols. The distributions were obtained using 300 realizations of added random errors. The median values of the errors are shown by a line in the boxplot along with the $25-75th$ percentiles indicate by a box and $5-95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

are around 0.05. Some underestimations respect to the actual errors are observed at short wavelengths. The mean values for IRI errors are around 60% and the errors for SSA show a clear tendency to increase with the wavelengths between 0.04 to 0.09. Mean values of LR errors provided by GRASP decrease with the wavelength between 18% to 14%.

The mean values of error estimates provided by GRASP for dust present good agreement in case of both mixtures. In general, the mean values for RRI error estimates vary between 0.07 to 0.09 and they do not present much variability with the wavelength, while the smaller values of errors are seen for Urban-Dust mixture case. The mean values of IRI error estimates are around 50%, while for BB-Dust mixture we observe some underestimations of actual errors by GRASP calculations. The errors of SSA show a clear tendency that decrease with the wavelengths between 0.04 to 0.009. The

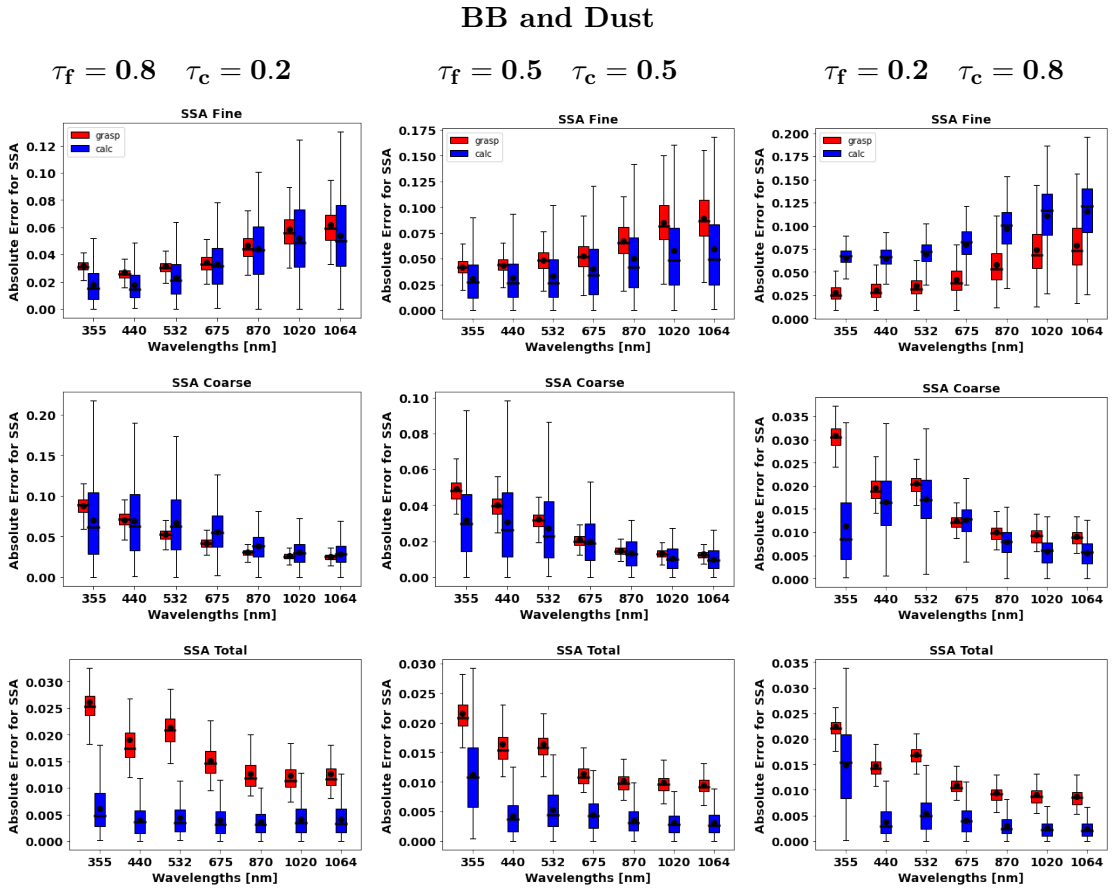


Figure 4.16: The comparison of estimated and actual error distributions for aerosol SSA retrieved from measurements by sun/sky photometer simulated and lidar for a mixture of BB-Dust aerosols. The distributions were obtained using 300 realizations of added random errors. The mean values are represented by the black dot and the median values of the errors are shown by a line in the boxplot along with the $25-75th$ percentiles indicate by a box and $5-95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

mean values for LR retrievals increase with the wavelength from 37% to 60%, with bigger errors observed for BB-Dust mixture.

Once again, it is important to note that the errors of the parameters characterizing total aerosol are generally accurately estimated. For both cases of Urban-Dust and BB-Dust mixtures, the mean values of total SSA error estimates vary between 0.02 to 0.009 and the mean values of total LR error estimates are in the range of 23% to 55%.

Figures 4.19 and 4.20 show the relative errors of AVP retrievals for fine and coarse aerosol modes for Urban-Dust and BB-Dust aerosols mixture. The errors estimated by GRASP are a bit higher that the errors obtained by simulations of random errors, correspondingly the GRASP errors can be safely used as upper estimates of actual retrieval uncertainties. Table 4.5 summarizes the evaluation of the errors estimates for all the scenarios discussed above. The GRASP estimates of the retrieval errors for both

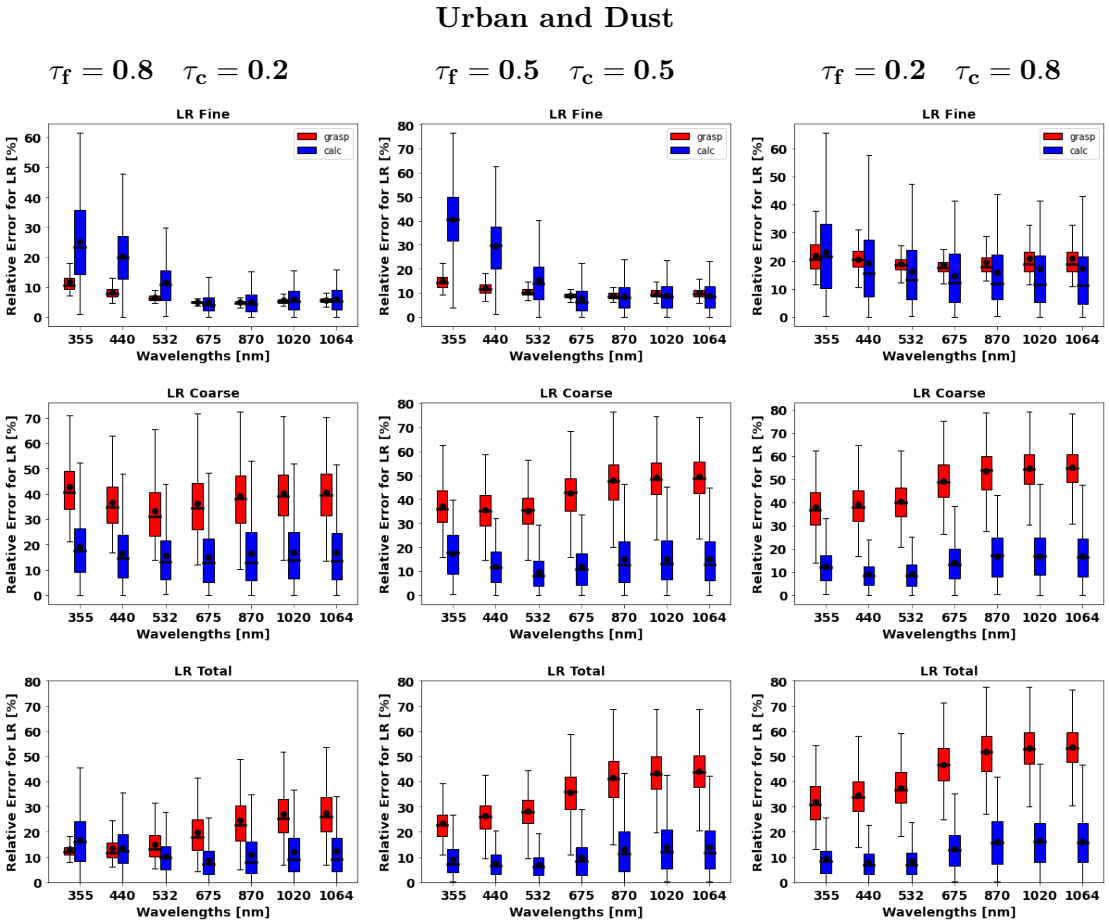


Figure 4.17: The comparison of estimated and actual error distributions for aerosol LR retrieved from measurements by sun/sky photometer simulated and lidar for a mixture of Urban-Dust aerosols. The distributions were obtained using 300 realizations of added random errors. The mean values are represented by the black dot and the median values of the errors are shown by a line in the boxplot along with the $25-75th$ percentiles indicate by a box and $5-95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

mixtures are between $50-70\%$ for the fine mode and $50-57\%$ for the coarse mode.

Finally, a lower sensitivity to the retrieval of fine mode properties can be observed as a clear tendency in the evaluation of the retrieval errors for the cases when mixed aerosols are analyzed. In particular, quite high errors were obtained for the complex refractive index. Then, these errors consequently propagate to the errors of other optical properties such as SSA of fine mode, as was found in the earlier study by Lopatin et al. (2013).

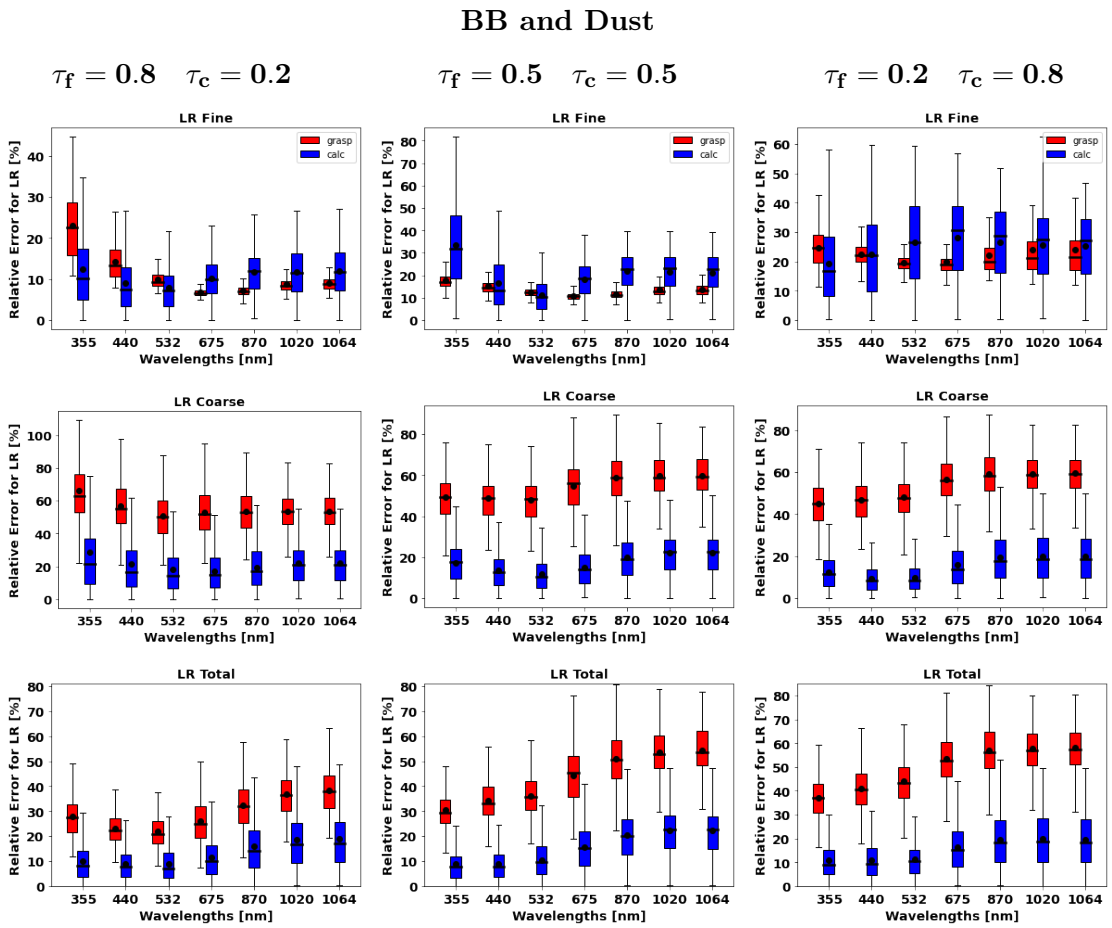


Figure 4.18: The comparison of estimated and actual error distributions for aerosol LR retrieved from measurements by sun/sky photometer simulated and lidar for a mixture of BB-Dust aerosols. The distributions were obtained using 300 realizations of added random errors. The mean values are represented by the black dot and the median values of the errors are shown by a line in the boxplot along with the $25-75th$ percentiles indicate by a box and $5-95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

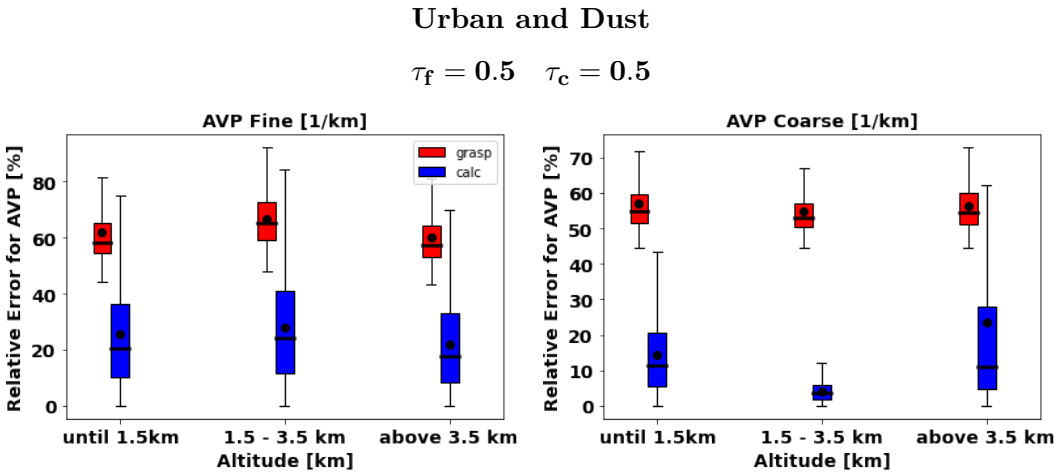


Figure 4.19: The comparison of estimated and actual error distributions for AVP retrieved from measurements by sun/sky photometer simulated and lidar for a mixture of Urban-Dust aerosols. The distributions were obtained using 300 realizations of added random errors. The mean values are represented by the black dot and the median values of the errors are shown by a line in the boxplot along with the $25-75th$ percentiles indicate by a box and $5-95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

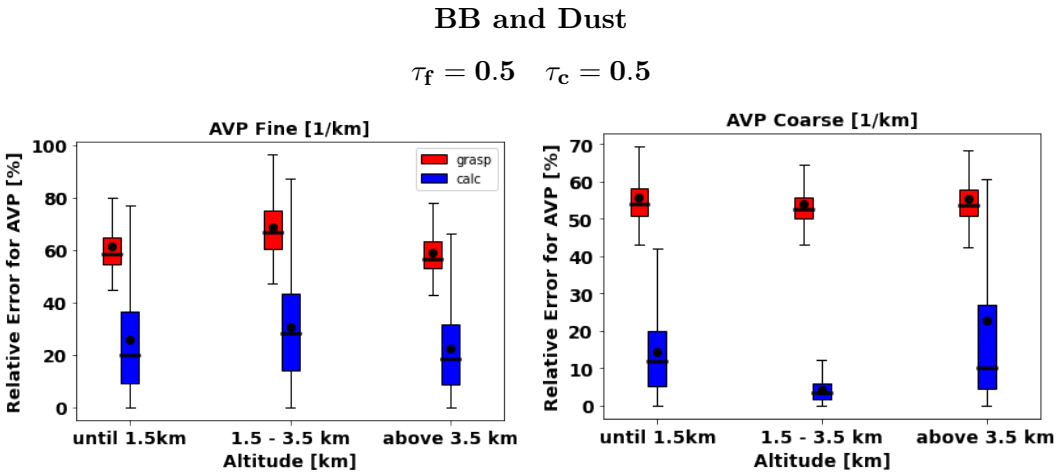


Figure 4.20: The comparison of estimated and actual error distributions for AVP retrieved from measurements by sun/sky photometer simulated and lidar for a mixture of BB-Dust aerosols. The distributions were obtained using 300 realizations of added random errors. The mean values are represented by the black dot and the median values of the errors are shown by a line in the boxplot along with the $25-75th$ percentiles indicate by a box and $5-95th$ percentiles indicated using whiskers. The red color shows the error estimates provided by GRASP and the blue shows the calculated actual errors (Eq. 4.1).

Table 4.3: The mean values of RRI, IRI, SSA and LR retrieval errors estimated by GRASP for the synthetic test for a mixture of Urban-Dust aerosol mixture. The mean values represent distributions obtained using 300 realizations of added random errors for the situation with total $\tau(440) = 1.0$ with $\tau_f = \tau_c = 0.5$ and SZA = 75°. The absolute errors are provided for RRI and SSA, and relative errors for IRI and LR. Mean values for the actual errors are provided in parenthesis.

	Urban-Dust						
	355 nm	**440 nm**	**532 nm**	**675 nm**	**870 nm**	**1020 nm**	**1064 nm**
$\mathbf{RRI}_f$	0.050	0.049	0.048	0.045	0.045	0.045	0.045
	(0.033)	(0.030)	(0.029)	(0.031)	(0.036)	(0.036)	(0.036)
$\mathbf{RRI}_c$	0.076	0.073	0.071	0.080	0.085	0.086	0.086
	(0.028)	(0.028)	(0.028)	(0.021)	(0.017)	(0.015)	(0.015)
$\mathbf{IRI}_f$[%]	73.31	71.18	70.67	70.33	71.34	72.18	72.26
	(103.03)	(103.4)	(94.19)	(81.46)	(76.62)	(76.78)	(76.79)
$\mathbf{IRI}_c$[%]	50.59	45.00	44.58	45.12	49.01	51.22	51.51
	(36.17)	(26.45)	(22.32)	(18.50)	(18.23)	(17.92)	(17.89)
$\mathbf{SSA}_f$	0.024	0.026	0.029	0.033	0.045	0.057	0.061
	(0.017)	(0.019)	(0.020)	(0.022)	(0.027)	(0.034)	(0.037)
$\mathbf{SSA}_c$	0.039	0.031	0.024	0.015	0.011	0.009	0.009
	(0.042)	(0.025)	(0.014)	(0.007)	(0.004)	(0.003)	(0.003)
$\mathbf{SSA}_T$	0.017	0.015	0.014	0.010	0.008	0.008	0.008
	(0.009)	(0.004)	(0.004)	(0.005)	(0.004)	(0.003)	(0.003)
$\mathbf{LR}_f$[%]	14.93	12.11	10.41	8.88	8.92	9.79	9.96
	(40.45)	(29.39)	(15.37)	(7.58)	(8.44)	(8.96)	(8.63)
$\mathbf{LR}_c$[%]	37.04	35.71	35.23	42.39	47.77	48.99	49.32
	(17.42)	(12.01)	(9.48)	(11.81)	(14.90)	(15.17)	(14.99)
$\mathbf{LR}_T$[%]	23.31	26.46	28.21	35.61	41.45	43.36	43.97
	(9.01)	(7.48)	(6.85)	(9.71)	(13.3)	(13.9)	(13.9)

Table 4.4: The mean values of RRI, IRI, SSA and LR retrieval errors estimated by GRASP for the synthetic test for a mixture of BB-Dust aerosol mixture. The mean values represent distributions obtained using 300 realizations of added random errors for the situation with total $\tau(440) = 1.0$ with $\tau_f = \tau_c = 0.5$ and SZA = 75°. The absolute errors are provided for RRI and SSA, and relative errors for IRI and LR. Mean values for the actual errors are provided in parenthesis.

	BB-Dust						
	355 nm	**440 nm**	**532 nm**	**675 nm**	**870 nm**	**1020 nm**	**1064 nm**
$\mathbf{RRI}_f$	0.052	0.050	0.048	0.044	0.045	0.046	0.046
	(0.064)	(0.067)	(0.061)	(0.052)	(0.046)	(0.047)	(0.047)
$\mathbf{RRI}_c$	0.084	0.083	0.079	0.085	0.089	0.091	0.091
	(0.019)	(0.019)	(0.019)	(0.017)	(0.016)	(0.014)	(0.014)
$\mathbf{IRI}_f$[%]	60.51	57.61	57.91	58.17	60.33	61.48	61.59
	(35.65)	(35.40)	(37.56)	(41.63)	(43.74)	(44.14)	(44.13)
$\mathbf{IRI}_c$[%]	48.75	44.57	43.81	43.89	47.43	49.56	49.87
	(33.75)	(41.39)	(48.20)	(60.36)	(62.61)	(55.55)	(55.44)
$\mathbf{SSA}_f$	0.041	0.044	0.048	0.052	0.067	0.085	0.089
	(0.031)	(0.031)	(0.033)	(0.040)	(0.050)	(0.057)	(0.059)
$\mathbf{SSA}_c$	0.049	0.040	0.032	0.021	0.015	0.013	0.013
	(0.032)	(0.031)	(0.027)	(0.019)	(0.013)	(0.010)	(0.009)
$\mathbf{SSA}_T$	0.022	0.016	0.016	0.011	0.010	0.009	0.009
	(0.011)	(0.004)	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)
$\mathbf{LR}_f$[%]	17.99	15.17	12.53	10.74	11.75	13.48	13.82
	(33.43)	(16.46)	(11.29)	(18.05)	(21.76)	(21.43)	(21.30)
$\mathbf{LR}_c$[%]	49.38	48.73	47.91	54.82	58.57	59.54	59.76
	(17.24)	(13.65)	(11.60)	(15.04)	(19.94)	(22.16)	(22.23)
$\mathbf{LR}_T$[%]	30.30	34.11	36.09	44.35	50.82	53.51	54.18
	(8.83)	(9.03)	(10.52)	(15.65)	(20.38)	(22.13)	(22.14)

Table 4.5: The mean values of AVP retrieval errors estimated by GRASP for the synthetic test for a mixture of Urban-Dust aerosol mixture. The mean values represent distributions obtained using 300 realizations of added random errors for the situation with total $\tau(440) = 1.0$ with $\tau_f = \tau_c = 0.5$ and SZA $= 75°$. The shown relative errors for AVP [1/km] are represented by the mean values for three layers. Mean values for the actual errors are provided in parenthesis.

		until 1.5 km	**1.5-3.5 km**	**above 3.5 km**
Urban-Dust	**AVP**$_f$ [%]	61.91	66.90	59.95
		(25.79)	(27.73)	(22.06)
	AVP$_c$ [%]	57.25	54.91	56.34
		(14.43)	(4.11)	(23.51)
BB-Dust	**AVP**$_f$ [%]	61.46	68.50	59.17
		(25.80)	(30.53)	(22.28)
	AVP$_c$ [%]	55.74	53.90	55.30
		(14.36)	(4.08)	(22.91)

4.4.2 The analysis of the retrieval in presence of the systematic uncertainties

In the Section 4.4.1 was presented the evaluation and validation of the errors of the different aerosol properties considering propagation of the random noise from measurements into retrieval. The analysis confirmed rather satisfactory performance of the approach adapted in GRASP for the estimation of retrieval errors in the presence of random noise. This section discusses the approach for estimating possible contributions of the systematic errors in the retrieval uncertainties. In principle, each retrieval methodology assumes that there is no systematic uncertainties neither in measurements nor in the used forward model. If any systematic bias is identified it is corrected in measurements or in their interpretation. However, in practice the systematic uncertainties may remain unidentified and make significant contribution in the retrieval uncertainties.

As mentioned above, in Eq. 3.53 the apparent misfit was used as an indicator of bias, however in real situations not all biases can be seen from the misfit. Thus, in this section the results are presented considering a possible solution to this problem. Therefore, commonly the contribution of potential bias is included in the estimation of the retrieval errors (e.g. see Dubovik et al. (2000), Sinyuk et al. (2020)). Using similar logic, in the present methodology was added an extra term, Eq. 3.56, that accounts for propagation of possible bias from the measurements. The propagation is accounted for linear approximation in similar manner as the systematic term in Eq. 3.53 accounts for bias from misfit. Thus, this section analyses the potential effect of realistic biases and their overall importance for reliable estimations of the retrieval errors in practice.

The potential effect of the systematic errors is analyzed in series of the numerical tests with possible assumed systematic errors. Following previous studies by Dubovik et al. (2000), Torres et al. (2017) and Sinyuk et al. (2020), in ground-based photometric and radiometric data, we consider two types of potential main biases in measured AOD and sky-radiances. These biases could be originated from miss-calibrations of direct Sun or diffuse sky sensors (Eck et al. (1999)). The biases are assumed wavelength independent, since spectral systematic deviations are easier to identify in direct analysis of observation, they are likely to be manifested in misfit and may compensate each other influence on the retrievals. Specifically, two possible levels of biases nominal and maximum are considered as follows:

- in AOD, nominal bias of ± 0.01 and maximum bias of ± 0.02, and
- in radiances nominal bias of $\pm 3\%$ and maximum bias of $\pm 5\%$.

To evaluate the effects of biases, the above values were added to synthetic direct measurements of AOD and sky-radiance by AERONET like ground-based radiome-

ter. These data were inverted by GRASP code and the retrieved values of aerosol parameters were compared to the values assumed in synthetic simulation as a 'truth'. In addition, the deviations of retrieved values from the 'true' ones are compared to the errors estimates generated by GRASP based on Eq. 3.56 using known values of added biases. In similar manner, the influence of the potential systematic errors in aerosol retrieval from combined observations of ground-based radiometer and lidar was analyzed. In these tests, the biases in lidar attenuation measurements were assumed for each wavelength following the studies by Lopatin et al. (2013, 2021): $\varepsilon_{355} = \pm 0.2$, $\varepsilon_{532} = \pm 0.15$ and $\varepsilon_{1064} = \pm 0.1$. It should be noted that conducted synthetic tests not only allow to verify the accuracy of the systematic error estimates by GRASP and also to analyze the effects of biases on the retrievals for different retrieval scenarios in diverse situations.

4.4.2.1 Effects of measurement bias in retrieval of single aerosol component from radiometer measurements

In this section the study is focused on the analysis of the effects of the biases and on estimating contribution of systematic errors in retrievals of aerosol from ground-based observation by radiometer. In similar manner as in the analysis of random errors, first was consider the observations dominated by two types of aerosols: BB and dust. The effect of measurement biases is expected to be manifested in the situations with low and moderate aerosol loading, therefore, the analysis is focused on the scenarios with AOD(440) = 0.1, 0.3 and 0.6.

Two situations were considered:

- when a single bias in AOD or radiances is present;
- when the biases can be present in both AOD and radiances simultaneously. In this case, the different combinations of positive and negative biases in AOD and radiances are considered.

The estimations of the errors introduced by the biases were calculated as:

$$\sigma^2_{bias} = \sigma^2_{lm} + \sigma^2_{misfit} + \frac{1}{N}\sum_{k=1}^{N}\sigma^2_k \tag{4.2}$$

where σ^2_{lm} corresponds to contributions from systematic errors introduced by the Levenberg-Marquardt procedure and σ^2_{misfit} are the errors manifested by the miss-fit estimated by Eq.3.49, and each σ^2_k is the contribution adding + bias and − bias in the measurements.

Figures 4.21 to 4.24 illustrate the results of the analysis for the different retrieved properties for situation with bias of ±0.01 and ±0.02 in AOD only. These results show

specific effects from AOD bias. The figures have two blocks: on the left the retrievals with added positive bias in AOD and on the right retrievals with negative bias are illustrated. In both cases, the error bars represent the systematic component adding the positive or negative bias respectively. In all the figures, the solid lines show the assumed value of the parameters in the simulation, the dotted lines show the retrieved values and the magnitudes of the estimated bias are shown by the shaded areas. It should be noted that for the case of BB with AOD(440) = 0.1, the results with negative bias are not shown. This is because, the AOD for BB decreases very strongly with the wavelength and for the case of AOD(440) = 0.1, the AOD at 1020 nm is ~ 0.01.

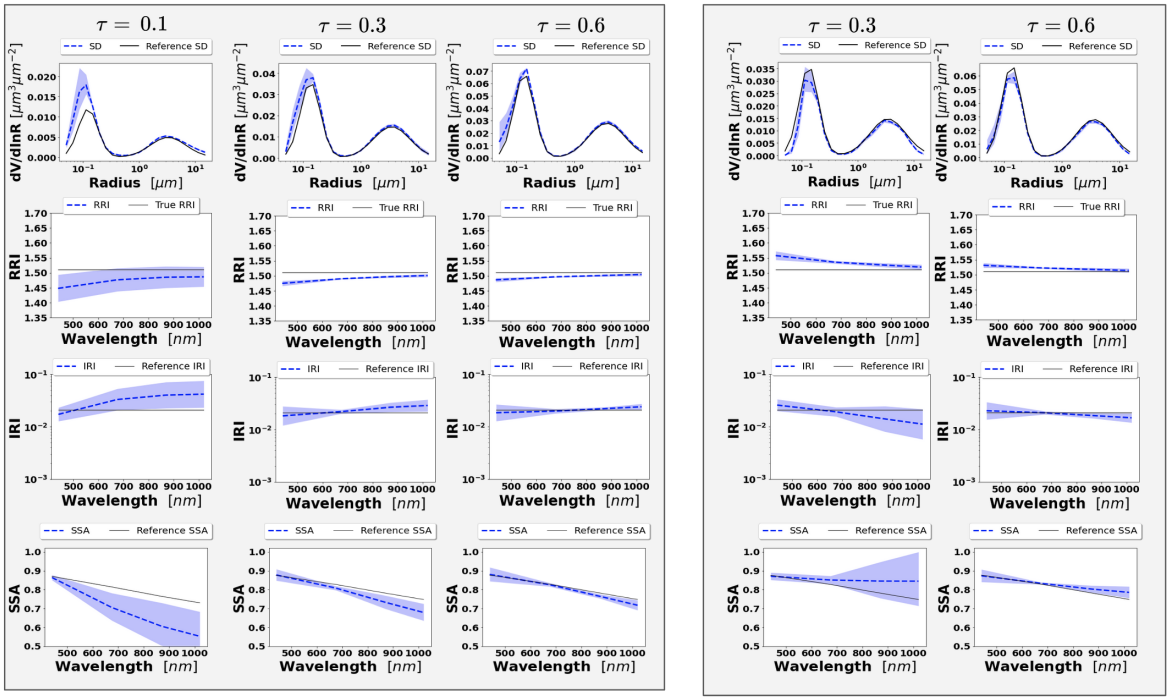


Figure 4.21: Aerosol properties retrieved from simulated sunsky photometer data with assumed bias in AOD simulated data for BB aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +0.01 are represented in the block on the left and negative bias −0.01 in the block on the right. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic errors estimated by GRASP algorithm.

The figures show different and clear tendencies which are in agreement with general expectations and with the tendencies already observed in previous studies by Dubovik et al. (2000) and Torres et al. (2014). For example, it can be seen that bias in AOD most strongly affects the estimate of the parameters characterizing aerosol absorption such as imaginary part of the refractive index and single scattering absorption. This is an anticorrelation: the positive bias results in overestimation of absorption (higher RRI and lower SSA) and the negative in underestimation absorption (lower RRI and higher SSA) respectively. The result was quite expected since radiance values in this first

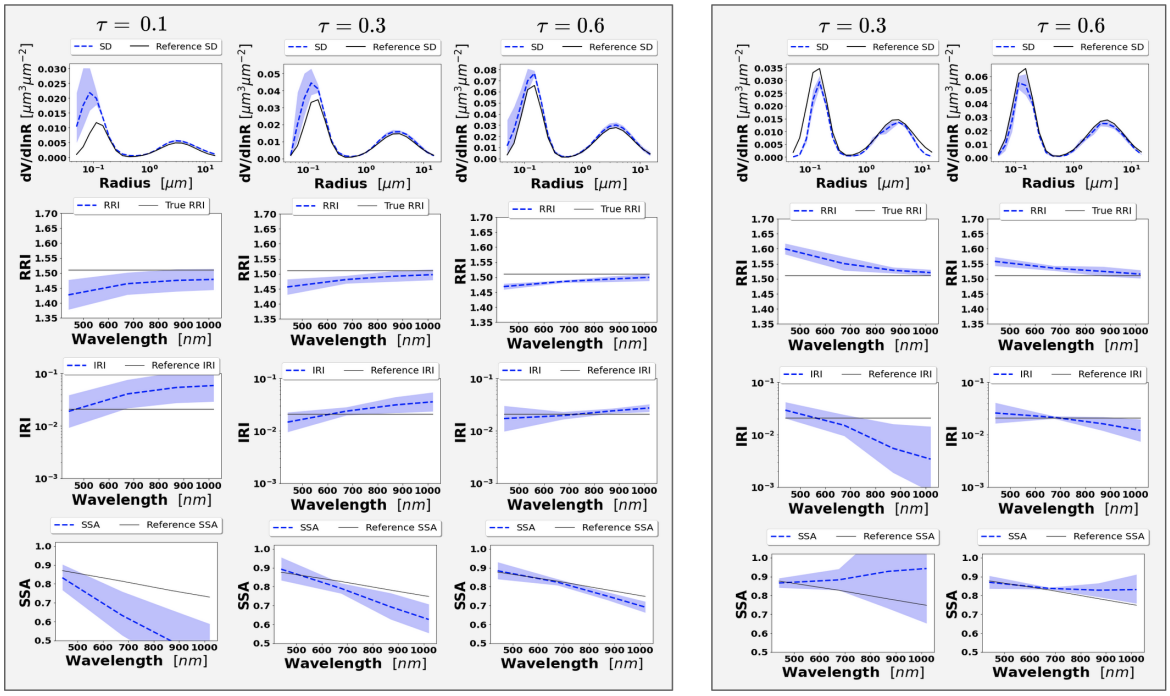


Figure 4.22: Aerosol properties retrieved from simulated sun/sky photometer data with assumed bias in AOD simulated data for BB aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +0.02 are represented in the block on the left and negative bias −0.02 in the block on the right. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic errors estimated by GRASP algorithm.

experience do not vary. Thus, if we keep the scattering component (which is derived from radiances) but we enlarge the extinction component (by enlarging the AOD), necessary the retrieval understands that the absorption should be larger (imaginary part of the refractive index). Conversely, if we reduce the value of extinction the retrieval would reduce the value of absorption. Also, the strongest effect is observed for optically thin situations when a small absolute error in optical thickness becomes comparable with the magnitude of aerosol optical thickness. This is especially clear for BB observations where AOD(1020) is always rather small as earlier discussed also by Dubovik et al. (2000). For observations of dust aerosol, the effect of biases in AOD are significantly smaller than for BB. It can be explained since dust has a small value of Ångstrom exponent and therefore larger values of AOD at longer wavelengths. For the retrieval of the size distribution, the bias in AOD has rather minor effect, though we found a general overestimation for positive BIAS values and underestimation for negative values.

Overall, the estimated systematic error agrees well with actual manifestations of the bias in the retrieval. The quantitative estimations are also quite convincing and shown in Figs. 4.21 and 4.23 for biases of ±0.01. In some cases, some underestimations of the

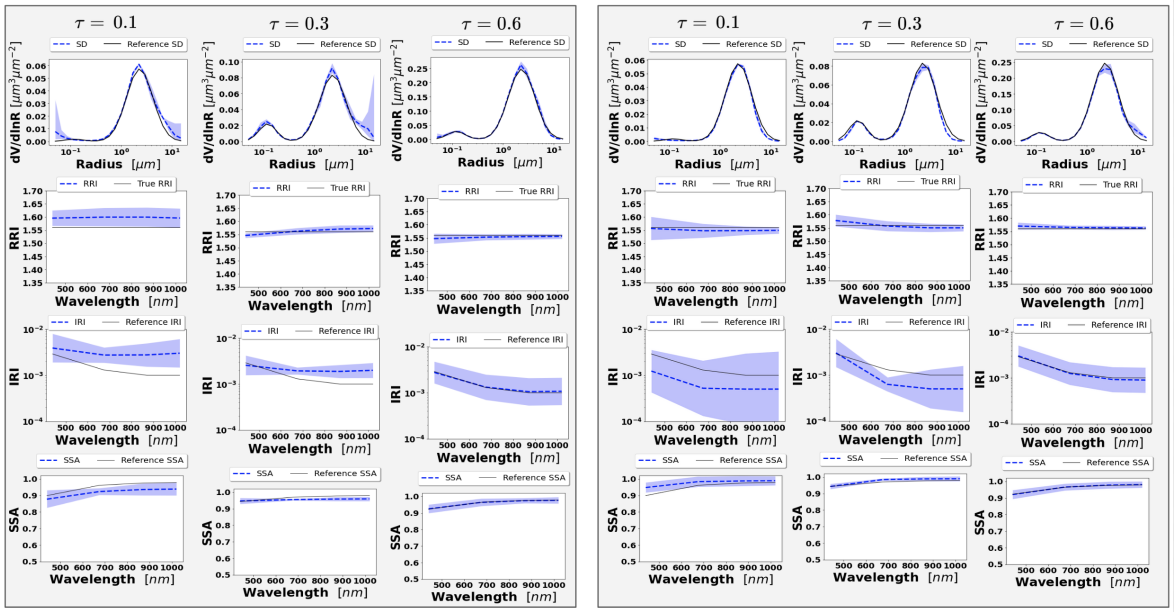


Figure 4.23: Aerosol properties retrieved from simulated sun/sky photometer data with assumed bias in AOD simulated data for dust aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +0.01 are represented in the block on the left and negative bias −0.01 in the block on the right. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic errors estimated by GRASP algorithm.

bias effects can be observed. For example, the largest differences are identified for the case of higher value of bias (±0.02) shown in Figs. 4.22 and 4.24 while a significant increase in the systematic component of the retrieved error is also well captured by the error estimates.

It can be seen that among all considered aerosol parameters, the main differences between the bias effects and the obtained error estimates are observed for Real part of Refractive Index (RRI). In these cases, the bias is not fully covered by the systematic component of the retrieved error results. Similarly, apparent underestimation of RRI errors was also seen by Sinyuk et al. (2020), who attributed these underestimations to different factors as, for example, the effect of not accounted pointing bias. In the present simulations, there is no pointing bias considered, and discrepancy is likely coming from the fact that Eqs. 3.49 and 3.56 relies on the derivatives estimated in the vicinity of the solution and based on linear approximation. Indeed, the dependence of both AOD and radiances scattered by aerosol is very complex and non-linear. Therefore, both taking the derivatives in the vicinity of obtained solution instead of vicinity of 'true values', as well as, non-liner character of AOD and radiances may explain the differences. At the same time, it is important to note that, as can be seen from analysis of the random component of RRI error (Section 4.4.1) the random error effect is likely to dominate over effect of AOD bias, and therefore, the estimation of the total error (described

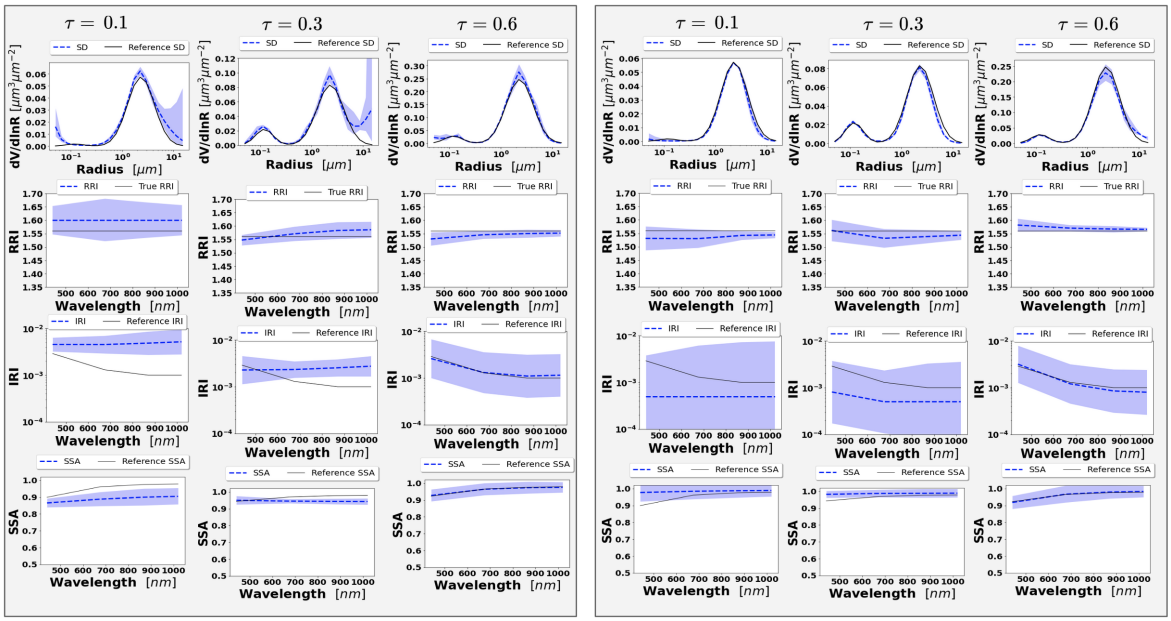


Figure 4.24: Aerosol properties retrieved from simulated sun/sky photometer data with assumed bias in AOD simulated data for dust aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +0.02 are represented in the block on the left and negative bias −0.02 in the block on the right. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic errors estimated by GRASP algorithm.

below in this section) seems to allow us to make an objective and complete observation on this parameter.

Figures 4.25 to 4.28 show the effects of the biases in the radiances of two magnitudes of ±3% and ±5% for the observations of BB and dust. In general, it can be seen from the results that in both cases, BB and dust, the retrievals are less affected by bias in radiances than by the biases in the AOD, even when the bias in radiances is ±5%. Similar tendency was also reported in studies by Dubovik et al. (2000), Torres et al. (2014) and Sinyuk et al. (2020). At the same time, it should be noted that the present analysis is focused on measurement configuration corresponding to solar almucantar the SZA is 75 degrees (see Table 4.1) when the measurements include are taken in a wide range of scattering angles. In this respect, Dubovik et al. (2000) showed that the effect of sky radiance bias increases when the range of observed scattering angles is limited, i.e. in almucantar observation corresponding to SZA less than 60 degrees. Moreover, according to recent tendencies in observational practices, the use of such measurements is limited and most analyses are focused on observational scenarios with sufficient range of observed scattering angles. For example, AERONET start to establish so-called 'hybrid' observational scenario during high SZA times (Giles et al. (2019)). In regard to the performance of the error estimation, the effect of the bias in the sky radiances seems to be well captured by the GRASP error estimates.

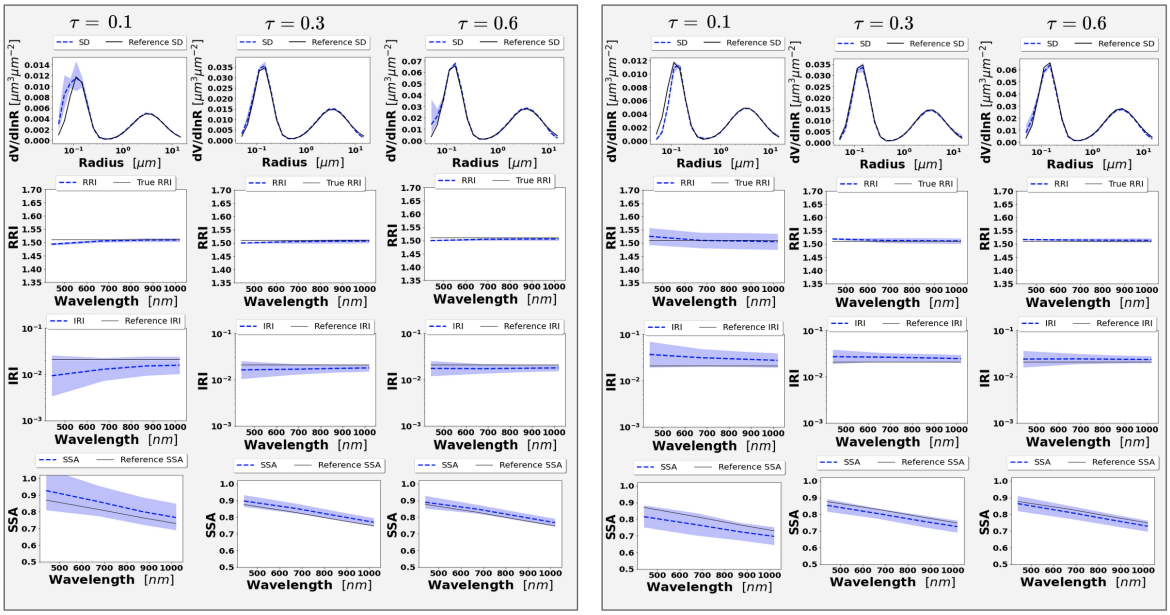


Figure 4.25: Aerosol properties retrieved from simulated sun/sky photometer data with assumed bias in RAD simulated data for BB aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +3% are represented in the block on the left and negative bias −3% in the block on the right. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic errors estimated by GRASP algorithm.

It should be also noted that we observe an anticorrelation between the radiances BIAS and the retrieval of the imaginary part of the refractive index. This effect is opposite to the one observe in the case of AOD and with differences significantly smaller. Thus, when the BIAS are positive (+3% and +5%) there is a decrease in the imaginary part of the refractive index. The fact that the value of AOD remains the same and there is an increase in the value of the scattering is interpreted by the code as a decrease of the aerosol absorption. Conversely, the negative BIAS in the radiances produce an increase in the imaginary part of the refractive index which can be explained by the same reason.

Figure 4.29 shows the results of the analysis of the situation when the systematic biases present in both AOD and radiances are simultaneously assumed. The results for BB are on the left and for Dust on the right. The illustrations are shown for the specific situation with two positive biases: +0.01 in AOD and +5% in radiances. It should be noted, that the tests were produced for both the situations with the biases of the same and opposite signs. This case with the biases of the same signs showed the most interesting results with the strongest manifestation of bias effects and therefore they are presented here. In the situation with biases of opposite signs effects on the retrievals are rather minor due to internal compensations of the influences of the biases. Additionally, the misfit of observations is more pronounced that helps to identify the

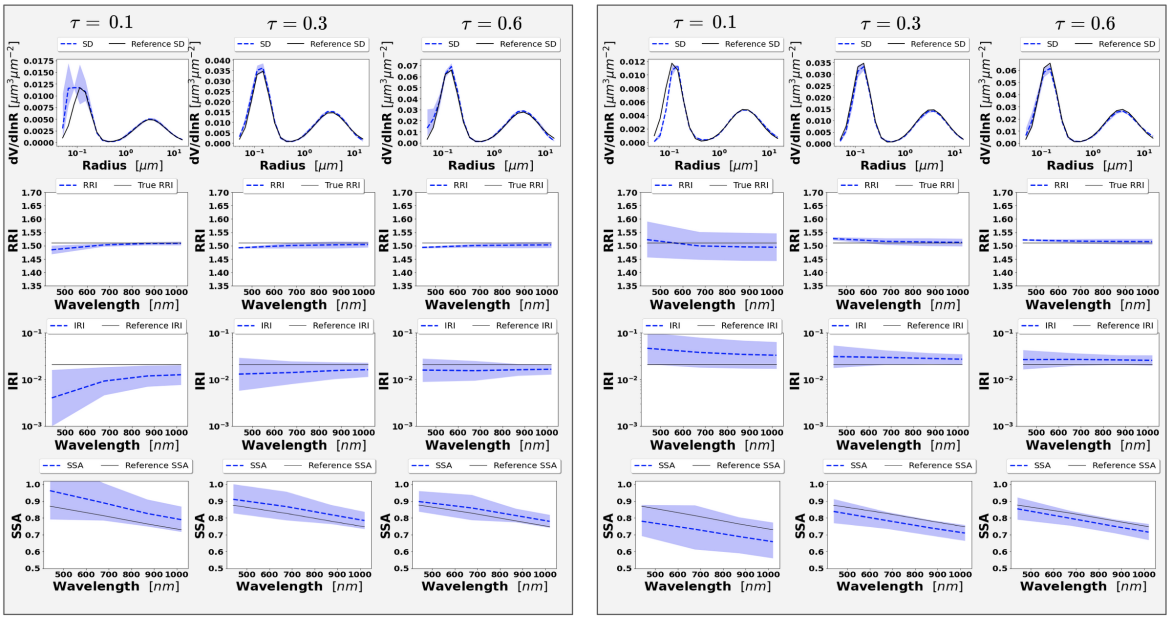


Figure 4.26: Aerosol properties retrieved from simulated sun-photometer data with assumed bias in RAD simulated data for BB aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +5% are represented in the block on the left and negative bias −5% in the block on the right. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic errors estimated by GRASP algorithm.

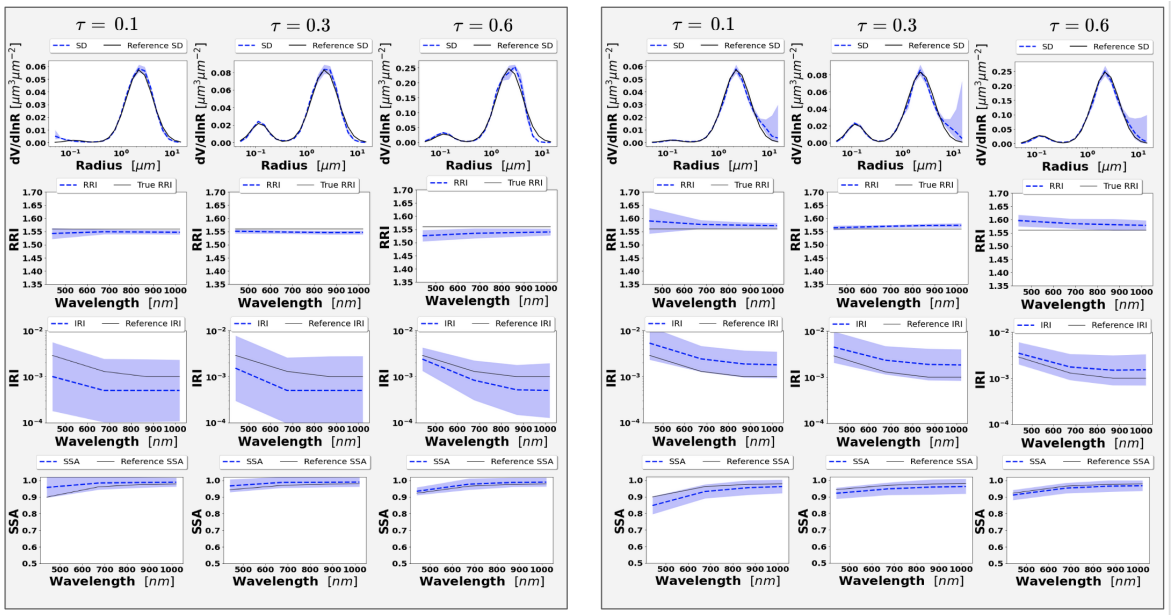


Figure 4.27: Aerosol properties retrieved from simulated sun/sky photometer data with assumed bias in RAD simulated data for Dust aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +3% are represented in the block on the left and negative bias −3% in the block on the right. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic errors estimated by GRASP algorithm.

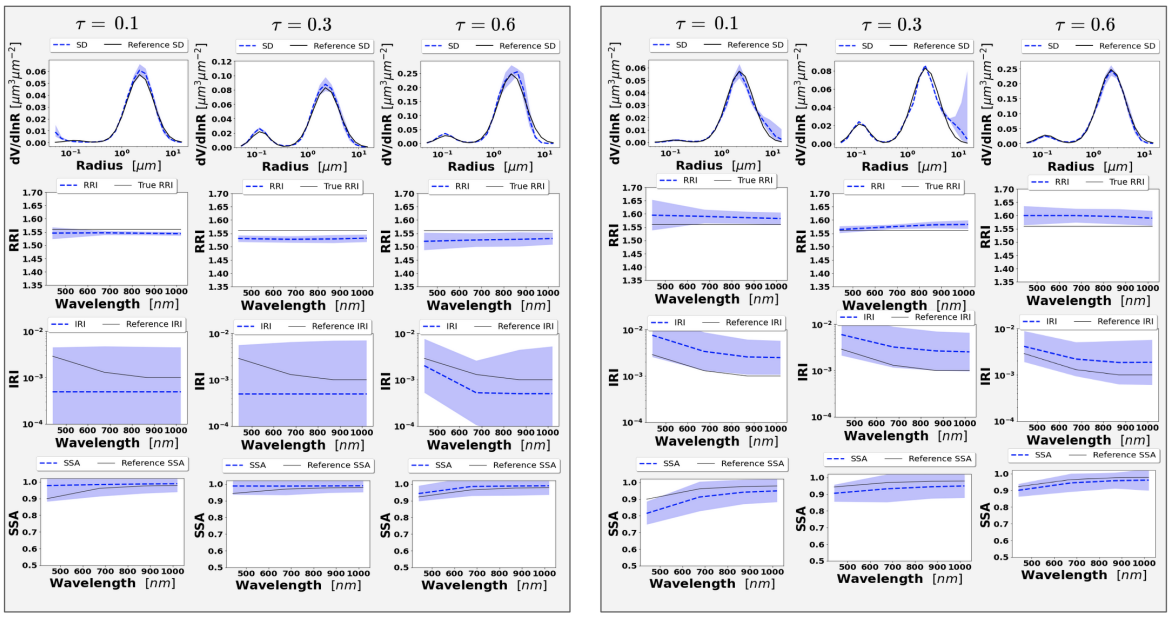


Figure 4.28: Aerosol properties retrieved from simulated sun/sky photometer data with assumed bias in RAD simulated data for Dust aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +5% are represented in the block on the left and negative bias −5% in the block on the right. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic errors estimated by GRASP algorithm.

issues and account for the biases in the error estimates. Also, the analysis here is focused on the simultaneous biases of moderate values (±0.01 in AOD and ±5% in radiances), since the appearance of simultaneous biases, of the highest bias values (i.e., ±0.02 in the AOD and ±5% in radiancies), lead to very strong effects in the retrievals. Those situations can be easily seen and screen out by quality filters (e.g., by high value of misfit). Also, it is quite unlikely to have such strong systematic errors in practical observations as those by AERONET.

As can be seen from Fig. 4.29 the biggest differences and highest bias values in the retrieval are found for low AOD (0.1). As seen earlier for this situation the errors for the RRI remain notably underestimated. As already mentioned, the situation is expected to be improved once the effects of both random and systematic errors are considered.

Figure 4.30 illustrate such situation for the retrieval of BB and dust for the three different aerosol loads (0.1, 0.3 and 0.6) when total error estimate includes both random and systematic components as:

$$\sigma_{tot} = \sqrt{\sigma_{ran}^2 + \sigma_{bias}^2} \tag{4.3}$$

where σ_{bias}^2 is calculated as was indicated in Eq. 4.2.

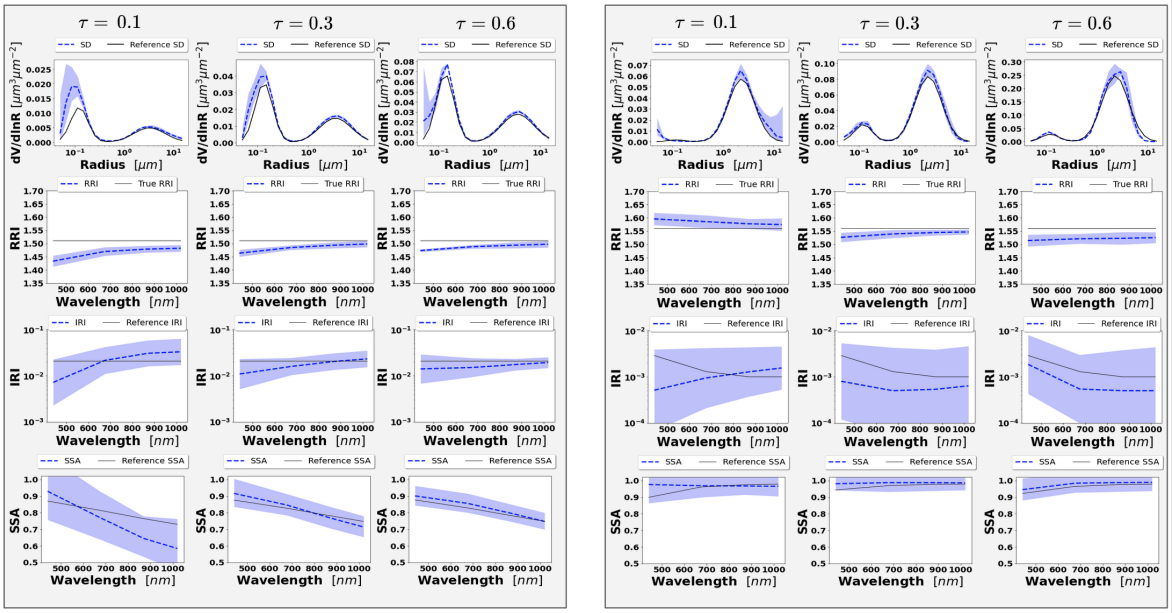


Figure 4.29: Aerosol properties retrieved from simulated sun/sky photometer data with assumed bias in AOD and radiances simulated data for BB (left) and Dust (right) aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +0.01 in AOD and +5% in radiances in both cases. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates systematic error estimated by GRASP algorithm.

It can be seen that the total error estimates capture the deviations for all parameters in the presence of random and systematic noises. These results confirm that the estimations using Eq. 3.56 based on the additional assumptions of potential presence of bias in the measurements improve the results of error estimated compare to the approach discussed in Section 4.4.1 when the effects of biased were taken into account only based on the value of the observation misfit. The observed tendencies in the effects of biases on the retrieval are consistent with all the results previously described in earlier studies. The obtained results are expected to be representative for most of practical situations, while some additional tests and analysis could certainly be useful. Therefore, in the examples presented below and for the real cases analyzed, the total error will be used as described in Eq.4.3. It means, the representation of the error will take into account the contribution of the random and systematic component. This last component contains the contribution of Levenberg-Marquardt and the misfit and the measurements, in which the contributions of ± bias added in the measurements are considered (Eq.4.2). These values of the assumed biases in our applications are consistent with AERONET as was aforementioned: ±0.01 in AOD and ±5% in radiances.

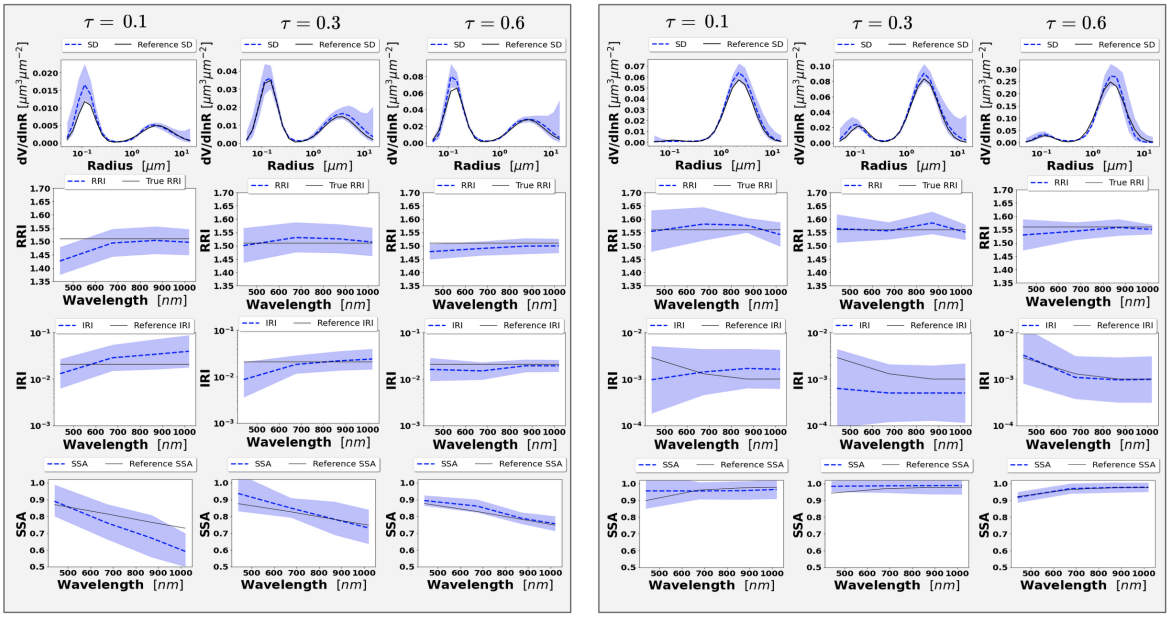


Figure 4.30: Aerosol properties retrieved from simulated sun/sky photometer data with assumed random noise and bias in AOD and radiances simulated data for BB (left) and dust (right) aerosol for $\tau(440) = 0.1$, 0.3 and 0.6 (left to right). Retrievals after adding positive bias +0.01 in AOD and +5% in radiances in both cases. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates total errors provided by GRASP algorithm.

4.4.2.2 Effects of measurement bias in retrieval of mixed aerosol properties from measurements of radiometer only

The present section tries to understand how the bias affects when inhomogeneous aerosol are observed. The example is not commonly considered in practical application, e.g. in AERONET operational processing. At the same time, since GRASP can consider this type of bi-component inversion that are fundamentally of high interest, we are analyzing this situation in the presence of biases. In the Section 4.4.2.1 we have shown different examples, considering bias in each measurement separately, and we have also illustrated the complete example with presence of both random errors and bias in all the measurements. Here we illustrate directly the results considering the presence of both random and bias in all the measurements since this is complete situation that is most close to the most practical situation.

Different tests were performed for this study. In particular, we focus on the case of BB-Dust since the Section 4.4.2.1 has already demonstrated the bias affects when each type of aerosols is observed separately. The effects of each bias separately were analyzed while the corresponding illustrations are not shown since the results presented similar tendencies to those previously discussed are observed for observation of each type of aerosol separately in the last section. Figure 4.31 illustrates the examples of BB-Dust

when $\tau(440) = 1.0$, for different aerosol loads ($\tau_f = 0.8$ and $\tau_c = 0.2$, $\tau_f = \tau_c = 0.5$ and $\tau_f = 0.2$ and $\tau_c = 0.8$) assuming bias and random noises. The shaded areas represent the estimated total errors, as shown in Equation 4.3. An important observation is that the error estimates for all retrieved and derived parameters well characterize the actual errors. As it can be gained from the figure, the retrieval of the properties of minor component appears as the most challenging. As a matter of fact, the biggest errors in the retrieval are observed for the fine mode properties, particularly, in the case of $\tau_f = 0.2$ and $\tau_c = 0.8$. The largest discrepancies of estimated errors with the actual ones are observed in this situation. On the other hand, the properties of coarse mode are well represented in almost all cases showing a good accuracy compared to the properties of the fine mode even in the most challenging cases with the smallest presence of coarse mode. This can probably be explained by the fact that desert dust AOD has rather moderate spectral changes.

On the other hand, in this section are also provided some illustrations for the lidar ratio in order to demonstrate how the retrievals and the error estimates are affected by the bias in the measurements. Figure 4.31 illustrates the lidar ratios in this situation of mixed aerosol. The retrieval results and estimation of LR errors are rather satisfactory, with exceptions of low AOD cases, mainly in the case where the fine mode has only very minor presence (of $\tau_f = 0.2$ and $\tau_c = 0.8$). Also, it should be emphasize that the error estimated for total SSA and RI are rather adequate while in the Section 4.4.1, where only random errors were considered, the results showed some apparent underestimations.

4.4.2.3 Effects of measurement bias in retrieval of mixed aerosol properties from observations of radiometer in combination with lidar

This section considers the same example as in Section 4.4.2.2 and analyses the effects of measurement biases in the synergy retrieval using co-incident measurements sun/sky photometer and lidar measurements. At the same time, the results are presented for most practical situation when both random and biased are present in measurements and accounted in the error estimates.

Figure 4.32 shows the results for the example of BB-Dust described in the previous section for $\tau(440) = 1.0$ and assuming the presence of both bias and random noises in all the synthetic measurements for AOD, radiometer and lidar. As can be seen, the results of these different tests illustrate the positive influence of using radiometer and lidar synergy. The error estimates seem to be rather accurate too. For example, the most notable enhancement is in the lidar ratio accuracy especially when the mode fine is the smallest, i.e. for the case with: $\tau_f = 0.2$ and $\tau_c = 0.8$. This behavior was also seen by Lopatin et al. (2013) who explained that these were expected results since lidar

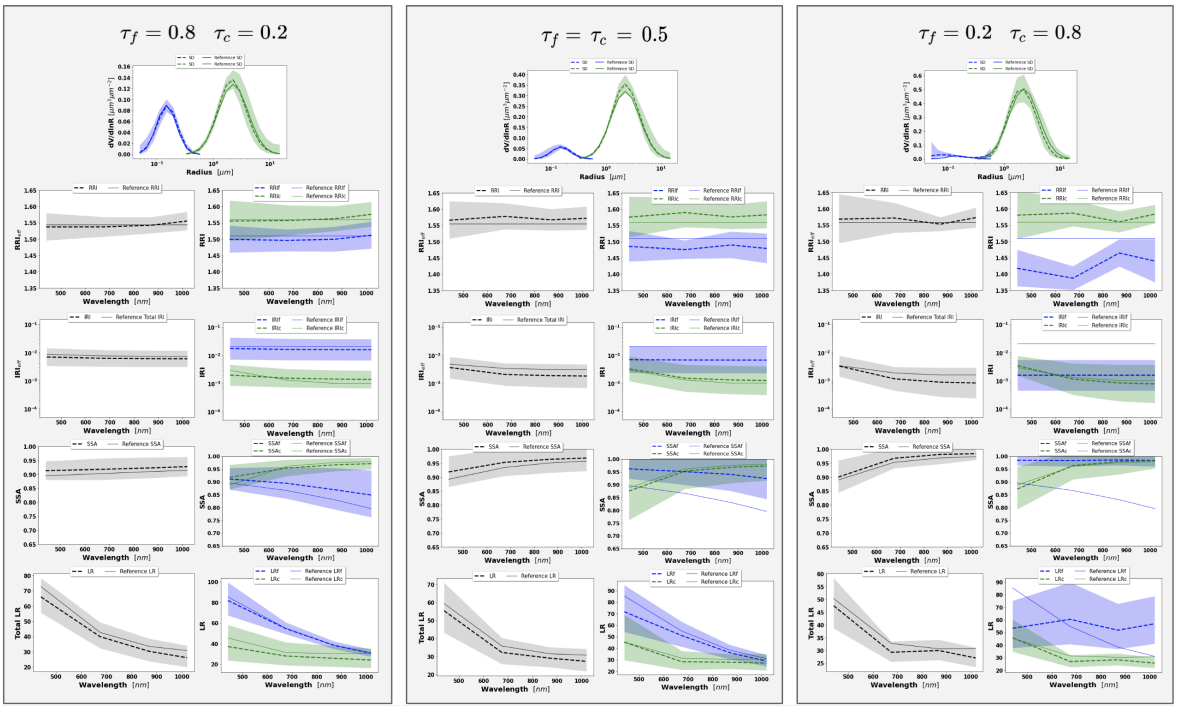


Figure 4.31: Aerosol properties retrieved from simulated sun/sky photometer data with assumed random noise and bias in AOD and radiances simulated data for BB-Dust for $\tau(440) = 1.0$. Retrievals after adding positive bias +0.01 in AOD and +5% in RAD in both cases. The solid lines are the simulated properties (SD, RRI, IRI, SSA and LR), the dashed lines are the retrieved parameters. The shaded area indicates total errors provided by GRASP algorithm.

ratio has a high sensitivity to lidar signal. Nevertheless, we can see some improvements in retrieval of complex refractive index using both lidar and photometer data.

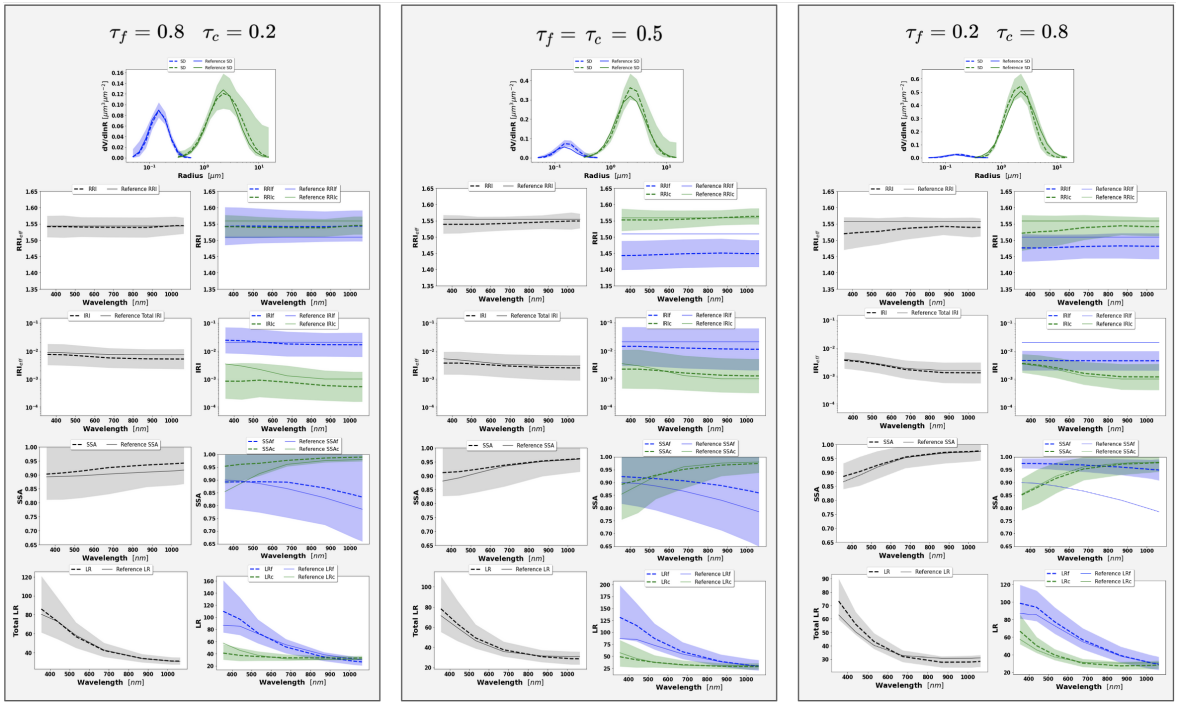


Figure 4.32: Aerosol properties retrieved from simulated sun/sky photometer and lidar data with assumed random noise and bias in AOD, radiances and lidar simulated data for BB-Dust for $\tau(440) = 1.0$. Retrievals after adding positive bias +0.01 in AOD, +5% in RAD and +0.2, +0.15 and 0.1 for each lidar wavelength: 355 nm, 532 nm and 1064 nm respectively. The solid lines are the simulated properties (SD, RRI, IRI, SSA and LR), the dashed lines are the retrieved parameters. The shaded area indicates total errors estimated by GRASP algorithm.

In regard to the accuracy of the error estimation, in the Section 4.4.1, we have illustrated the retrieval of error estimates for LR and showed some apparent underestimation when only random errors were considered. Figure 4.32 illustrates an important improvement in the estimation of the errors, once both random noise and bias considered and Eq. 3.56 were used for accounting the effect of the systematic component.

Thus, using the synergy of both instruments can provide more accurate retrievals of LR and the error can be estimated rather accurately using the developed methodology for both aerosol components even for aerosol mode with the lower presence. Figure 4.33 illustrates the retrievals of the vertical aerosol profile in all three cases. The results show similar tendencies as in Section 4.4.1.

4.4.3 Illustration and description of the correlation matrices

As introduced in section 3.3.3, the values of non-diagonal elements of covariance provide important and interesting information about the retrieved parameters. For example,

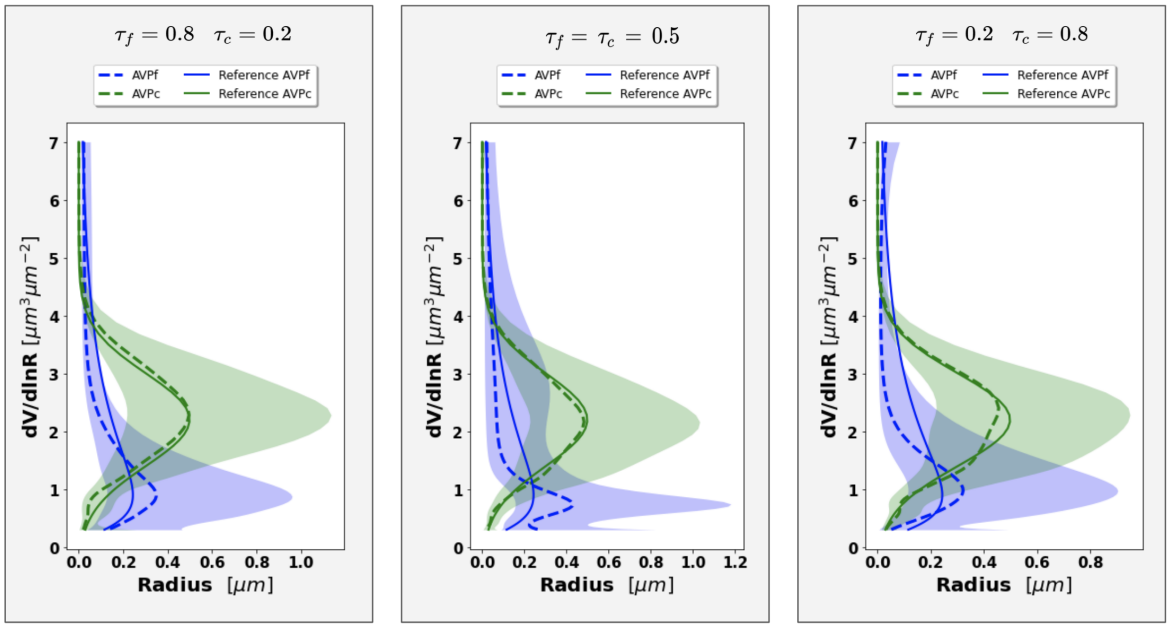


Figure 4.33: Aerosol vertical profiles (AVP) from simulated sun/sky photometer and lidar data with assumed random noise and bias in AOD, radiances and lidar simulated data for BB-Dust for $\tau(440) = 1.0$. Retrievals after adding positive bias +0.01 in AOD, +5% in RAD and +0.2, +0.15 and 0.1 for each lidar wavelength: 355 nm, 532 nm and 1064 nm respectively. The solid lines are the simulated properties (SD, RRI, IRI, SSA and LR), the dashed lines are the retrieved parameters. The shaded area indicates total errors estimated by GRASP algorithm.

if the values $\rho_{ii'} \neq 0$ are close to ± 1 the similitude of the influences of the parameters a_i and $a_{i'}$ on the inverted measurements $\mathbf{f}^*$ may explain the large variances of the retrieval error for these parameters. Also, knowledge about $\rho_{ii'} \neq 0$ is highly useful for the situation when several parameters from a set simultaneously retrieved parameters a_i need be jointly used in the applications. This can be easily seen from Eq.3.54. For example, let us consider the estimates of two parameters a_1 and a_2 which have errors Δa_1 and Δa_2 characterized by covariance matrix:

$$\mathbf{C}_\Delta = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 \end{pmatrix} \tag{4.4}$$

where $\mathbf{a}$ is a vector defined as $\mathbf{a} = (a_1, a_2)^T$. Correspondingly if in applications one needs to use the characteristics m that is a liner function of $m = K_1 a_1 + K_2 a_2$, the variance σ_m^2 can be obtained from Eq. 3.54 as:

$$\sigma_m^2 = \mathbf{K}\mathbf{C}_{\Delta \mathrm{a}}\mathbf{K}^T = \begin{pmatrix} K_1 & K_2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 \end{pmatrix} \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} \\ = K_1^2\sigma_1^2 + K_2^2\sigma_2^2 + 2K_1K_2\sigma_1\sigma_2\rho_{12} \tag{4.5}$$

From this equation, the importance of the correlation coefficient ρ_{12} is quite evident. Specifically, if $\rho_{12} = 0$ then the variance σ_m^2 is just a simple sum $K_1^2\sigma_1^2 + K_2^2\sigma_2^2$. Therefore, the error propagation from $\Delta \mathrm{a}_i$ to Δm is straightforward, and namely only the values of sensitivities K_i^2 determine the contribution of $\Delta \mathrm{a}_i$ (decreasing or increasing) to Δm.

When $\rho_{12} = 0$ the situation is more complex. However, the estimation of main tendencies can be simplified in some cases. For instance:

$$\text{if} \quad \sigma_1^2 = \sigma_2^2 \quad \text{and} \quad K_1 = K_2, \quad \text{then} \quad \sigma_m^2 = 2K_1^2\sigma_1^2(1 + \rho_{12}) \tag{4.6}$$

or

$$\text{if} \quad \sigma_1^2 = \sigma_2^2 \quad \text{and} \quad K_1 = -K_2, \quad \text{then} \quad \sigma_m^2 = 2K_1^2\sigma_1^2(1 - \rho_{12}) \tag{4.7}$$

From these equations, it can be seen that if correlation coefficient $\rho_{12} \to 1$ or $\rho_{12} \to (-1)$ then depending on the case in Eqs. 4.6 and 4.7 σ_m^2 can be close to zero or up $4K_1^2\sigma_1^2$. Therefore, the knowledge about the non-zero correlation coefficients $\rho_{ii'} \neq 0$ is very important to understand how the error is propagated to derived parameters obtained from the primary set **a**.

In practical cases, when the derived parameter m is a function of a large number parameters a_i, the contributions to σ_m^2 become increasingly very complex with the increase of the number of involved parameter a_i. Therefore, unfortunately, the general qualitative analysis, similarly to the one demonstrated by Eqs. 4.5-4.7, becomes very difficult and often practically impossible. Nonetheless, as it will be shown below the visualization of the correlation matrices Eq. 3.59 can be very useful for analysis of the retrieval tendencies.

4.4.3.1 Retrieval of single aerosol component from radiometer measurements

Figure 4.34 shows the correlation matrices of random retrieval errors for BB (spherical particles) and dust (non-spherical particles) for conventional AERONET-like inversion. The first 22 parameters (22x22) represent the SD. It is followed by two blocks of 4x4. These two blocks are related to the RRI and IRI for 4 wavelengths ($440\,nm$ to

1020 nm). The last parameter is the sphericity fraction (1 x 1). The colors represent the values of correlation coefficients, where the red color denotes positive correlations, blue color indicates negative correlations. The density of the colors indicates values of the correlation coefficients changing from zero (the white color) to dense red or blue colors corresponding to values 1 and -1 accordingly.

The correlation for biomass burning case is shown in Fig. 4.34a. As it can be seen, in general size bins retrievals have rather moderate correlation between them, though large positive correlations between the retrieval errors of neighbours can be observed. This is more evident for size bins at the smallest and largest particle sizes. This indicates that size distribution values for those size have tendency to be overestimated or underestimated together, which can be mostly explained by the use of typical smoothness constraints imposed to the size distributions as explained in section 3.3.3. The errors of RRI and IRI are negatively correlated with the SD parameters. The correlations seem especially pronounced between RRI and values of size distribution for the fine mode. Correspondingly the overestimations of size distribution values may tend to be accompanied by underestimation of RRI and vice versa. The errors of the fraction of spherical particle seem to show positive correlation with the SD retrieval errors. This correlation is more evident when there is a fine mode dominated aerosol, since scattering of fine mode particles has a quite similar shape for sphere and spheroids. Therefore when there is fine mode domination is more difficult to differentiate between sphere or spheroids. The positive sign of correlataion can be explained by the fact that extinction cross sections for the equivalent radii are a bit higher for spheroids. Thus, a higher percentage of spheres can be optically compensated by an increase in the volume concentration, without a big impact in the total residual. The fraction of spherical particle shows a negative strong correlation with the errors of the refractive index. Strong positive correlations can be seen between spectral values of RRI. The positive correlation are presents but lower between spectral values of IRI. As notice already this likely relates with the use of rather strong smoothness constraint of spectral variability of RRI and weaker constraint of spectral variability of IRI in the retrieval (Dubovik and King (2000)). The essential positive correlation can be noticed also between errors of RRI and IRI.

Figure 4.34b shows the correlation matrix for retrieval of dust aerosol. The structure of the correlation for SD exhibits some differences compare to the BB case. Specifically, the positive correlations between neighboring size bins for the smallest and largest particle sizes are even more pronounced. Also, somewhat stronger negative correlations can be seen the intermediate sized. The strong negative correlation between RRI and SD retrieval errors remains only between concentrations of very small particles and values of RRI at shortest wavelengths. The notable positive correlation presents only between spectral values of RRI at the shortest wavelengths and between spectral values

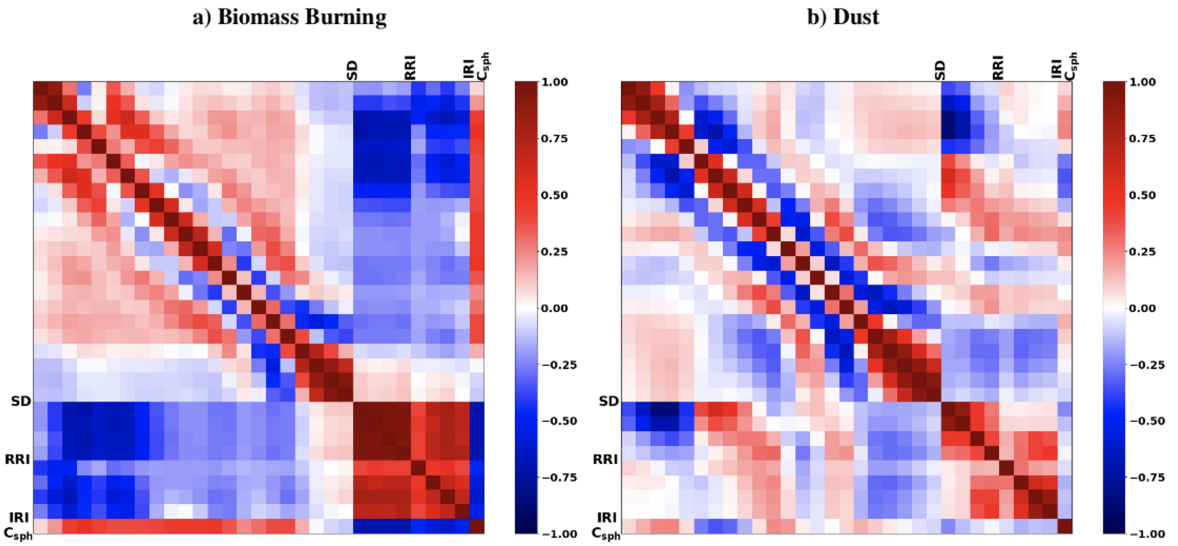


Figure 4.34: Correlation matrices of the estimated errors for aerosol retrieval from Sun/sky-radiometer observations a) for biomass burning aerosols, and b) for desert dust using GRASP algorithm. The values close to 1 or −1 mean stronger correlations between the properties, positive or negative, respectively.

of RRI at the longest wavelengths. At the same time, overall the correlation of retrieval errors of both RRI and IRI between themselves and with other parameters decreases compared to the case of BB. The errors of the fraction of spherical particle for dust case correlate much less with the errors of the other parameter compared to BB case. This can be explained by the fact the light scattering of large particles is significantly more sensitive to deviation of aerosol particles from spheres compared to spheroids, than the light scattering of fine fraction particles (Dubovik et al. (2006)). Therefore, when coarse particles dominate the discrimination between spheres and spheroids becomes more evident.

Thus, the analysis of the correlation matrices itself provide very useful inside that helps to understand and interpret retrieval results. For example, such artifacts as appearance of 'tails' (unrealistically high concentrations) at extremes of size distribution has been noticed and widely discussed (Dubovik et al. (2002b, 2006); Torres et al. (2014), etc.). Such retrieval artifacts as underestimation RRI accompanied by overestimation size distribution of very fine particles has been widely discussed in studies by Dubovik et al. (2000, 2002a,b). These artifacts were strongly reduced by accounting for particle non-sphericity of desert dust particles (Dubovik et al. (2002a, 2006)), but nonetheless the less pronounced appearance of such artifacts remains in AERONET like retrieval (Torres et al. (2014, 2017)). These artifacts are clearly related with the observed above presence of strong negative correlation between values of RRI and size distribution of very fine particles. It should be noted that the presence of high correlations is an indication that adding information about one of the correlated pa-

rameters should improve retrieval not only the constrained parameter itself but also the parameters that strongly correlated with this parameter. For example, addition of polarimetric observations to the traditional set of AERONET observations results in clear improvement in the retrieval of RRI and size distribution of very fine particles (Li et al. (2009b); Fedarenka et al. (2016)). Indeed, the degree of linear polarization is known to be very sensitive to amount and especially the RRI of fine particles (Dubovik et al. (2006)). This is why, the addition of polarimetric observations help to reduce the correlations between the errors of RRI and size distribution of fine particles that helps to the overall improvement of the retrieval accuracy of these particles.

4.4.3.2 Retrieval of mixed aerosol component from measurements of radiometer only

In this section, the correlation matrix for bi-component aerosol retrieved from synthetic observations of sun/sky radiometer of two aerosol mixtures is illustrated : BB-dust and Urban-dust. The structure of this matrix consists in 25 parameters related to the SD that are separated into two blocks: 10 parameters for fine mode and 15 parameters for coarse mode. The following 4 blocks of 4x4 are related to the RRI and IRI of fine and coarse modes at 4 wavelengths. These blocks are followed by a single value of the sphericity fraction.

The area of the correlation matrix that contains SD, RRI, IRI and sphericity fraction is quite similar to the correlation matrix obtained for aerosol AERONET like retrieval. The main difference is the separation into two modes since strong negative correlations can be observed between the corresponding parameters of fine and coarse mode. For example, strong negative correlations can be observed between IRI fine and coarse. These correlations mean that overestimating the amount or absorption of one aerosol mode is likely compensated by underestimation of the amount or absorption of another aerosol mode. Another interesting anticorrelation can be observed for the last three bins of SD fine mode and the first three bins of SD coarse mode. Actually, both volume distributions have these three bins in common. This overlap zone is never easy to properly separate for the code but at the same time, it coincides to a local minimum value of most size distributions found in the real retrievals.

4.4.3.3 Retrieval of mixed aerosol properties from measurements of radiometer in combination with lidar

Figure 4.36 shows the correlation matrix for the case when bi-component aerosol retrieved from synthetic observations sun/sky photometer and lidar of aerosol mixtures. In the figure the different blocks are identified. The first 25 parameters represent the SD that are separated into two blocks: 10 parameters for fine mode and 15 parameters

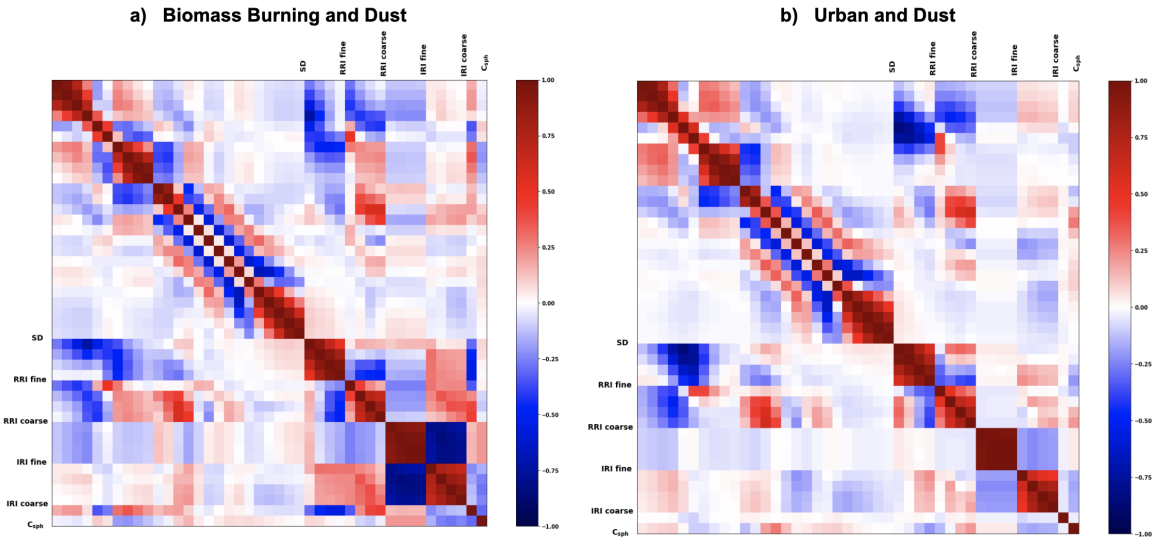


Figure 4.35: Correlation matrices of the estimated errors for mixed aerosol retrieval from Sun/sky-radiometer observations a) for the mixture of biomass burning and dust aerosols, and b) for the mixture of urban and dust aerosols using GRASP algorithm. The values close to 1 or −1 mean stronger correlations between the properties, positive or negative, respectively.

for coarse mode. The following 4 blocks of 7x7 are related to the RRI and IRI of fine and coarse modes at 7 wavelengths. These blocks are followed by a single value of the sphericity fraction. Two last and largest blocks of 60x60 parameters each correspond to the AVP values of two modes given at 60 different altitudes.

The area of correlation matrix that contains SD, RRI, IRI and sphericity fraction is quite similar to the correlation matrix described in previous section considered the aerosol mixture retrieval from AERONET like observation only.

As it could be expected, the block of the correlations of AVP retrieval shows strong negative correlations between errors of the retrieved parameters of fine and coarse mode. Thus, an overestimation of one mode is higly correlated with an underestimation of another AVP mode. Furthermore, an strong positive correlations can be observed between AVP values corresponding to the same fine or coarse mode; i.e. the AVP values of each mode have tendency to be simultaneously overestimated or underestimated. This is related to the limited sensitivity of used lidar data for distinguishing the contributions of different modes, and also, to the use of smoothness constraints on vertical variations of AVP of each fraction. On the other hand, nearly zero correlation can be seen for AVP parameters at the altitudes with significant presence of one or both aerosol components. Correspondingly, there is a high sensitivity of both lidar and radiometer observation to the aerosol parameters at those altitudes. It should be noted that all above discussed retrieval suggested from the analysis of correlation matrices where actually observed in the retrievals from real data as discussion by Lopatin et al.

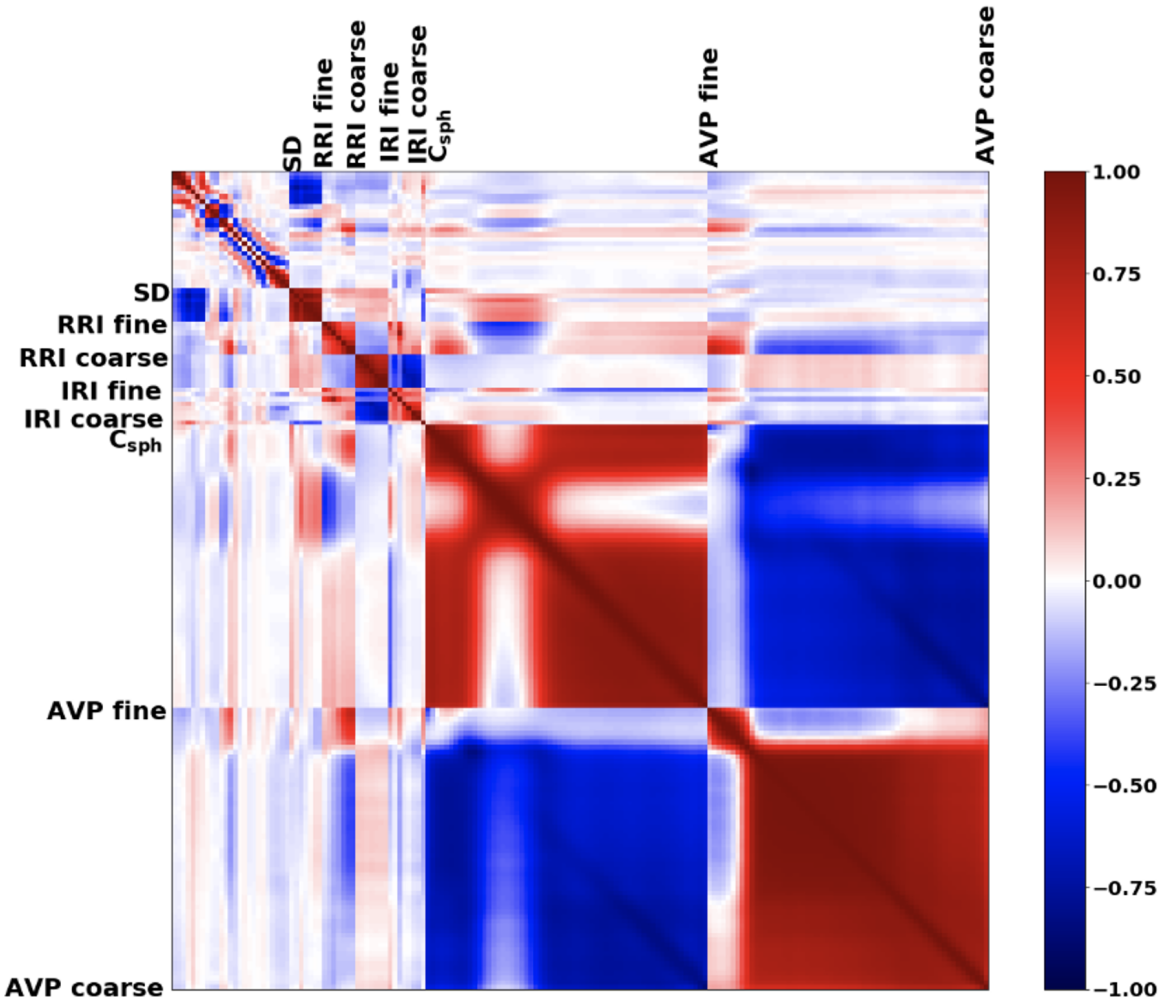


Figure 4.36: Correlation matrices of the estimated errors for aerosol retrieval from joint Sun/sky-radiometer and lidar observations for a mixture of urban and desert dust using GRASP algorithm. The values close to 1 or −1 mean stronger correlations between the properties, positive or negative, respectively.

(2013, 2021).

4.5 Conclusions

In this chapter has been evaluated the performance of the GRASP error estimates for aerosol parameters retrieved from ground-based observations. In order to achieve the results, AERONET like retrievals form observations by Sun/sky-scanning radiometer and GRASP synergy aerosol retrieval from joint observations by radiometer and lidar have been considered. Moreover, the capability of GRASP to generates the full covariance matrices has shown, which are used to provide the error bars for retrieved parameters and also an interesting inside for understanding retrieval tendencies. Therefore, the quantitative reliability of the obtained covariance diagonal elements and analyzed

the structure of correlation coefficient of covariance matrices have been studied.

The performance of GRASP estimates of error variances in presence of random errors was evaluated in series of numerical tests and illustrated the capabilities of GRASP algorithm to provide rigorous estimates of dynamic retrieval errors. In frame of these tests the synthetic proxy observations perturbed by 300 random noise generated realization were inverted using GRASP algorithm. Then, retrieved parameters were compared the those used for generation of the synthetic data and the obtained error estimates were compared with actual deviations of the retrieved parameters from assumed values. This analysis was realized for synthetic observations for three different types of aerosols, as well as, for the mixture of them. Observations of desert Dust were modeled using AERONET retrieval climatology at Solar Village (Saudi Arabia) site. The AERONET retrieval climatologies from African savanna (Zambia) and the GSFC (Maryland, USA) were used simulating Urban and BB aerosol observations respectively. The Urban-Dust aerosols and BB-Dust mixture were considered for modeling properties of mixed aerosols. For each observed aerosol type or mixture of them different aerosol loads were tested. First, the observation of aerosols of only one type aerosol for $\tau(440) = 0.3$, 0.6 and 0.9 was modelled. For aerosol mixtures was also considered scenarios with different aerosol loads, while are presented most of the illustrations for the situation with the large aerosol load ($\tau(440) = 1.0$, combining different aerosol loads for each mode $\tau_f = 0.2$ and $\tau_c = 0.8$, $\tau_f = \tau_c = 0.5$ and $\tau_f = 0.8$ and $\tau_c = 0.2$). The data were perturbed by random noise before applying the retrieval algorithm. For all these simulations it was used the SZA at 75 degrees.

The tests evaluated the situations when only radiometer data were inverted and then radiometer data were inverted jointly with co-incident lidar data. Two GRASP retrieval set ups were tested: (i) then the retrieval assumes that aerosol is composed by homogeneous particles and parameters of only one aerosol component are retrieved and (ii) then the aerosol is assumed as external mixture of two aerosol components and parameters of each component are retrieved separately. In case the when lidar data were used the vertical profiles of concentration were also retrieved for each aerosol component. The illustrations for aerosol retrieval from all simulated data sets were provided using both approaches. However, the full statistical analysis was provided only on the two conventional retrieval scenarios: - AERONET like single component retrieval from radiometric observations and - GRASP bi-component aerosol retrieval from combinations of radiometer and lidar data.

The results of the tests showed that the complete set of aerosol parameters for each aerosol component can be robustly derived with acceptable accuracy in almost all considered situations. The retrieval of bi-component aerosol was evaluated using radiometer only simulated measurements and then by adding lidar observations. These tests allowed us to observe that by using the synergy of two instruments, some im-

provements were achieved in the retrieval of aerosol properties of each component of observed aerosol mixture and in the estimations of the retrieval errors. The test for selected cases with different presence of different aerosol components ($\tau_f = 0.2$, $\tau_c = 0.8$ and $\tau_f = 0.8$, $\tau_c = 0.2$) showed that optical properties of the dominant mode can be retrieved significantly more accurately as can be expected. It is interesting to note that in all situation using only a radiometer data or adding lidar simulated measurements, such properties as total SSA and effective refractive index can be retrieved rather accurately, even in cases where the retrieval of properties of each mode separately is questionable.

The results of the statistical tests with randomly generated noise showed that GRASP error estimates in most cases are comparable or exceed the actual errors by the 20 to 30% and therefore can be safely used for assuring uncertainties of actual retrieval products. In addition, the observation of typical error values was summarized for different situations and retrieval scenarios. Namely, the study confirmed that the detailed properties of aerosol mixtures can be rather reliably retrieved from a combination of radiometer and lidar data provided that there is sufficient amount of both aerosol components. For example, for the case when total $\tau(440) = 1.0$ with comparable presence of both components $\tau_f(440) = \tau_c(440) = 0.5$, and SZA is 75 degrees, the mean values for RRI errors are ~ 0.05 for BB and Urban and vary between 0.073 to 0.083 for Dust, IRI errors are around 58% for BB, 72% for Urban and 45% for Dust. SSA errors vary between 0.024 to 0.061 for Urban, 0.041 to 0.089 for BB and 0.04 to 0.009 for Dust, showing a clear tendency to increase with the wavelengths for values of fine mode and decrease for coarse mode. However, even for this case the separation of the LR values for both modes showed high uncertainties at short wavelengths in particular for the fine mode, while the values of the total LR errors were found reasonable in the range of 20% to 55%. The relative error estimates for AVP for both aerosol mixtures (Urban-Dust and BB-Dust) cases where ranging for fine mode vary between $50 - 70\%$ and for coarse Dust mode between $50 - 55\%$.

The effects of the systematic errors on the retrievals were also analyzed in a series of limited dedicated numerical tests. As was mentioned, the apparent misfit was used as an indicator of bias in the GRASP algorithm. However, in the real situations not all biases can be seen from the misfit. Thus, this chapter has been presented the results considering a possible solution to this problem. Therefore, the contribution of potential bias was included in the estimation of the retrieval errors. In this regard , an extra term in the equation of the systematic errors was added which accounts for propagation of possible bias from the measurements.

The estimated systematic error over all agrees well with actual manifestations of the bias in the retrieval. The estimations reflect very well all qualitative effect of biases on the retrieval. The results confirm improvements in the error estimates adding the new

term to the equation of the systematic errors. It means, the additional assumptions of potential presence of bias in the measurements help to model the contribution of bias in the retrieval uncertainties.

The observed tendencies in the effects of biases on the retrieval are consistent with all the results already described in other studies. The obtained results are expected to be representative for most of practical situation, while some additional tests and analysis can be certainly useful.

In addition, to the evaluation of error bars estimates and effects of systematic errors, in this paper we illustrated and discussed the correlation structures of the error covariance matrices for all main considered retrieval scenarios. The results showed that analysis of the correlation structure can be very useful for understanding the observed retrieval tendencies and optimizing retrieval. For example, for conventional AERONET like aerosol retrievals from radiometer data only, the strong negative correlation between errors of real part of the refractive index and size distribution values for small sizes. This suggests agrees well with the tendency commonly observed in actual retrieval when the underestimations of the real part are coincident with overestimation of fine mode size distribution. Also, the presence of high positive correlation between the errors of size distribution for extreme sizes and between the errors of refractive index at different wavelengths agrees well with known possibilities of possible overestimations of aerosol concentrations for very small or very large particles and joint overestimations/underestimations of refractive index value at different wavelengths. For bi-component retrievals strong negative correlations can be observed between nearly all corresponding parameter of fine and coarse mode. This means, for example, that the overestimation of amount or absorption of one aerosol mode is likely compensated by underestimation of amount or absorption of another aerosol mode. The decrease of some of these correlations was observed when inverted radiometer data were inverted simultaneously with the lidar data. The high positive correlations were seen for the errors of the vertical profile of the fine and coarse concentrations with the exception of the values for the altitudes where one or both of aerosol modes had substantial loads. These and other less obvious, but quite interesting, correlations structures and tendencies can be identified using the analysis of the correlation matrix structure. Thus, the availability and analysis of not only error variances but also correlation patterns appear to be a useful and promising approach for optimizing observation schemes and retrieval setups.

Chapter 5

GRASP error estimates using real observations

The previous chapter allowed to validate the estimation of errors from different tests using synthetic ground-based observations. Based on the previous tests, it was possible to see how the errors are related to several aspects, for example, one of the simplest to observe is the aerosol load. These validations allowed to know the error estimates for the different scenarios, their tendencies, which configurations present lower errors, etc. However, an important part is to provide the practical outcome of the error estimates approach in real applications.

Different works were published over the last years using the GRASP algorithm in applications of real data from several remote sensing instruments. Particularly, in ground-based, GRASP has been successfully applied for different configurations. For example, GRASP has been used to retrieve columnar properties from only spectral aerosol depth measurements, as is shown by Torres et al. (2017); Torres and Fuertes (2021). Furthermore, the well known application combining the aerosol optical depth and sky radiances can be seen in different works as Torres et al. (2014), Lopatin et al. (2013); Fedarenka et al. (2016). Also, the combination of aerosol optical depth and sky/cameras was explored and the results were presented by Román et al. (2017). Another successful GRASP application was shown by Espinosa et al. (2017) from nelphelometer measurements.

In addition, the synergy between instruments has provided the possibility to retrieved extra information. Thus, the aerosol retrieval synergies from diverse combinations of ground-based passive Sun/sky photometer measurements with collocated active lidar ground-based was presented by Lopatin et al. (2013). This synergy was applied by different authors (Benavent-Oltra et al. (2017, 2019, 2021); Hu et al. (2019); Tsekeri et al. (2017)) showing the capabilities of GRASP to provide information also vertically distributed. Moreover, some studies have showed the synergy using also

ceilometer measurements (Herreras et al. (2019); Román et al. (2018), Cayuela et al. (2021)). The recent work published by Lopatin et al. (2021) provide the description of the evolution of GARRLiC/GRASP approach showing a wide spectrum of possibilities for realizing the processing of ground-based observations.

These works show the successful performance of the GRASP algorithm against different configurations and measurements. In addition, some of these studies have explored sensitivity analyzes such as Lopatin et al. (2013), Torres et al. (2014) and Román et al. (2018), which allows to have a reference on the accuracy of some of the retrieved parameters. Unlike the aforementioned works, in this section GRASP has been applied to the synergy of real lidar and sun/sky photometer observations with the objective to verify the performance of error estimates in these real cases. Thus, in this chapter is provided and illustrated the error of the different retrieved parameters. The three selected cases were chosen over Buenos Aires (Argentina) for three different aerosol loads: i) aod(440) = 0.1, ii) aod(440) = 0.3 and iii) aod(440) = 0.5. The error was obtained as explained in the Equation 4.3 in the Chapter 4. Thus, the systematic component of the error estimates is calculated as in the Equation 4.2 where the same described values of bias in the contribution of the measurements are considered.

5.1 Illustrations for ground-based observations

This section illustrates the GRASP error estimates performance for the retrieval from real data. With that purpose the lidar and sun/sky photometer measurements collected at Aeroparque ($34^\circ\ 33'\ 51''S$, $58^\circ\ 25'\ 02''W$) and Villa Martelli ($34^\circ\,33'21''S$, $58^\circ\,30'\,23''W$) stations, in Buenos Aires, Argentina have been used. These instruments are part of LALINET (Latin America Lidar Network, Guerrero-Rascado et al. (2016)) and AERONET networks. Both sites are located in an industrialized city dominated by continental and urban/industrial aerosols, and during winter and spring are affected by biomass burning from north and center of the country and neighboring countries mainly Brazil. Aeroparque station is located at the airport Jorge Newbery within the limits of the city. This station does not have a co-located sun-photometer but its location is 7 km from the Villa Martelli station where the sun-photometer is installed. On the other hand, Villa Martelli station is found in the limit of Buenos Aires city, in a highly populated and industrialized area.

The observation from two different biomass burning events in Argentina were selected for the illustrations. Specifically, three days were chosen with different aerosol loads and SZA. The lidar range-corrected signal (RCS) corresponding to each event are shown in Figure 5.1. They have been calculated from the lidar signal, with background and dark current correction, multiplied by the height squared. In addition, the

back-trajectories calculated from the HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory; Stein et al. (2015); Rolph et al. (2017)) models are presented in order to confirm where the air masses come from (Fig. 5.2).

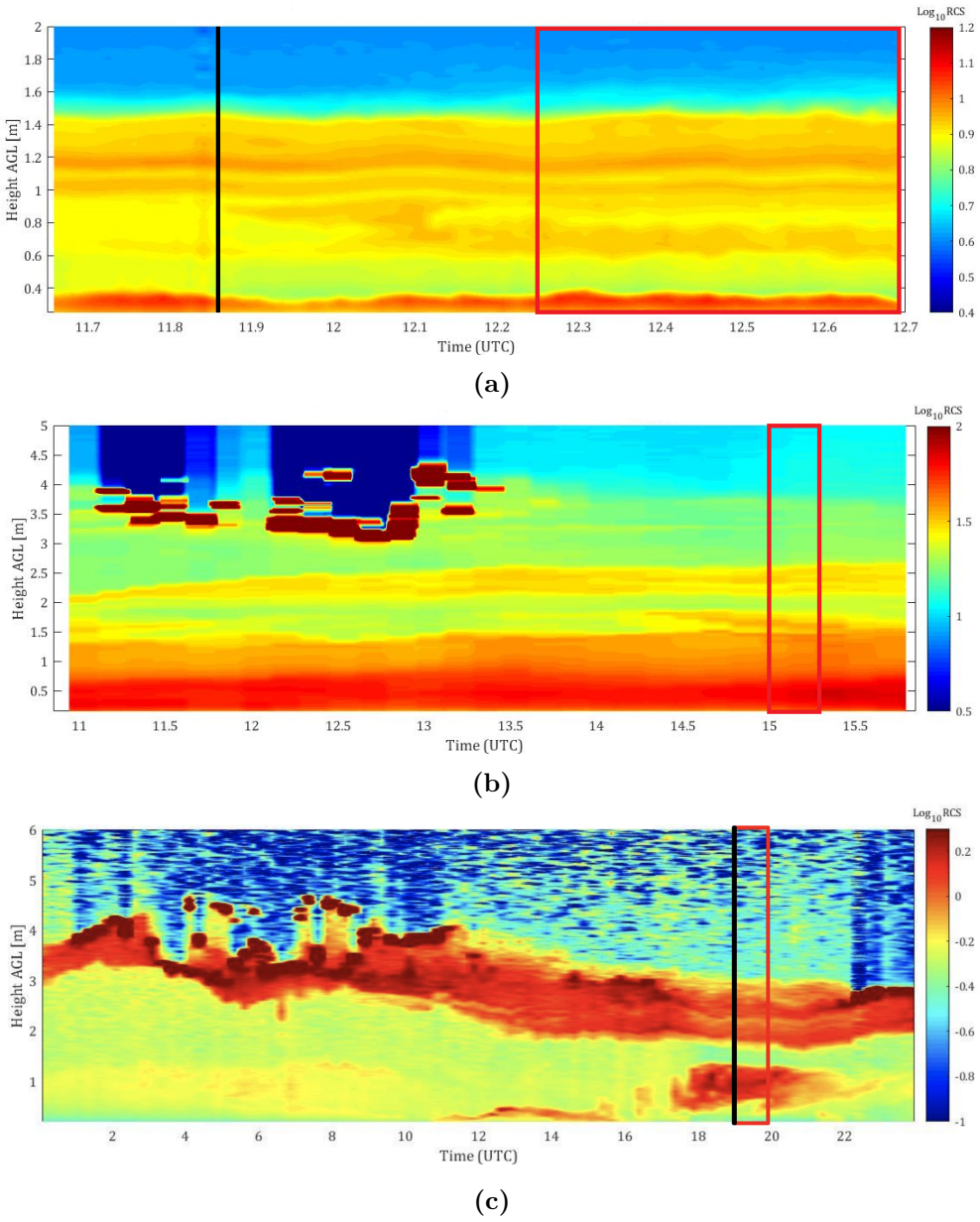
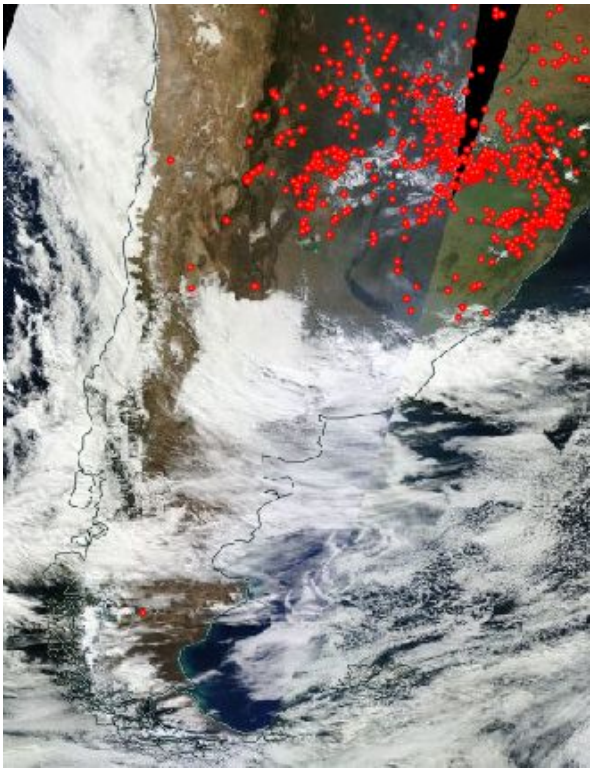
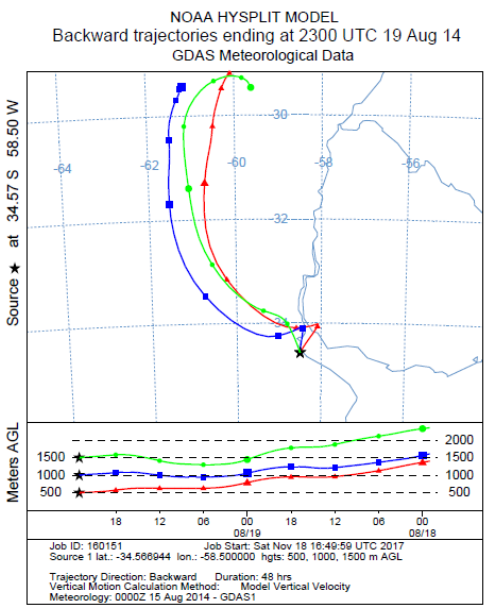


Figure 5.1: RSC at 1064 nm in arbitrary units from Villa Martelli, Argentina on 19 August 2014 (a) and 22 August 2014 (b), and from Aeroparque station (c), Argentina on 25 September 2017. The two red lines indicate the analyzed lidar data interval. The black solid line indicates the sun-photometer measurements.

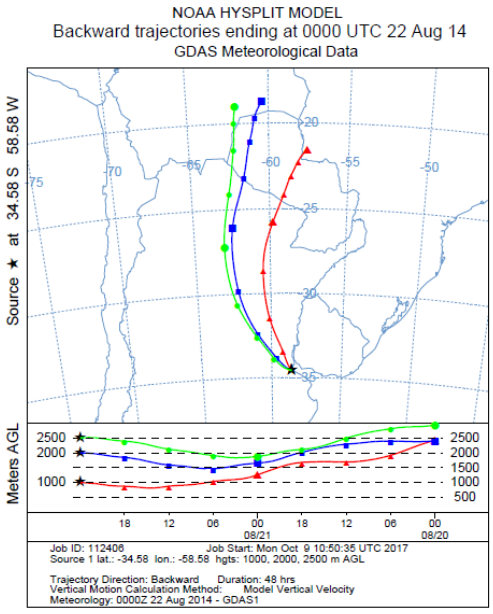
The two first cases selected correspond to an important event of biomass burning occurred in bordering countries in the north of Argentina in August 2014, particularly south of Brazil and Paraguay. It was detected in Buenos Aires, between 19 to 23 August. For the illustrations are used the measurements corresponding on August 19



(a)



(b)



(c)

Figure 5.2: a) Satellite image with hot spots corresponding to August 22, 2014, b) air mass back trajectories for the Villa Martelli measurement site on August 19, 2014 and c) air mass back trajectories for the Villa Martelli measurement site on August 22, 2014.

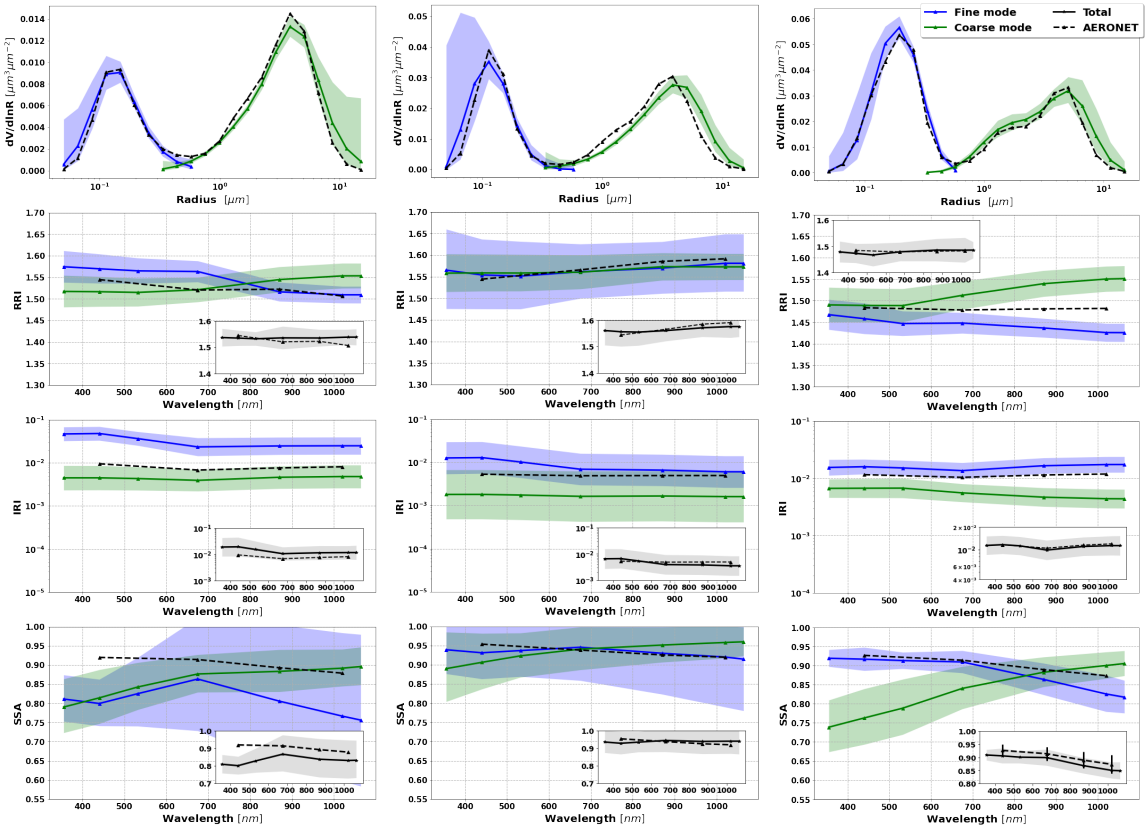
which present low aerosol load at 440 nm (~ 0.11) and SZA > 50 degrees. Figure 5.1a indicated the presence of several layers of aerosols up to 1.5 km. The lidar measure-

ments on August 19 were taken between 12:15 UTC and 12:45 UTC, and AERONET measurements correspond to 11:25 UTC. Figure 5.2 shows the HYSPLIT back trajectories that validate the source of the air masses. The measurements corresponding to August 22 are shown in Figure 5.1b, where several layers of aerosol up to 3 km are observed. For this day the aerosol load increases (AOD at $440\,nm \sim 0.31$) and the SZA is < 50 degrees. The inversion was realized with the average lidar data between 15 UTC to 15:20 UTC, and the AERONET measurements were considered to 16:59 UTC. The satellite image corresponding to August 22 (Fig. 5.2a) shows the presence of aerosols that extend from north of Argentina towards the center, passing through the province of Buenos Aires. Moreover, MODIS hotspots are detected in the satellite image. The source of the air masses can be validated from the HYSPLIT back trajectories (Fig. 5.2c). The last selected case corresponds to the biomass burning event on 25 September 2017, occurred in the north of Argentina and bordering countries (Figure 5.1c). In this work, lidar measurements from Aeroparque station between 19:20 UTC to 20:10 UTC, and the AERONET measurements corresponding to 19:20 UTC, whose AOD value at $440\,nm$ is 0.57 and SZA > 50 degrees were used.

Figure 5.3 illustrate the retrieved columnar properties for each day obtained by GRASP from a combination of radiometer and lidar data and the comparison with the corresponding standard AERONET retrievals. The results provided by GRASP are represented in solid lines: blue for the fine mode and green the coarse mode. Shaded areas represent the error provided by GRASP for each retrieved and derived property. Zoomed plots represent the effective refractive index and total SSA for GRASP (black solid line) and AERONET (black dashed line). From the illustrations can see almost all the GRASP retrieved properties in the three cases present good agreement with AERONET retrievals.

The error tendencies for SD that can be seen from Figure 5.3 agree with those identified above in present study and with results presented in previous sections and also with the results in some works as Dubovik et al. (2000) and Lopatin et al. (2013). For example, the retrieval errors clearly increase at the extremes of SD. Moreover, one clear and known tendency can be mentioned. The size distribution shift towards higher radii in the three cases could be explained by the use of lidar data in the inversions that provide additional information at scattering angles of 180 degrees (Lopatin et al. (2013); Bovchaliuk et al. (2016); Benavent-Oltra et al. (2017)). As it can be seen in the Figure 5.3, these deviations in almost all cases are included in the error (shaded areas).

The errors of RRI, IRI and SSA were retrieved for each mode separately by GRASP and they are significantly higher than the error for RRI, IRI and SSA of the total components. The effective RRI and IRI and the total SSA obtained by GRASP are in the middle of the retrieved values for fine and coarse mode separately. The total values



(a) August 19, 2014 **(b)** August 22, 2014 **(c)** September 25, 2017

Figure 5.3: Comparison of columnar properties retrieved by GRASP from a combination of sun/sky photometer and lidar data and retrieved conventionally by AERONET. SD, RRI, IRI and SSA retrieved by GRASP are shown in solid lines: blue (fine mode) and green (coarse mode). The shaded area in colors blue and green represent the total error provided by GRASP and black shaded areas are the uncertainties provided by AERONET. Zoomed panels show the RI effective and total SSA provided by GRASP (black solid line) and AERONET (black dashed line). Their associated errors are represented in grey sahded area for GRASP and with error bars for AERONET.

showed in the zoomed plots agree well with RRI, IRI and SSA provided by AERONET.

On the other hand, the case corresponding to September 25 has an AOD > 0.4 and SZA > 50 degrees allowing us to have the uncertainty of the SSA provided by AERONET. Thus, in the Figure 5.3 for this particular case we can observe the comparison of the uncertainty of SSA from AERONET and the SSA error provided by GRASP. Note, it is one advantage of GRASP that provide the errors for each parameter in all the situations. Regarding the values of SSA and the tendencies of their variability, the results show SSA representative of biomass burning cases. Namely the values of SSA decreasing with the wavelength agree with AERONET climatology by Dubovik et al. (2002b). As expected, based on the results of previous studies by Dubovik et al. (2000), Lopatin et al. (2013) and Tsekeri et al. (2017) the best agreements are obtained as the aerosol load increases. More specifically, we observe the

estimated errors of total SSA, in the two first cases increase. This can be associated with not favorable configurations of observation, on August 19 the AOD at $440\,nm$ is ~ 0.11 (SZA is > 50 degrees) and on August 22 the AOD at $440\,nm \sim 0.31$ (SZA is < 50 degrees). Thus, the measurements in situations with low amount of aerosol and with small SZA may not contain enough information to adequately retrieve the SSA (Dubovik et al. (2000), Lopatin et al. (2013); Torres et al. (2014)). Therefore, the case of September 25 corresponds to the most favorable situation for realizing reliable aerosol retrieval since AOD at $440\,nm$ value is > 0.4 and SZA is > 50 degrees. Indeed, the GRASP and AERONET retrievals have the best agreement for this day.

Figure 5.4 shows the retrieved vertical distributions of fine and coarse modes. The vertical structure of the aerosols of different types is clearly discriminated and shows good agreements with the back trajectory analysis for each day. Furthermore, the error estimates show good agreements with previous results provided in last sections for simulated cases.

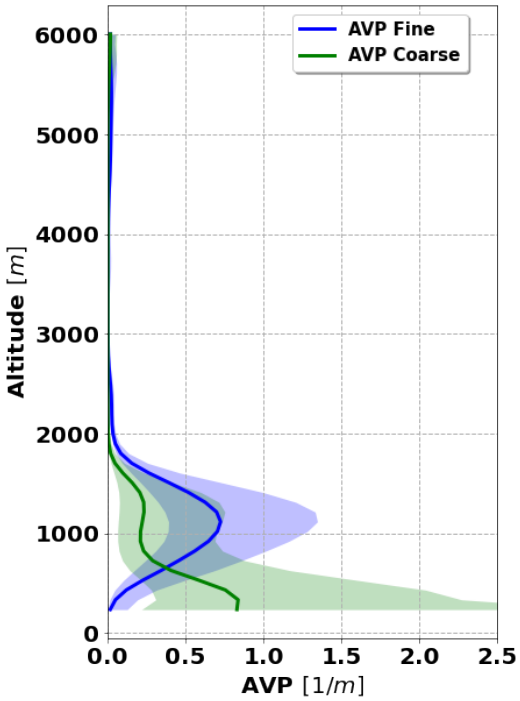
Moreover, Figure 5.5 show the retrieved LR and their error estimates using GRASP algorithm. The zoomed plots show the total LR provided by GRASP (black solid line) and by AERONET (black dashed line). The associated errors are represented in shaded areas for GRASP. As can be see only is possible to compare the values with AERONET retrieval for the Fig.5.5c which corresponds to higher aerosol load and SZA$> 50°$ on September 25, 2017.

Thus, the retrieved parameters and error estimates from GRASP application to the real data and their comparisons to the AERONET retrieval results showed an encouraging agreement between columnar properties of aerosol. At the same time, GRASP provide the error estimates for the retrieved properties in both fine and coarse mode and also for the total components. Moreover, GRASP has also the advantage of provide the dynamic error estimates in all the configurations. As was seen AERONET error estimates are only provided in some particular situations, when the AOD at 440 is greater than 0.4 and SZA> 50 degrees.

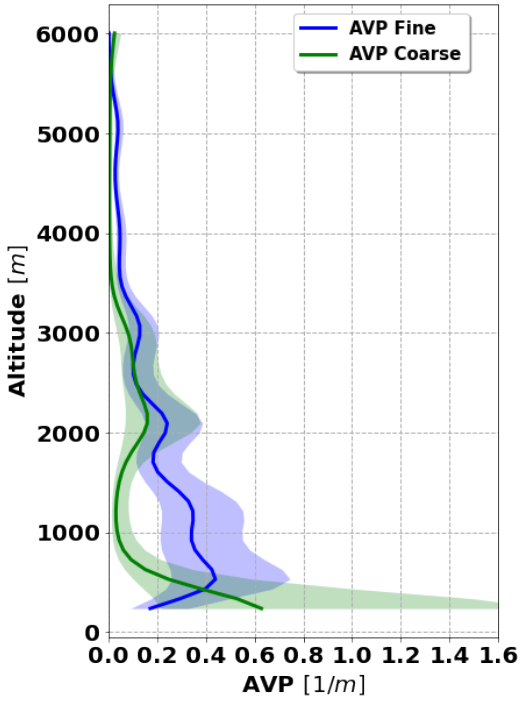
5.2 Conclusions

This chapter shows the capability of GRASP to provide not only retrieved parameters but also the dynamic error estimates for different configurations using real measurements.

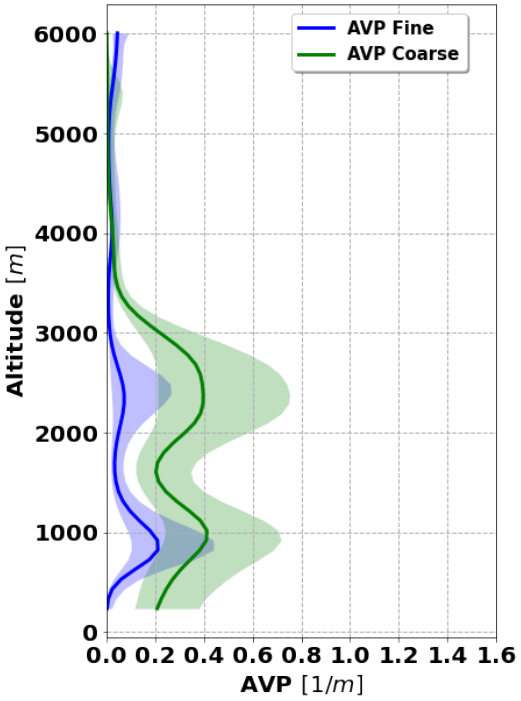
In this chapter the utilization of GRASP for deriving detailed aerosol properties and their error estimates was demonstrated for the coincident lidar and sun/sky photometer observations at Buenos Aires, Argentina. It was applied in three different cases of biomass burning at different aerosol loads (AOD(440) = 0.1, 0.3 and 0.5).



(a) August 19, 2014



(b) August 22, 2014



(c) September 25, 2017

Figure 5.4: Retrieved aerosol vertical profiles (AVP) by GRASP from a combination of sun/sky photometer and lidar data. Blue solid line represents the AVP fine mode and green the AVP coarse mode. The shaded areas correspond to the total error provided by GRASP.

The GRASP retrievals and the error estimates of the columnar aerosol properties showed to be fully adequate in comparative analysis with the aerosol products

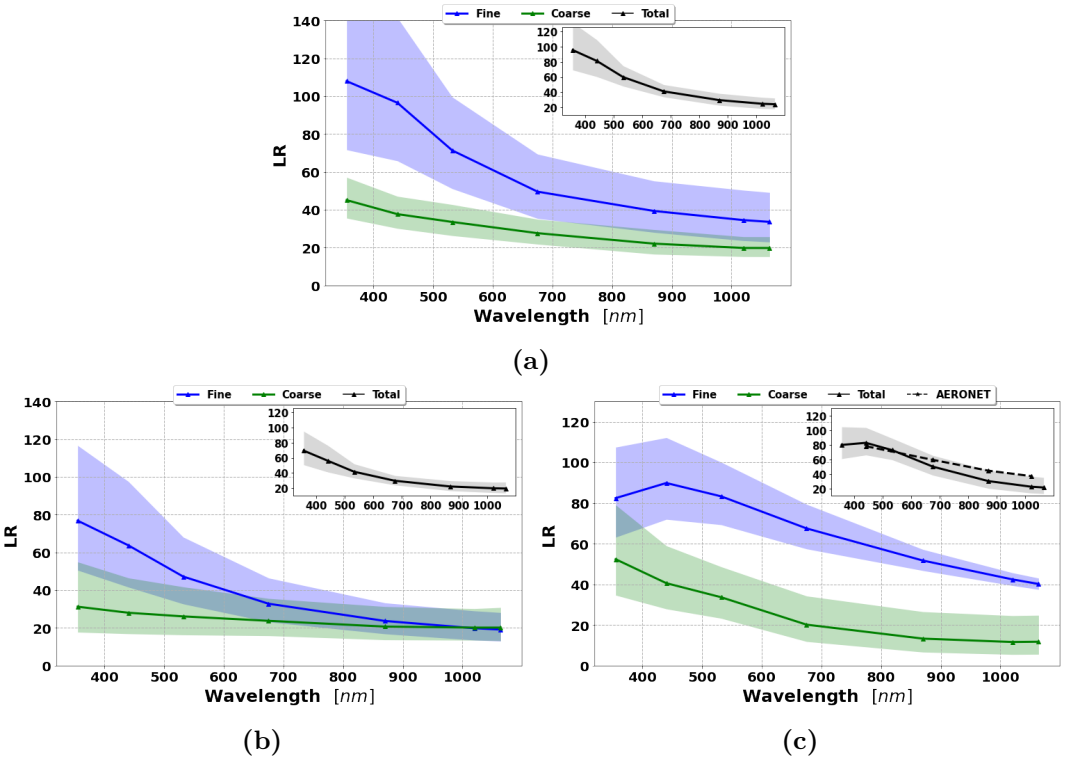


Figure 5.5: Retrieved LR by GRASP from a combination of sun/sky photometer and lidar data. Blue solid line represents the LR fine mode and green the LR coarse mode. Zommed plots show: in black solid lines the LR provided by GRASP and in dashed line the LR provided by AERONET. Their GRASP associated errors are represented in shaded areas.

available from AERONET operational retrievals. The retrieval of vertical profiles of fine and coarse aerosol modes showed consistent results with the expectations and the predictions of back-trajectories analysis.

Furthermore, GRASP has the capability to provide the dynamic error estimates for retrieved parameters in all the configurations and also for fine and coarse mode. It is an advantage since AERONET only provide the errors in the particular situation when the AOD at 440 is greater than 0.4 and SZA$>$ 50 degrees (for the Level 2 inversion products).

Chapter 6

GRASP error estimates using satellite observations: preliminary results

As it is widely known, the amount of available satellite data and the algorithms to obtain products from them have increased in the last two decades. In particular, multi-angular multi-spectral polarimetry (MAP) is proven to be one of the ideal measurements for aerosol and surface characterization from space. Different studies have been carried out from the use of different algorithms for the retrieval of aerosol and surface properties from this information (Hasekamp and Landgraf (2007); Waquet et al. (2009); Tanré et al. (2011); Dubovik et al. (2019); Chen et al. (2020))

However, one of the challenges is to provide the error estimates of these properties. As was already mentioned, there are several methods capable of providing error estimates and some of them have been mentioned for different instruments. Though some efforts have been already made, there are still some aspects to be analyzed.

GRASP has been adapted over the last decade for the operational processing of polarimetric satellite observations and several aerosol products for POLDER/PARASOL observations have been generated, which are released and publicly available (for example, `http://www.icare.univ-lille.fr` and `http://www.grasp-open.com/products/`). Moreover, the GRASP algorithm provides the error estimates for the retrieved and derived properties. Thus, an important further step is being able to provide the error estimates for the different POLDER/PARASOL products. Therefore, the objective of the current Chapter is the evaluation of the errors for the GRASP aerosol and surface products from the PARASOL-like measurements. Note, the analysis were made based in the Equations 3.52 and 3.53 for error estimates, which the systematic component of the error estimates is based in the misfit of the observation. This means that in the follow examples is not considered the improvements made in the systematic components,

as was shown in ground-based results. Therefore, these are preliminary results and more studies considering bias as explained in Chapter 4 are needed to have a complete study of the evaluation of error estimates in satellite observation applications.

First part of this chapter presents the description of the aerosol models and the configuration considered for the simulations. Then, the different tests and results are described. Moreover, a brief description of the correlation matrix for PARASOL-like measurements is illustrated. Finally, the application in real data is described and illustrated over Banizoumbou and Mongu.

6.1 Aerosol models and configurations for numerical tests

In order to evaluate the error estimates for atmospheric properties using satellite data, a series of numerical tests were conducted using synthetic data. They were performed based on the observation geometry, the aerosol and the surface characteristics for POLDER observations over land. The retrievals were considered using a multi-pixel approach for two configurations: GRASP/HP (5 log-normal bins) and GRASP/Models (5 aerosol models), which are explained below. A comparison with single pixel approach in the case of GRASP/HP configurations is provided with the objective to understand what is the contribution to multi-pixel approach in the retrievals and their error estimates.

The sensitivity tests were performed for two distinct and widely studied AERONET sites, Mongu and Banizoumbou, respectively. Particularly for this study, the sensitivity tests were carried out from real POLDER/GRASP retrievals over each site, i.e., to simulate the data was considered each configuration, GRASP/HP and GRASP/Models, as initial approach for the retrievals. In this way, we can get realistic PARASOL-like observations at each site for approximately months. For example, the retrievals for Banizoumbou were considered from January 4, 2008 to March 26, 2008 and for Mongu from June 2, 2008 to August 18, 2008.

The directly retrieved parameters such SD, RI and BRDF were used for the simulations from the retrievals when using GRASP/HP approach and only C_v and BRDF was considered from GRASP/Models retrievals. In this last approach, the values of the size distribution and the complex refractive index at each wavelength were assumed from AERONET retrieval climatology for each aerosol component from Dubovik et al. (2002b), since they are not part of directly retrieved parameters.

Some variations can be seen from these simulations since they are based on retrievals from real data. For example, the RI can vary quite significantly in time due to low sensitivity, especially if AOD is low on a respective day. Even, some variations in the

Table 6.1: Summary of POLDER/PARASOL measurements and the retrieved properties by GRASP algorithm under two different configurations: GRASP/HP and GRASP/Models (it is based on Table 1 Dubovik et al. (2011).

POLDER/PARASOL observations	
- $I(\Theta_j, \lambda_i)$ reflected total radiances - $Q(\Theta_j, \lambda_i)$ component of the Stokes vector - $U(\Theta_j, \lambda_i)$ component of the Stokes vector	
Retrieved characteristics	
GRASP/HP (5 bins)	**GRASP/Models (5 Models)**
Aerosol: - C_V - $dV(r_i)/d\ln r$ - C_{sph} - $n(\lambda_i)$ - $k(\lambda_i)$ - h_0	Aerosol: - C_V - h_0
Surface **Ross-Li model (Li and Strahler, 1992; Ross, 1981):** - BRDF parameters ($BRDF^*_{iso}$, $BRDF_{vol}$, $BRDF_{geom}$) **Maignan et al., 2009 model:** - BPDF	Surface **Ross-Li model (Li and Strahler, 1992; Ross, 1981):** - BRDF parameters ($BRDF^*_{iso}$, $BRDF_{vol}$, $BRDF_{geom}$) **Maignan et al., 2009 model:** - BPDF

* $BRDF_{iso}$ is called BRDF1 in the chapter.

BRDF with respect to the pixels can be seen since BRDF can be different even if the pixels are neighbours. The advantage in these simulations is that they provide a more evenly spread possible values range. The aerosol loadings values present in the sensitivity tests cover the range from $AOD(440\,nm) = 0.01$ to $AOD(440\,nm) = 1$.

In this application, only the random noise was considered to perturb the simulated data. Future additional studies will be needed to evaluate the error estimates adding biases. The random noise at the level of 3% for intensity and 0.01 for the polarization components Q and U of PARASOL signal have been added to the simulated synthetic PARASOL observations. Moreover, a filtering criteria of the 'residual relative' (mean root square of relative error in fitting the measurements by the algorithm) smaller than 3% was used in this work, to select the most reliable retrievals. It corresponds to the

same value as the assumed random noise of 3%.

These synthetic PARASOL-like observations were inverted by GRASP using two retrieval scenarios: conventional single-pixel and multi-pixel retrieval (more details are presented in Chapter 3). In the first scenario all synthetic observations were inverted fully independently. In both cases the simulated data corresponds to 4 pixels with 30 number of times (it is 2x2x30 pixels that corresponds to N_xxN_yxN_t).

Two different approaches have been used to simulate the data and also to invert them: GRASP/HP (High Precision) and GRASP/Models, in each case respectively. Table 6.1 describes the retrieved aerosol parameters of both approaches. In the first approach, GRASP/HP, the size distribution can be understood as a 5 bin simplification of the AERONET 22 bins retrieval conceived for satellite application. On the other hand, GRASP/Models approach further simplifies the assumption of external mixing of 5 components. The significant reduction of retrieved parameters makes it faster and gives good performances for satellite application, especially for single-viewing measurement. A brief description is provided below and more details can be found in Chen et al. (2020), Lopatin et al. (2021) and Dubovik et al. (2021).

6.1.1 Retrieval configuration for POLDER/GRASP

As was already mentioned in different chapters, GRASP has the flexibility of modeling aerosol and surface. It is important since in some situations the information content is limited and it might be desirable to add additional assumptions and reduce the number of directly retrieved parameters.

In this way, as described in Dubovik et al. (2011), Chen et al. (2020) and Lopatin et al. (2021), different retrieval configurations can be used. Particularly, in this chapter only two configurations are used. However, a brief description of the three configurations is provided below:

- PARASOL/GRASP Optimized: in the sense of radiative transfer calculations were optimized to find the best trade-off between speed of processing and accuracy of results;
- PARASOL/GRASP High-precision: radiative transfer calculations with high precision were used;
- PARASOL/GRASP Models: the simplest, fastest processing; aerosol is assumed to be an external mixture of several aerosol models.

The first two configurations are based on the same concept, in which the retrievals were performed using one aerosol component model with five bins of the size distribution and with spectrally dependent complex refractive index.

On the other hand, the Models approach uses different assumptions for modeling aerosol properties. The aerosol can be assumed to be an external mixture of several aerosol components:

$$n(\lambda) = \sum_{i=1}^{K} c_i \, n_i(\lambda) \quad \text{and} \quad k(\lambda) = \sum_{i=1}^{K} c_i \, k_i(\lambda) \tag{6.1}$$

or as an internal mixture of refractive indices of different components:

$$n(\lambda) = n_{mix}[c_1; c_2; ...; n_1(\lambda); n_2(\lambda); ...] \quad \text{and} \quad k(\lambda) = k_{mix}[c_1; c_2; ...; k_1(\lambda); k_2(\lambda); ...]. \tag{6.2}$$

In this configuration, only the respective concentrations are retrieved together with aerosol layer height (ALH) and spectral BRDF and BPDF parameters. Particularly, in this study is used the 5 models as is explained in Chen et al. (2020) and Lopatin et al. (2021): fine medium absorbing, fine non-absorbing, coarse spherical, coarse non-spherical and fine absorbing. The correspondent size distribution, complex refractive index and non-sphericity parameter for each aerosol model are derived from the results of AERONET aerosol climatology for the main distinct aerosol types (Dubovik et al. (2002b)) and improved in a series of sensitivity tests with satellite data as is described in Lopatin et al. (2021).

Note, GRASP/HP employs the most accurate radiative transfer calculations, while GRASP/Optimized and GRASP/Models are optimized to achieve the best trade-off between accuracy and speed. More details can be found in Dubovik et al. (2011); Chen et al. (2020); Lopatin et al. (2021) and Dubovik et al. (2021).

6.2 Test results for POLDER/PARASOL simulated observations

Series of tests have been performed to evaluate error estimates for inverting PARASOL-like measurements using the GRASP algorithm over Banizoumbou and Mongu. The set of main aerosol retrieved and derived parameters include:

- AOD, SSA, RRI and BRDF1 in the case of GRASP/HP configurations, and
- AOD, SSA and BRDF1 in the case of GRASP/Models.

RRI is not discussed for Models since it is not part of the retrieved parameters.

The first results of the sensitivity tests are organized as follows: i) single pixel with no noise added, ii) single pixel with adding noise and iii) multi-pixel with adding noise.

The main objective is to illustrate the differences using single-pixel and multi-pixel approach in situations where random noise is present.

Two types of illustrations can be seen in each figure throughout this section. At the top of each figure are the correlations between the assumed values and the retrieval provided by GRASP for each property at 443 nm and 670 nm. The x-axis represents the assumed values and the y-axis the retrieved property with their respective error bars provided by GRASP. The black solid line and the red solid line are the 1 : 1 reference line and the linear regression line. At the bottom, the scatter plots are illustrated with the errors of the retrievals depending on the AOD at different wavelengths. Each of them are represented by colors: 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red).

Figures 6.1 to 6.6 present the results of retrieving PARASOL-like synthetic measurements using GRASP/HP (High-precision) configurations.

Figures 6.1 to 6.3 show the results of the tests for retrieving aerosol and surface parameters over Banizoumbou from synthetic measurements. Top panels illustrate the AOD, SSA, RRI and BRDF1 at 443 nm and 670 nm. Figure 6.1 corresponds to the single-pixel with no noises added. Then, Figures 6.2 and 6.3 show the results of the retrieval considering random noise for I $\sim$ 3% and for the polarization components Q and U $\sim$ 0.01 in the case of single-pixel approach and multi-pixel approach respectively. At the bottom, the figures present scatter plots of the error estimates for each retrieved parameter, separated by wavelengths: in green 443 nm, in black 490 nm, in blue 670 nm and 1020 nm in red. This allows to see the dependency of the errors with the AOD for each wavelength. Figures 6.4 to 6.6 are analogous to Figures 6.1 to 6.3 in the case of simulated observations over Mongu.

The analysis of the results show that in situations where no noise is added, all properties are retrieved rather accurately, low values of error estimates agree with strong correlations (R) and low values of RMSE for each property (Figures 6.1 and 6.4). In general, the retrievals of RRI and BRDF1 parameters become unstable, particularly at short wavelengths even when no noise is added. All these tendencies become pronounced once the random noise is added. Moreover, in the cases of single-pixel approach and adding random noise almost all the retrieved properties deteriorate in most of the cases and big values of error estimates for each parameter can be seen. For example, very low correaltions and big dispersion of values can be seen for RRI in the Figures 6.2c and 6.5c. SSA also present big dispersion of values mainly in Fig. 6.5b over Mongu and BRDF1 over Banizoumbou (Fig. 6.2d). The error estimates in all these cases also take large values, which is expected for such retrievals. Moreover, the AOD retrievals and the error estimates are rather reliable even when random noise is added.

Figures 6.3 and 6.6 show a significant improvement of the retrieved parameters

and error estimates using a multi-pixel approach. As was explained by Dubovik et al. (2011), this is an expected tendency since these additional restrictions allow the propagation and consolidation of useful information from different observation situations. For example, in the situations with low aerosol loading the satellite observes mainly surface reflectance properties. Correspondingly, once the constraints limiting time variability of the surface reflectance are applied, this information is supplied into the interpretations of observations corresponding to moderate and high aerosol loading over the same pixel. Similarly, the constraints of horizontal variability of aerosol properties help to improve the retrieval of aerosol by benefiting from observations of the same or similar aerosol properties over several pixels with somewhat different conditions of observations (geometry, surface reflectance, aerosol loading).

Thus, as can be seen, in the presence of noise the multi-pixel configuration presents the best results for GRASP/HP retrievals. Therefore, the results for GRASP/Models retrievals are directly illustrated using multi-pixel configuration.

Figures 6.7 and 6.8 show the correlations of the retrieved parameters with their error estimates for the inversion of GRASP/Models configuration. These cases represent the retrieval using a multi-pixel approach and adding random noise (I $\sim$ 3% and Q and U $\sim$ 0.01). Note that in the Figures 6.7 and 6.8 are illustrated the AOD, SSA and BRDF1.

The results, in the cases of multi-pixel approach with added noises, show high correlations for all the illustrated properties and present good agreements with the error estimates since the errors provided by GRASP are comparable with the difference between the assumed and the retrieved values (also called 'actual values'). However, some pixels show larger errors. These correspond to a low aerosol load.

On the other hand, the scatter plots at the lower part of the figures (Figs. 6.1 to 6.8) provide more detailed information about the errors. These illustrations present the behavior of the error of each parameter as a function of the AOD and the wavelengths. In general, it can be noted that the AOD errors and SSA errors increase as the AOD decreases. Similar tendencies can be founded for the RRI, particularly in Fig. 6.3 and 6.6 where multi-pixel is considered. In the cases of single pixel with added noise, the RRI error present big dispersion of values.

The highest dispersions of both retrieval and error are observed in BRDF1 mainly at 443 nm and 490 nm showing the highest errors at Banizoumbou station. Moreover, BRDF1 error increases with the AOD. This is because the radiance arriving to the satellite comes from aerosol scattering and ground reflectance. Thus, in the situations with low aerosol loading the satellite observes mainly surface reflectance properties.

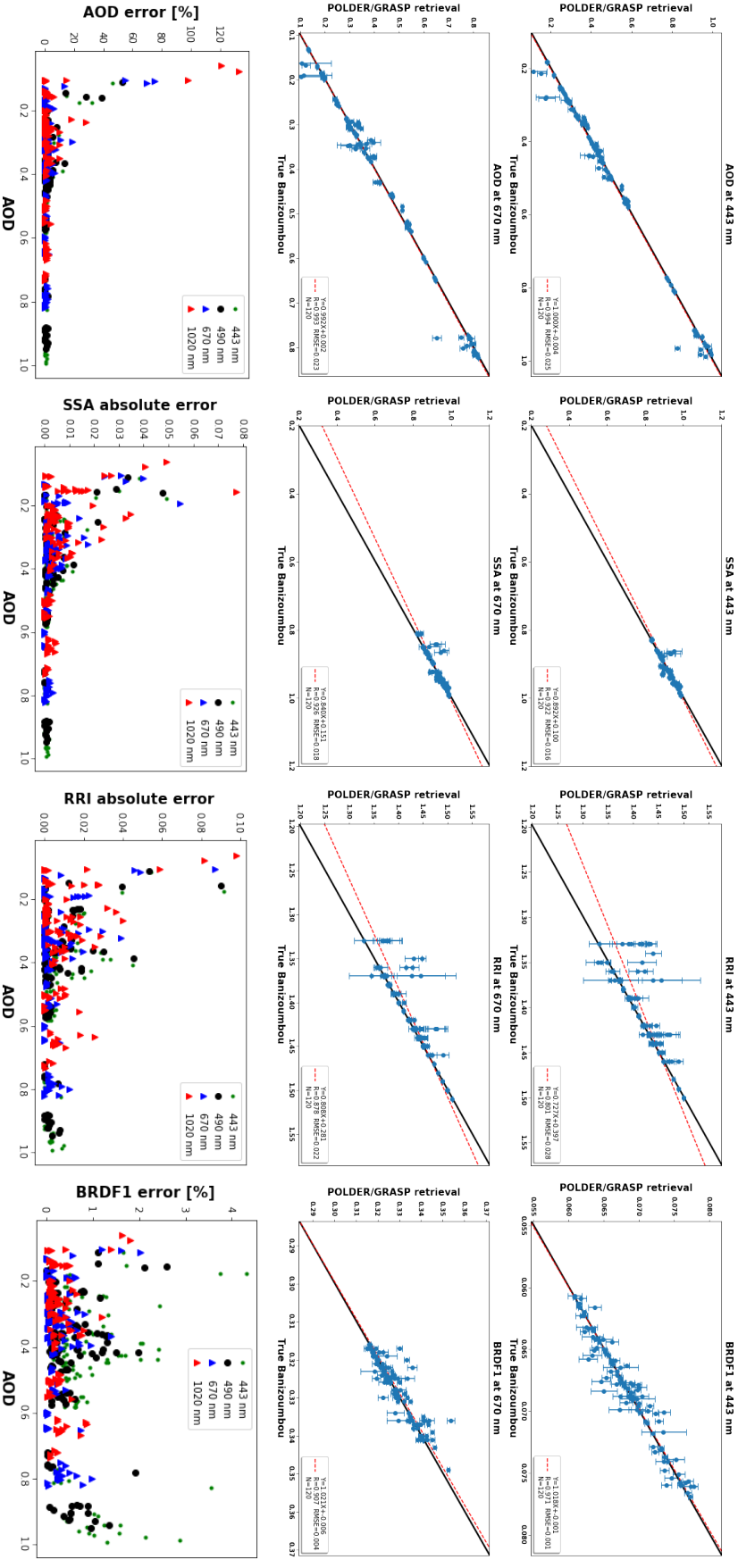


Figure 6.1: Validation of retrieving PARASOL-like synthetic measurements over Banizoumbou using GRASP/HP configuration for each property at 443 nm and 670 nm. The retrievals are considered for single-pixel approach without random noise. At the top are the correlations for each property and the bottom show their respective scatter plots at 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red): a) AOD; b) SSA; c) RRI and d) BRDF1.

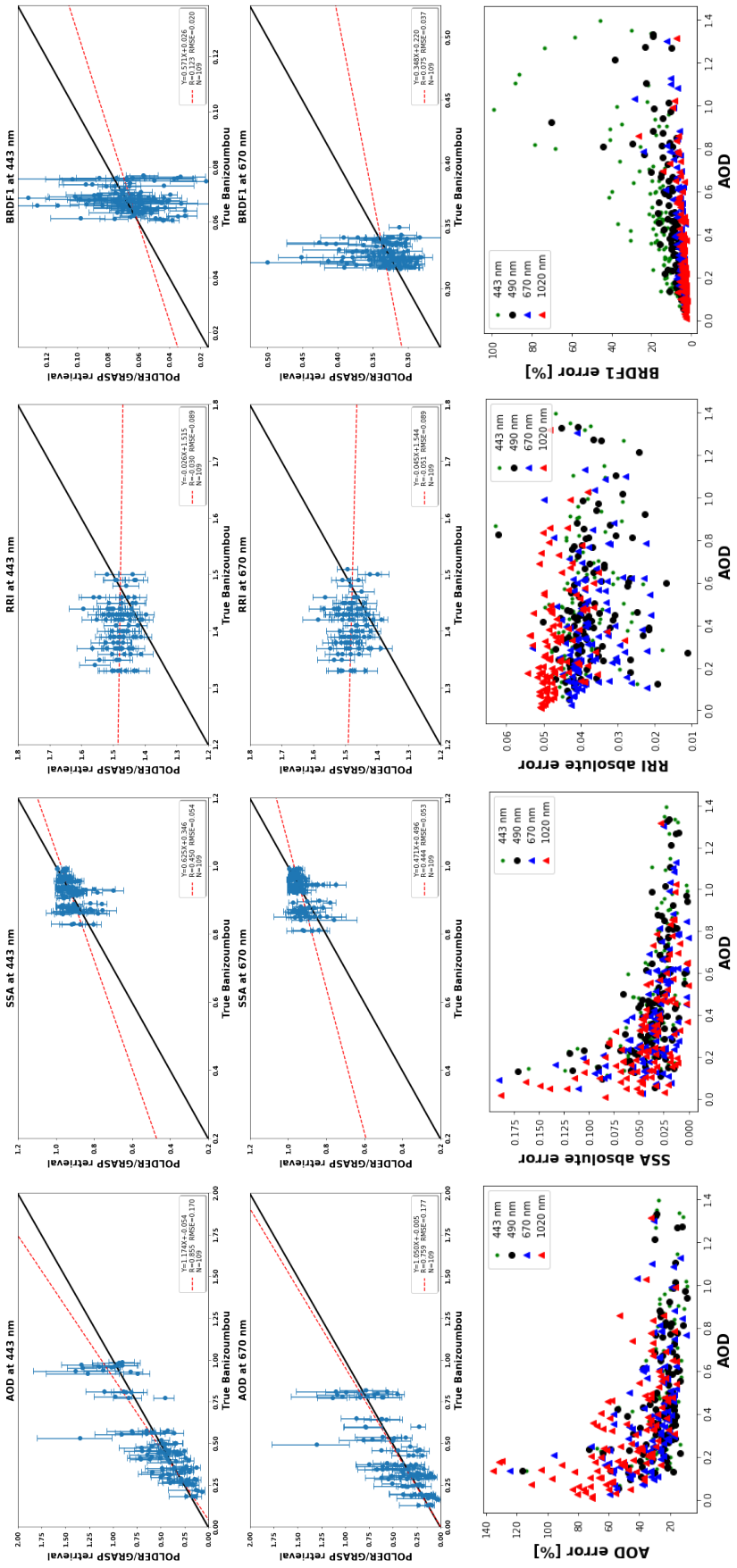


Figure 6.2: Validation of retrieving PARASOL-like synthetic measurements over Banizoumbou using GRASP/HP configuration for each property at 443 nm and 670 nm. The retrievals are considered for single-pixel approach under random noise I$\sim$ 3% and polarization components Q and U $\sim$ 0.01. At the top are the correlations for each property and the bottom show their respective scatter plots at 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red): a) AOD; b) SSA; c) RRI and d) BRDF1.

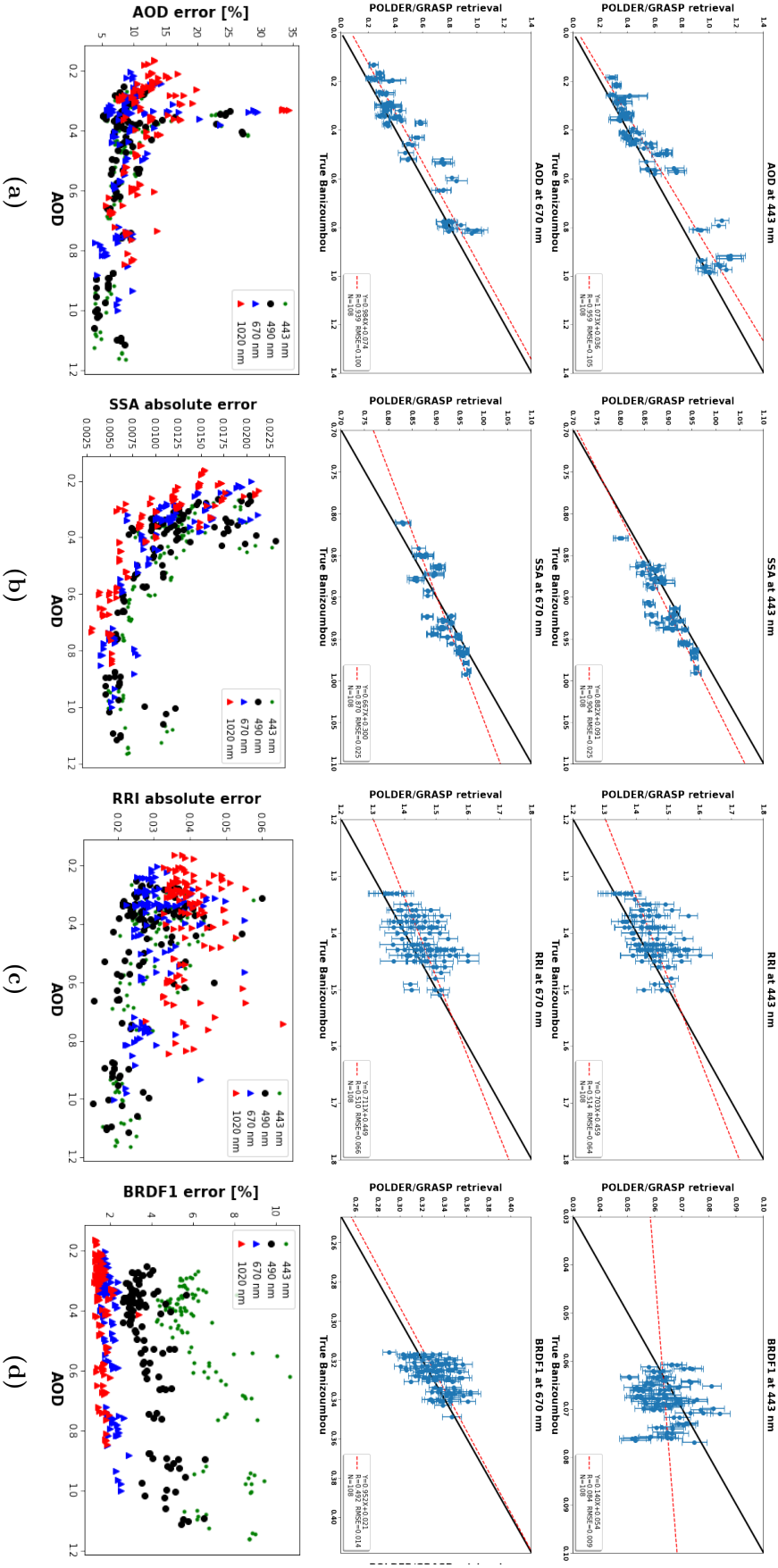
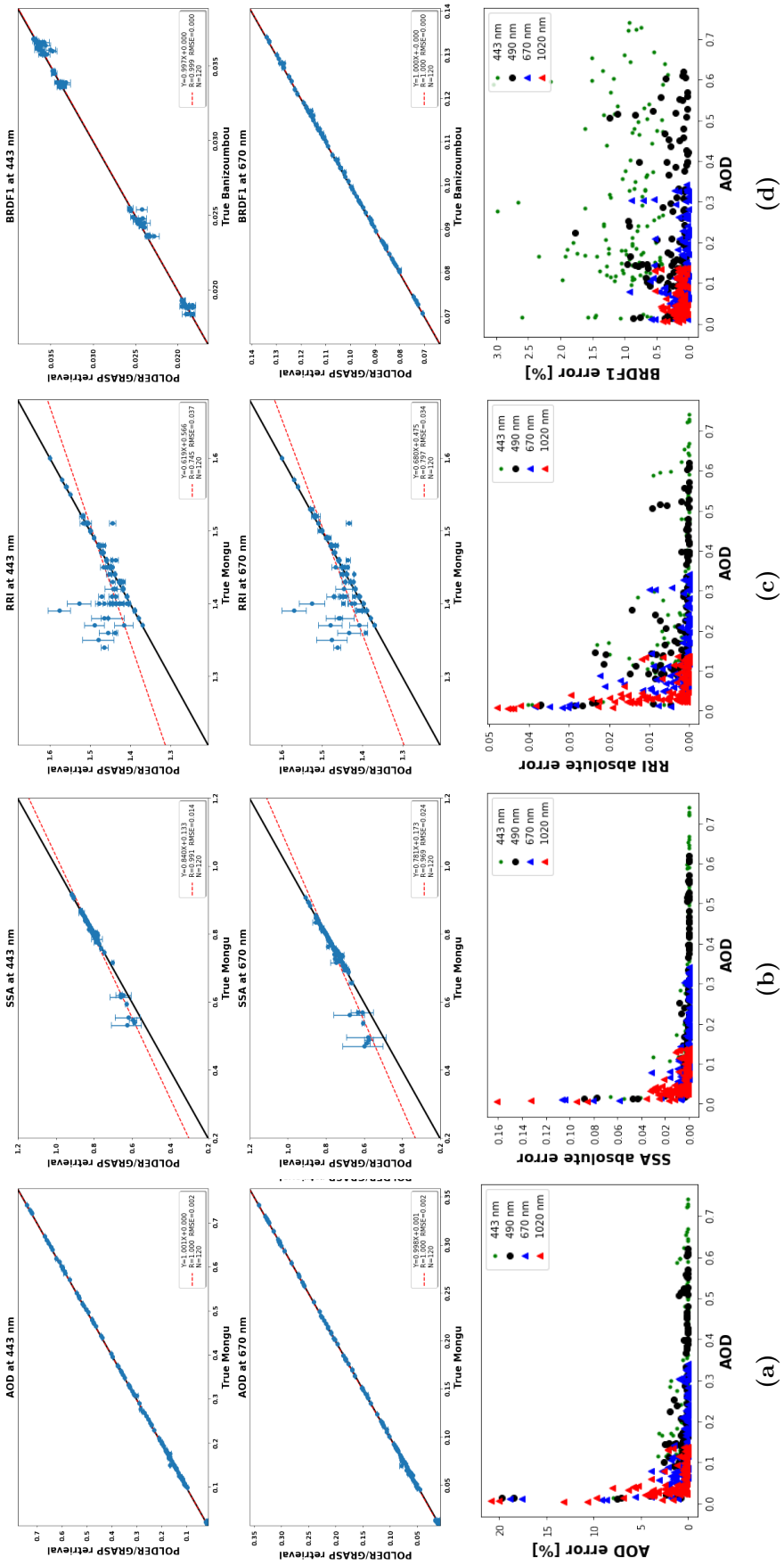
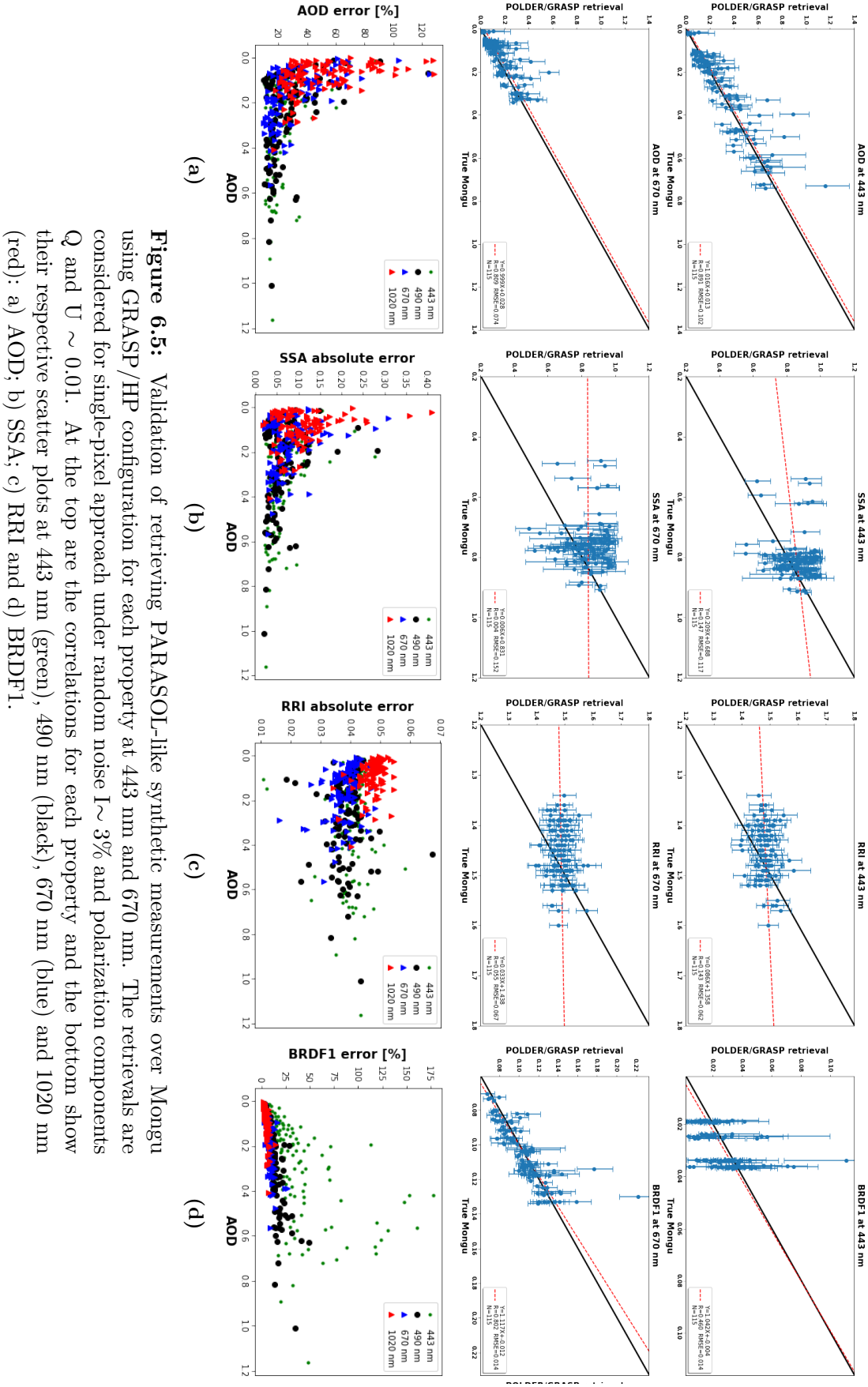


Figure 6.3: Validation of retrieving PARASOL-like synthetic measurements over Banizoumbou using GRASP/HP configuration for each property at 443 nm and 670 nm. The retrievals are considered for multi-pixel approach under random noise I∼ 3% and polarization components Q and U ∼ 0.01. At the top are the correlations for each property and the bottom show their respective scatter plots at 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red): a) AOD; b) SSA; c) RRI and d) BRDF1.



(a) (b) (c) (d)

Figure 6.4: Validation of retrieving PARASOL-like synthetic measurements over Mongu using GRASP/HP configuration for each property at 443 nm and 670 nm. The retrievals are considered for single-pixel approach without random noise. At the top are the correlations for each property and the bottom show their respective scatter plots at 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red): a) AOD; b) SSA; c) RRI and d) BRDF1.



(a) (b) (c) (d)

Figure 6.5: Validation of retrieving PARASOL-like synthetic measurements over Mongu using GRASP/HP configuration for each property at 443 nm and 670 nm. The retrievals are considered for single-pixel approach under random noise I$\sim$ 3% and polarization components Q and U $\sim$ 0.01. At the top are the correlations for each property and the bottom show their respective scatter plots at 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red): a) AOD; b) SSA; c) RRI and d) BRDF1.

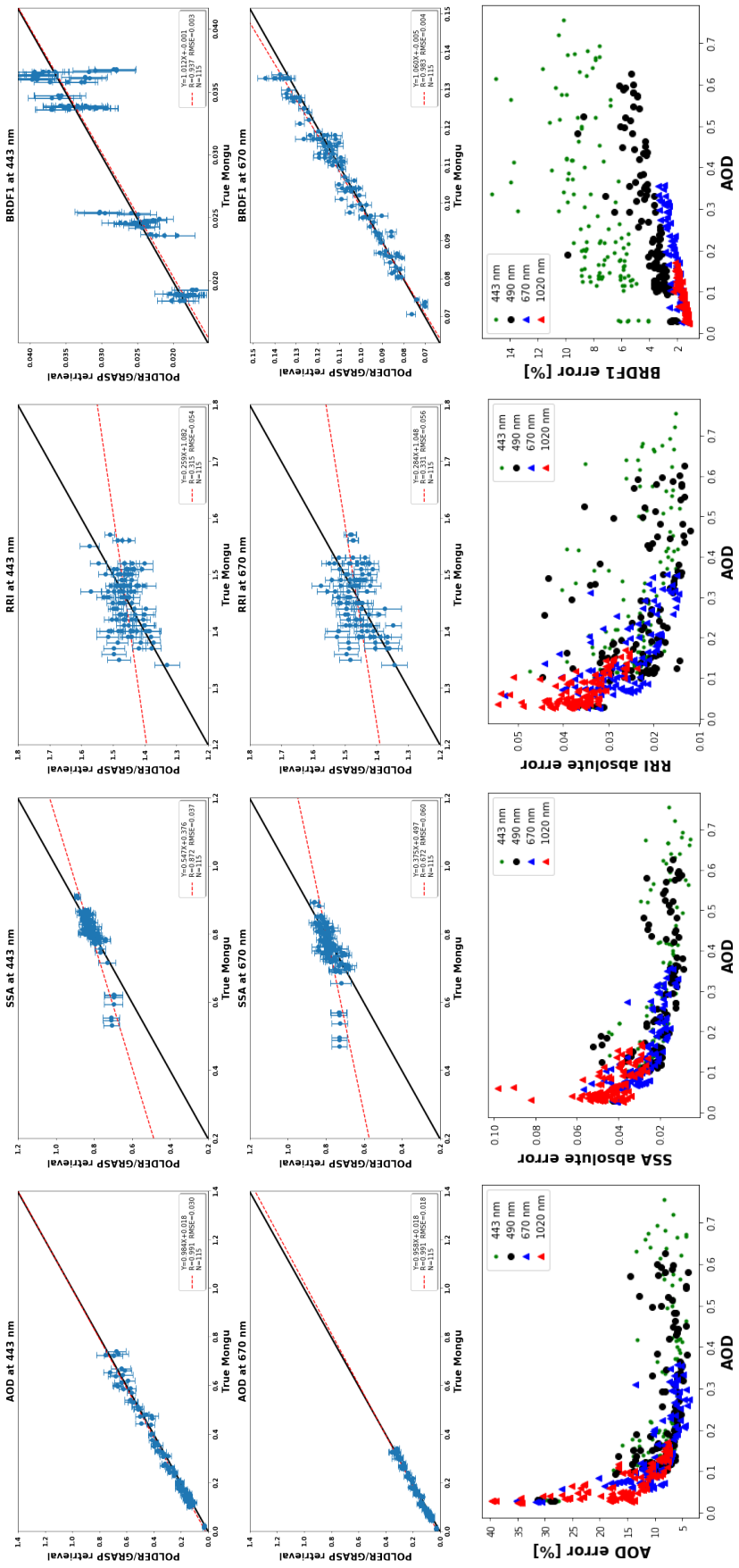
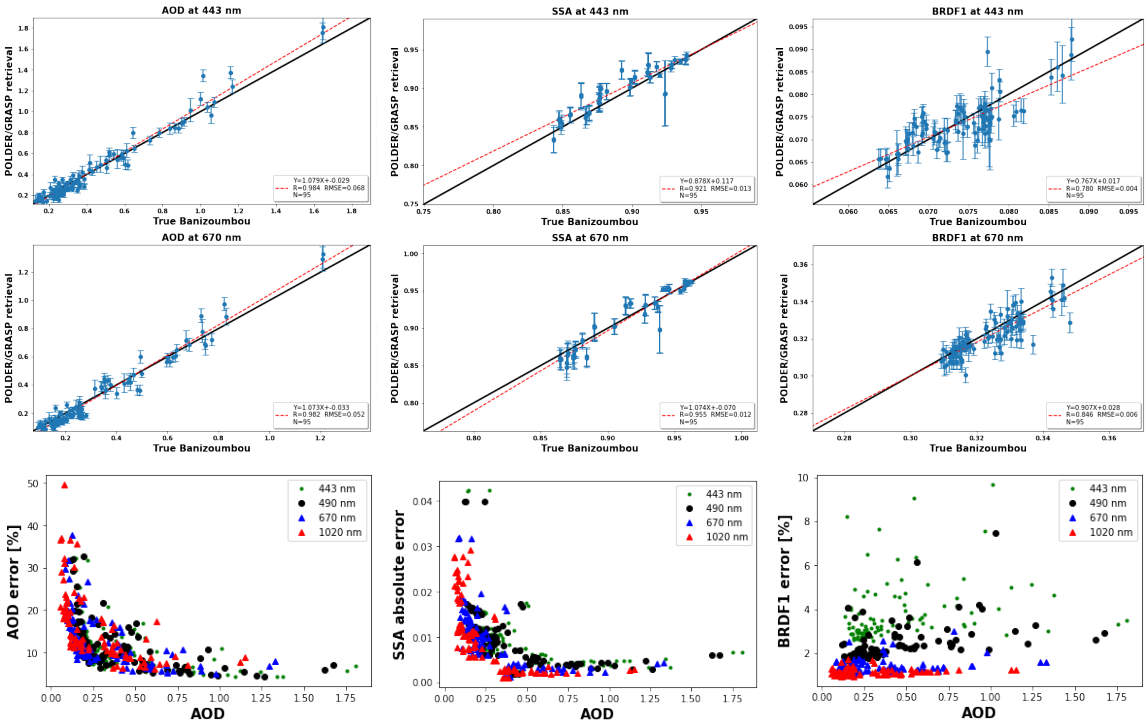
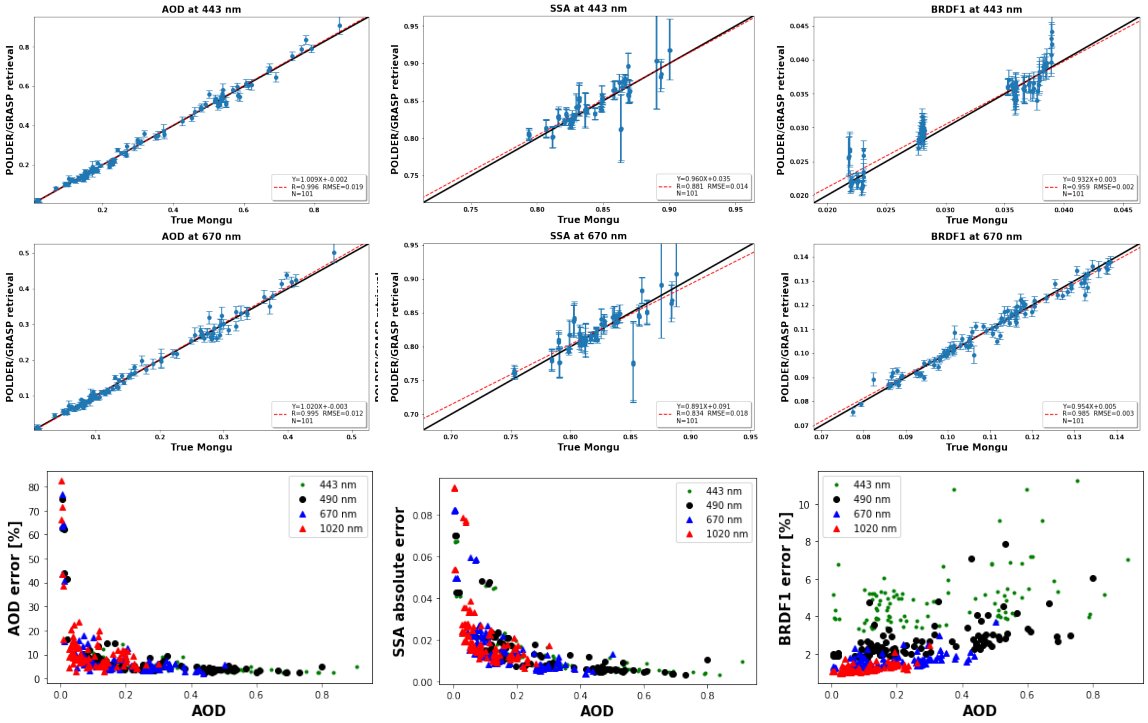


Figure 6.6: Validation of retrieving PARASOL-like synthetic measurements over Mongu using GRASP/HP configuration for each property at 443 nm and 670 nm. The retrievals are considered for multi-pixel approach under random noise I$\sim$ 3% and polarization components Q and U $\sim$ 0.01. At the top are the correlations for each property and the bottom show their respective scatter plots at 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red): a) AOD; b) SSA; c) RRI and d) BRDF1.



(a) **(b)** **(c)**

Figure 6.7: Validation of retrieving PARASOL-like synthetic measurements over Banizoumbou using GRASP/Models configuration for each property at 443 nm and 670 nm. The retrievals are considered for multi pixel approach under random noise I$\sim$ 3% and polarization components Q and U $\sim$ 0.01. At the top are the correlations for each property and the bottom show their respective scatter plots at 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red): a) AOD; b) SSA; c) RRI and d) BRDF1.



(a) **(b)** **(c)**

Figure 6.8: Validation of retrieving PARASOL-like synthetic measurements over Mongu using GRASP/Models configuration for each property at 443 nm and 670 nm. The retrievals are considered for multi pixel approach under random noise I$\sim$ 3% and polarization components Q and U $\sim$ 0.01. At the top are the correlations for each property and the bottom show their respective scatter plots at 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red): a) AOD; b) SSA; c) RRI and d) BRDF1.

In order to quantify the results, Tables 6.2 to 6.8 summarize the performance of the retrievals in the most robust cases, multi-pixel approach, for both GRASP/HP (5 log-normal bins) and GRASP/Models (5 aerosol Models) configurations over land in Banizoumbou and Mongu. In the Tables are represented the different metrics values which allow quantify the results: correlation (R), Slope, Offset, root-mean-square error (RMSE), Bias, standard deviation (STD) and GRASP/Errors.

Focus on the analysis of the retrieved parameters for the multi-pixel approach, it is possible to see that in the case of Mongu the best results are obtained, showing the best correlations, low RMSE values and error estimates that agree with those provided by Dubovik et al. (2011).

Overall, based on the obtained metrics (R, RMSE, STD, Bias and GRASP/Errors), the configuration of GRASP/Models and GRASP/HP present slight differences between them, though both show stability in the retrieved solutions. For example, over Mongu observations, for AOD at $443\,nm$ GRASP/Models shows slight better performance than GRASP/HP: R = 0.996 compared to 0.991 and RMSE = 0.019 for GRASP/Models compared to 0.029 for GRASP/HP at $443\,nm$ (see in Figures 6.6 and 6.8). In general, error estimates have good agreements with the differences of the assumed and the retrieved values. It can be seen from the following metrics: bias, std and GRASP/Error. For example, for the same case already described, the metrics for GRASP/Models are $\sim 0.0011 \pm 0.019$ and error estimates take values ~ 0.014, while in GRASP/HP the metrics are $\sim 0.014 \pm 0.026$ and error estimates are ~ 0.024.

These errors, in the Tables 6.2 to 6.8, are provided as reference values since a complete study is necessary taking into account the possible bias in the measurements. Moreover, Figures 6.9 and 6.10 show the error variances for AOD and SSA for the simulated data over Banizoumbou and Mongu for the retrieval of multi-pixel configuration (the retrievals represent one pixel in each case). In the figures can be seen the assumed values in solid lines, the retrievals in dashed lines and the error estimates in shaded areas showing good agreements between the retrievals and the assumed values. More studies are needed assuming also bias in order to provide a complete description and also the illustration of all the retrieved parameters.

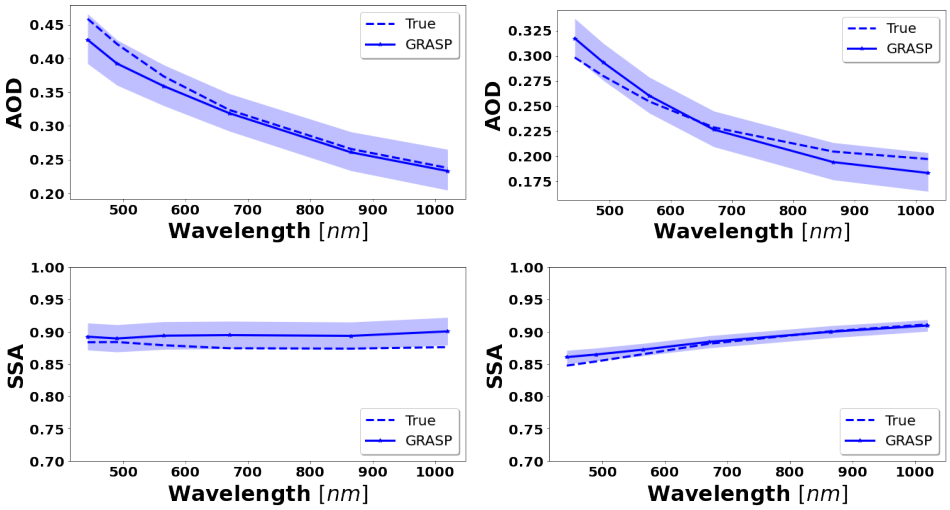


Figure 6.9: AOD and SSA retrieved for PARASOL-like synthetic measurements from simulated sun/sky photometer data with assumed random noise in I $\sim$ 3% and in the polarization components Q and U $\sim$ 0.01 for multi-pixel approach over Banizoumbou. a) using GRASP/HP and b) using GRASP/Models. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates total errors provided by GRASP algorithm.

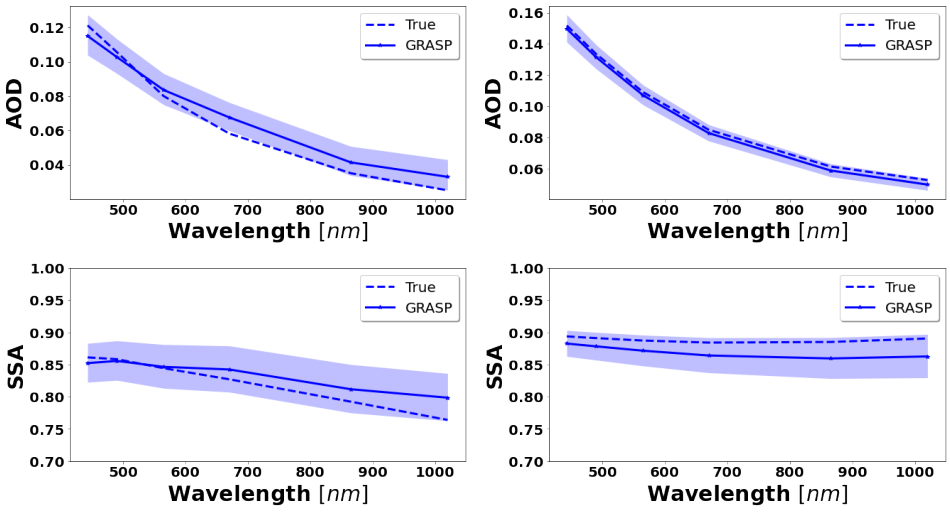


Figure 6.10: AOD and SSA retrieved for PARASOL-like synthetic measurements from simulated sun/sky photometer data with assumed random noise in I $\sim$ 3% and in the polarization components Q and U $\sim$ 0.01 for multi-pixel approach over Mongu. a) using GRASP/HP and b) using GRASP/Models. The solid lines are the simulated properties (SD, RRI, IRI and SSA), the dashed lines are the retrieved parameters. The shaded area indicates total errors provided by GRASP algorithm.

Table 6.2: Global statistics of PARASOL-like synthetic measurements using GRASP algorithm for AOD in Mongu over land.

Band (nm)	Products	R	Slope	Offset	RMSE	Bias	Std	Grasp Error
443 nm	**5 bins (HP)**	0.991	0.984	0.018	0.029	0.014	0.026	0.024
	Models	0.996	1.009	-0.0018	0.019	0.0011	0.019	0.014
490 nm	**5 bins (HP)**	0.992	0.981	0.018	0.025	0.014	0.021	0.020
	Models	0.996	1.011	-0.0019	0.017	0.0010	0.017	0.013
565 nm	**5 bins (HP)**	0.991	0.989	0.015	0.021	0.013	0.017	0.016
	Models	0.996	1.014	-0.0022	0.014	0.0009	0.014	0.011
670 nm	**5 bins (HP)**	0.991	0.958	0.018	0.018	0.012	0.013	0.013
	Models	0.995	1.020	-0.0025	0.012	0.0009	0.012	0.009
865 nm	**5 bins (HP)**	0.974	0.967	0.016	0.018	0.013	0.013	0.01
	Models	0.991	1.033	-0.0029	0.011	0.0009	0.011	0.008
1020 nm	**5 bins (HP)**	0.963	0.933	0.015	0.016	0.012	0.011	0.009
	Models	0.988	1.043	-0.0032	0.011	0.0009	0.011	0.008

Table 6.3: Global statistics of PARASOL-like synthetic measurements using GRASP algorithm for AOD in Banizoumbou over land.

Band (nm)	Products	R	Slope	Offset	RMSE	Bias	Std	Grasp Error
443 nm	**5 bins (HP)**	0.959	1.073	0.036	0.105	0.071	0.077	0.045
	Models	0.984	1.078	-0.029	0.068	0.006	0.067	0.041
490 nm	**5 bins (HP)**	0.955	1.047	0.048	0.102	0.069	0.074	0.043
	Models	0.984	1.077	-0.030	0.062	0.002	0.062	0.038
565 nm	**5 bins (HP)**	0.949	1.015	0.062	0.099	0.069	0.072	0.043
	Models	0.983	1.075	-0.032	0.056	-0.003	0.056	0.035
670 nm	**5 bins (HP)**	0.938	0.984	0.074	0.099	0.068	0.073	0.043
	Models	0.982	1.073	-0.033	0.052	-0.009	0.051	0.032
865 nm	**5 bins (HP)**	0.911	0.940	0.086	0.103	0.065	0.079	0.044
	Models	0.979	1.072	-0.035	0.051	-0.014	0.049	0.032
1020 nm	**5 bins (HP)**	0.893	0.908	0.094	0.103	0.063	0.082	0.046
	Models	0.978	1.072	-0.037	0.052	-0.017	0.049	0.032

Table 6.4: Global statistics of PARASOL-like synthetic measurements using GRASP algorithm for SSA in Mongu over land.

Band (nm)	Products	R	Slope	Offset	RMSE	Bias	Std	Grasp Error
443 nm	**5 bins (HP)**	0.872	0.547	0.376	0.037	0.010	0.036	0.022
	Models	0.881	0.960	0.035	0.013	0.0019	0.013	0.014
490 nm	**5 bins (HP)**	0.855	0.502	0.413	0.041	0.013	0.039	0.022
	Models	0.873	0.946	0.046	0.014	0.0017	0.014	0.015
565 nm	**5 bins (HP)**	0.771	0.446	0.447	0.049	0.015	0.036	0.025
	Models	0.858	0.922	0.066	0.016	0.0013	0.016	0.016
670 nm	**5 bins (HP)**	0.672	0.375	0.496	0.060	0.023	0.055	0.028
	Models	0.834	0.890	0.091	0.018	0.0008	0.018	0.017
865 nm	**5 bins (HP)**	0.527	0.266	0.566	0.084	0.045	0.071	0.036
	Models	0.807	0.872	0.104	0.023	-0.0001	0.023	0.021
1020 nm	**5 bins (HP)**	0.375	0.198	0.597	0.103	0.061	0.084	0.043
	Models	0.798	0.885	0.092	0.027	-0.0009	0.027	0.023

Table 6.5: Global statistics of PARASOL-like synthetic measurements using GRASP algorithm for SSA in Banizoumbou over land.

Band (nm)	Products	R	Slope	Offset	RMSE	Bias	Std	Grasp Error
443 nm	**5 bins (HP)**	0.904	0.882	0.091	0.025	-0.018	0.017	0.011
	Models	0.921	0.877	0.117	0.013	0.007	0.011	0.010
490 nm	**5 bins (HP)**	0.897	0.833	0.137	0.025	-0.017	0.018	0.011
	Models	0.931	0.919	0.077	0.012	0.005	0.011	0.010
565 nm	**5 bins (HP)**	0.891	0.762	0.206	0.024	-0.013	0.200	0.011
	Models	0.944	0.989	0.012	0.011	0.002	0.011	0.010
670 nm	**5 bins (HP)**	0.870	0.667	0.300	0.024	-0.008	0.232	0.010
	Models	0.955	1.074	-0.070	0.012	-0.002	0.012	0.010
865 nm	**5 bins (HP)**	0.859	0.585	0.384	0.026	-0.002	0.025	0.010
	Models	0.960	1.171	-0.165	0.014	-0.006	0.013	0.010
1020 nm	**5 bins (HP)**	0.834	0.535	0.434	0.027	-0.001	0.027	0.010
	Models	0.961	1.219	-0.213	0.015	-0.007	0.013	0.010

Table 6.6: Global statistics of PARASOL-like synthetic measurements using GRASP algorithm for BRDF1 in Mongu over land.

Band (nm)	Products	R	Slope	Offset	RMSE	Bias	Std	Grasp Error
443 nm	**5 bins (HP)**	0.937	1.012	-0.001	0.003	-0.0003	0.003	0.002
	Models	0.959	0.932	0.0025	0.002	0.0004	0.002	0.001
490 nm	**5 bins (HP)**	0.943	1.074	-0.004	0.003	0.0004	0.003	0.002
	Models	0.950	0.985	0.0008	0.002	0.00	0.002	0.001
565 nm	**5 bins (HP)**	0.939	1.026	-0.001	0.004	0.0011	0.004	0.003
	Models	0.965	0.982	0.0018	0.003	0.0003	0.003	0.002
670 nm	**5 bins (HP)**	0.983	1.059	-0.005	0.004	0.0013	0.003	0.002
	Models	0.985	0.954	0.0049	0.003	-0.0001	0.003	0.002
865 nm	**5 bins (HP)**	0.959	0.995	0.003	0.006	0.0019	0.006	0.004
	Models	0.955	0.957	0.0114	0.005	0.0001	0.005	0.003
1020 nm	**5 bins (HP)**	0.968	1.025	-0.004	0.007	0.0036	0.007	0.005
	Models	0.947	0.949	0.0161	0.006	0.0003	0.006	0.004

Table 6.7: Global statistics of PARASOL-like synthetic measurements using GRASP algorithm for BRDF1 in Banizoumbou over land.

Band (nm)	Products	R	Slope	Offset	RMSE	Bias	Std	Grasp Error
443 nm	**5 bins (HP)**	0.084	0.140	0.054	0.009	-0.0044	0.0079	0.0040
	Models	0.780	0.767	0.017	0.004	-0.0003	0.0036	0.0028
490 nm	**5 bins (HP)**	0.383	0.543	0.049	0.007	-0.0015	0.0072	0.0041
	Models	0.846	0.846	0.018	0.004	0.0004	0.0036	0.0028
565 nm	**5 bins (HP)**	0.643	1.133	-0.026	0.009	0.0012	0.0095	0.0051
	Models	0.797	0.770	0.047	0.005	-0.0004	0.0054	0.0034
670 nm	**5 bins (HP)**	0.492	0.952	0.021	0.014	0.0051	0.0129	0.0064
	Models	0.846	0.907	0.028	0.006	-0.0022	0.0059	0.0045
865 nm	**5 bins (HP)**	0.742	1.843	-0.363	0.016	0.0061	0.0151	0.0075
	Models	0.621	0.770	0.096	0.009	-0.0030	0.0081	0.0051
1020 nm	**5 bins (HP)**	0.746	1.704	-0.341	0.016	0.0048	0.0157	0.0080
	Models	0.640	0.870	0.058	0.010	-0.0048	0.0091	0.0054

Table 6.8: Global statistics of PARASOL-like synthetic measurements using GRASP algorithm for RRI for both Mongu and Banizoumbou stations over land. The retrieval configuration is provided for 5 bins (HP) configuration.

Band (nm)	Site	R	Slope	Offset	RMSE	Bias	Std	Grasp Error
443 nm	**Banizoumbou**	0.514	0.702	0.459	0.064	0.039	0.049	0.029
	Mongu	0.315	0.259	1.082	0.054	0.006	0.053	0.024
490 nm	**Banizoumbou**	0.506	0.696	0.467	0.064	0.039	0.050	0.028
	Mongu	0.349	0.280	1.053	0.053	0.006	0.052	0.023
565 nm	**Banizoumbou**	0.527	0.713	0.443	0.064	0.039	0.049	0.030
	Mongu	0.374	0.317	0.998	0.053	0.005	0.052	0.025
670 nm	**Banizoumbou**	0.510	0.710	0.449	0.065	0.041	0.051	0.031
	Mongu	0.331	0.284	1.048	0.056	0.006	0.056	0.027
865 nm	**Banizoumbou**	0.436	0.678	0.497	0.072	0.042	0.059	0.036
	Mongu	0.379	0.336	0.981	0.060	0.014	0.058	0.032
1020 nm	**Banizoumbou**	0.447	0.679	0.495	0.071	0.042	0.058	0.040
	Mongu	0.357	0.312	1.017	0.061	0.015	0.059	0.036

6.2.1 Illustration and description of the correlation matrix

As was already mentioned in the Chapter 3, correlation matrix provides additional information about the correlations of the retrieved parameters. Therefore, this section describes some details about the correlation matrix for GRASP/POLDER retrieval. Particularly, only one example is described: the case of GRASP/HP configuration over Banizoumbou. More studies are expected in a future work in order to see the different tendencies in each configuration.

Figure 6.11 shows the correlation matrices for an example of GRASP/POLDER retrieval for one pixel. In the example are retrieved aerosol and BRDF and BPDF together. The first 5 parameters (5x5) represent the concentration values. It is followed by two blocks of 6x6. These blocks are related to the RRI and IRI for 6 wavelengths ($440\,nm$ to $1020\,nm$). Then, the next parameter is the sphericity fraction (1 x 1). It is followed by BRDF_{iso}, BRDF_{vol} and BRDF_{geom} which are provided in 3 blocks each of them of 6x6, correspondingly at 6 wavelengths. Finally, the last block represents the BPDF at 6 wavelengths (6x6). The colors represent the degree of correlation between each property, where the red color is related to positive correlations, the white color represents uncorrelated properties and the blue color is for negative correlated parameters (or anticorrelated). Values close to 1 or -1 mean stronger correlations between the properties.

Some of the more obvious tendencies are outlined below which are also consistent with the already mentioned in Section 4.4.3.1. Strong negative correlations can be seen between C_1 to C_3 with the RRI, which are more pronounced at short wavelengths. This is an expected tendency since C_1 to C_3 represent the concentrations related to the fine mode. In addition, the concentration related to the coarse mode (C_5) exhibits strong negative correlations with the IRI at all wavelengths.

On the other hand, it can be mentioned that strong correlations between RRI and IRI are not observed. RRI for all the wavelengths presents strong negative correlations with the C_{sph}. Moreover, the C_{sph} presents strong positive correlations with the three first concentration parameters (C_1, C_2 and C_3) and very low correlations with C_4 and C_5 concentrations.

For the surface parameters, it is possible to see some moderate positive correlations between the BRDF_{iso} and BRDF_{geom}.

An interesting feature is that there are not too strong correlations between aerosol and surface parameters. This suggests in the GRASP/POLDER approach, aerosol and surface can be well distinguished.

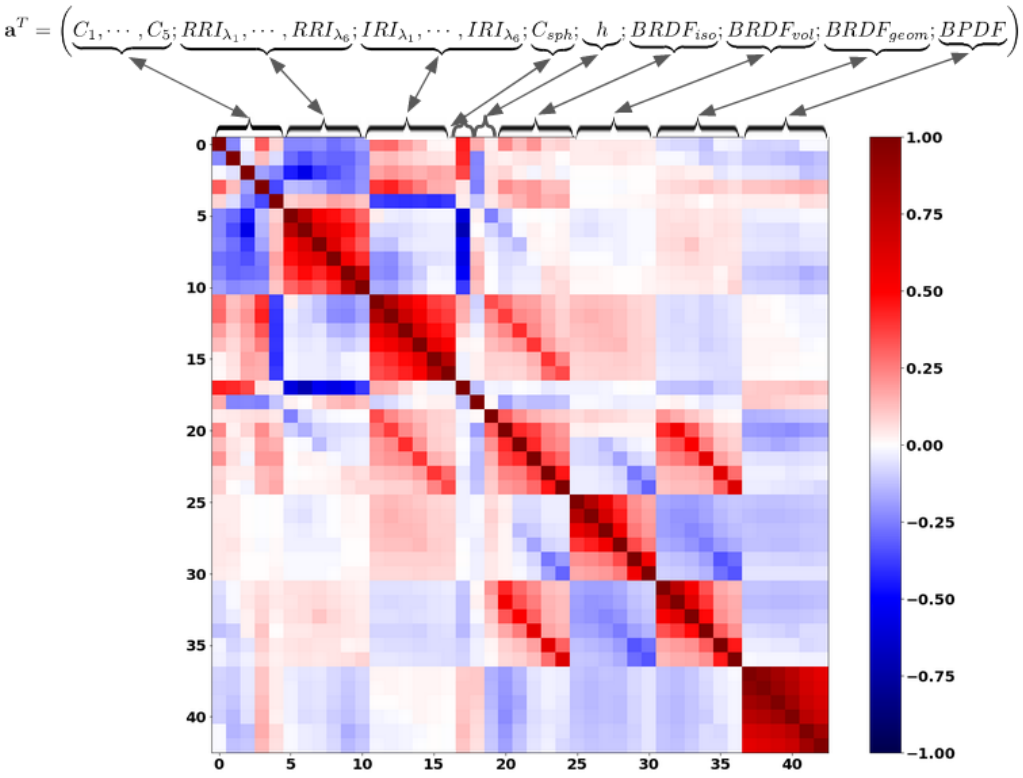


Figure 6.11: Correlation matrix of aerosol and surface properties for POLDER/PARASOL simulated data under noisy conditions using GRASP algorithm. The values close to 1 or -1 mean stronger correlations between the properties, positive or negative, respectively.

6.3 Applications in real POLDER/PARASOL observations

In order to verify the performance of GRASP error estimates in real applications, the set of some aerosol and surface parameters including AOD, SSA and BRDF were evaluated for the full 2008 year of POLDER/PARASOL observations. The illustrations of the scatter plots comparing the POLDER/GRASP retrievals with the AERONET observations are provided for AOD at Banizoumbou and Mongu. It is important to note that total errors are provided by Eqs. 3.52 and 3.53 as was described in Section 6.2. The scatter plots for SSA are not provided since very few coincident points were obtained at the Banizoumbou station and no coincident points were found at the Mongu station. This is because the SSA for almucantar geometry corresponding to the Version 2 Level 2 of AERONET is provided when AOD(440)$>$ 0.4 and an SZA$>$ 50 degrees.

GRASP was applied to the 3x3 pixel around the AERONET locations sites and for these retrievals was used the multi-pixel approach for different configurations: Optimized, High Precision, and Models. These configurations were described in Section 6.1.1. PARASOL/GRASP retrievals are available and validated at six wavelengths

(443, 490, 565, 670, 865 and 1020 nm) and in order to select the more reliable retrievals, it was filtered using the residual relative, at 5% as is suggested by general users, for PARASOL/GRASP products.

In addition, for this study were used the requirements proposed by the Global Climate Observing System (GCOS). These have been adopted in the Aerosol_cci study (Popp et al. (2016)) as well as the DB validation by Sayer et al. (2019) and POLDER/GRASP validation by Chen et al. (2020). Following the Aerosol_cci study by Popp et al. (2016), the uncertainty of 0.01 for AERONET AOD has been taken into account and GCOS is defined as

$$GCOS = \max(0.04 \quad \text{or} \quad 0.1\text{AOD}) \tag{6.3}$$

where, the GCOS fraction (%) is the percentage of retrieved AOD by satellite that satisfies the GCOS requirement. Furthermore, in order to quantify the results, the standard statistical parameters, the linear correlation coefficient (R), the root-mean-square error (RMSE), the slope and the offset of linear regression and the bias, are provided in the illustrations.

Figures 6.12 to 6.17 show the scatter plots for AOD of co-located PARASOL/GRASP against AERONET at different wavelengths (440, 670 and 1020 nm) for GRASP/Optimized, GRASP/HP and GRASP/Models respectively. The retrievals are provided over Banizoumbou and Mongu after filtering retrievals where the fitting residual is higher than 5%. The gray envelope indicates the described GCOS requirement. For each GRASP retrieved AOD was provided the error estimates which are represented by black error bars in the y-axis.

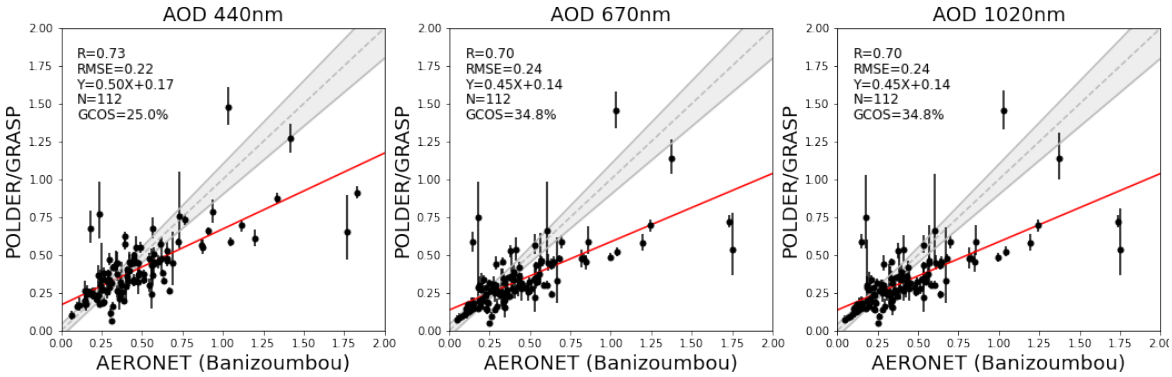


Figure 6.12: Evaluation of 2008 year of PARASOL/GRASP AOD at 440 nm, 670 nm and 1020 nm against AERONET over Banizoumbou using GRASP/Optimized configuration. The gray dashed line and the red solid line are the 1 : 1 reference line and the linear regression line. The gray envelope indicates GCOS requirement: max (0.04 or 0.1 AOD).

Figures 6.12 to 6.14 show the results for Banizoumbou. Figure 6.12 illustrates the GRASP/Optimized AOD at 440, 670 and 1020 nm. In general, based on the metrics

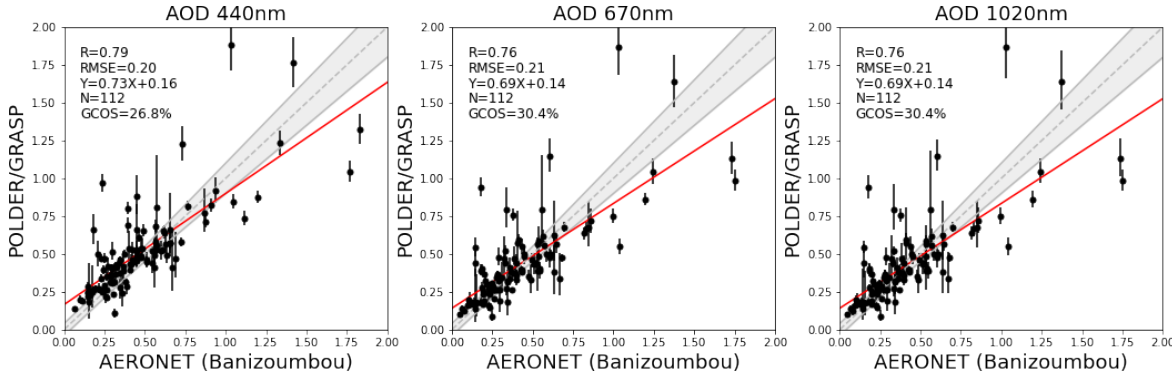


Figure 6.13: Evaluation of 2008 year of PARASOL/GRASP AOD at 440 nm, 670 nm and 1020 nm against AERONET over Banizoumbou using GRASP/HP configuration. The gray dashed line and the red solid line are the 1 : 1 reference line and the linear regression line. The gray envelope indicates GCOS requirement: max (0.04 or 0.1 AOD).

(R, RMSE, Slope, Offset, bias, errors provided by GRASP) the quality of the comparison with AERONET is the best for GRASP/Models. For example, the correlations for AOD at $440\,nm$ is R = 0.89 (GRASP/HP: 0.79 and GRASP/Optimized: 0.73), for AOD at $670\,nm$ is R = 0.89 (GRASP/HP: 0.76 and GRASP/Optimized: 0.70) and for AOD at $1020\,nm$ is R = 0.89 (GRASP/HP: 0.76 and GRASP/Optimized: 0.70). Moreover, RMSE, slope and offset values for GRASP/Models are the better in the comparison with other both configurations. It is possible to see that they follow the same tendencies except some discrepancies in $440\,nm$ for RMSE that show a bigger value than for GRASP/HP and GRASP/Optimized. GCOS fraction of AOD for GRASP/Models at $440\,nm$ is 36.8% while for GRASP/HP and GRASP/Optimized are 26.8% and 25% respectively. The error estimates present bigger values in the case of GRASP/Optimized and GRASP/HP than using GRASP/Models. However, more studies about the error estimates are needed considering some contributions of systematic errors that could be originated from biases in the measurement or some modifications implemented in the algorithm to improve retrieval convergence of non-linear solutions.

Figures 6.15 to 6.17 show the results over Mongu for GRASP/Optimized, GRASP/HP and GRASP/Models respectively. Unlike Banizoumbou in these cases AERONET results at 1020 nm are not available for the year 2008. In general, good agreements can be seen for Mongu station with AERONET. Overall, based on the different evaluated metrics the quality of the comparison with AERONET shows slight differences over Mongu for the three retrieval configurations. However, better results can be seen for GRASP/HP. The correlation values for AOD at $440\,nm$ are quite similar for all the configurations, for example, R= 0.99 in the case of GRASP/HP and GRASP/Optimized while R= 0.98 for GRASP/Models. Some differences can be seen

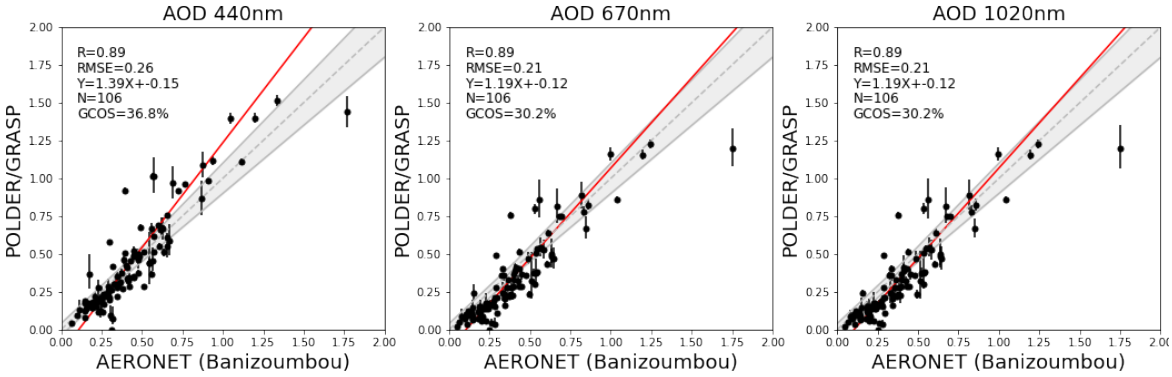


Figure 6.14: Evaluation of 2008 year of PARASOL/GRASP AOD at 440 nm, 670 nm and 1020 nm against AERONET over Banizoumbou using GRASP/Models configuration. The gray dashed line and the red solid line are the 1 : 1 reference line and the linear regression line. The gray envelope indicates GCOS requirement: max (0.04 or 0.1 AOD).

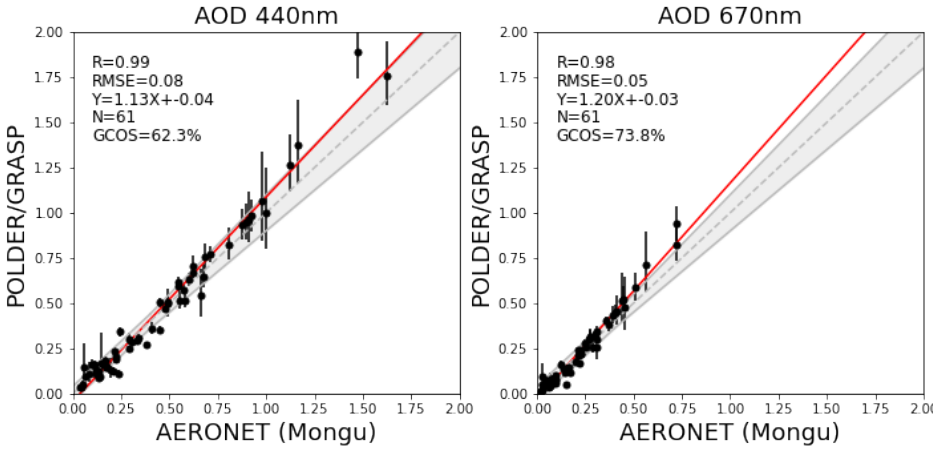


Figure 6.15: Evaluation of 2008 year of PARASOL/GRASP AOD at 440 nm and 670 nm against AERONET over Mongu using GRASP/Optimized configuration. The gray dashed line and the red solid line are the 1 : 1 reference line and the linear regression line. The gray envelope indicates GCOS requirement: max (0.04 or 0.1 AOD).

from RMSE, slope and offset. For example, RMSE for GRASP/Optimized is 0.08, for GRASP/HP 0.11 and GRASP/Models 0.08. The slope and offset for GRASP/Models are = 0.98 and = −0.01 respectively, while for GRASP/HP are 1.19 and −0.03 and for GRASP/Optimized the slope and offset values are 1.13 and −0.04.

GCOS fraction shows bigger values for GRASP/Optimized: 62.3% while values for GRASP/HP is 50% and GRASP/Models is 58.9%.

Figures 6.18 and 6.19 show the error estimates (relative and absolute) provided by GRASP for AOD at four different wavelengths: 443 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red); over Banizoumbou and Mongu for the three different configurations (GRASP/Optimized, GRASP/HP and GRASP/Models). As it is possible to see all the wavelengths show a dependency with the AOD, error estimates

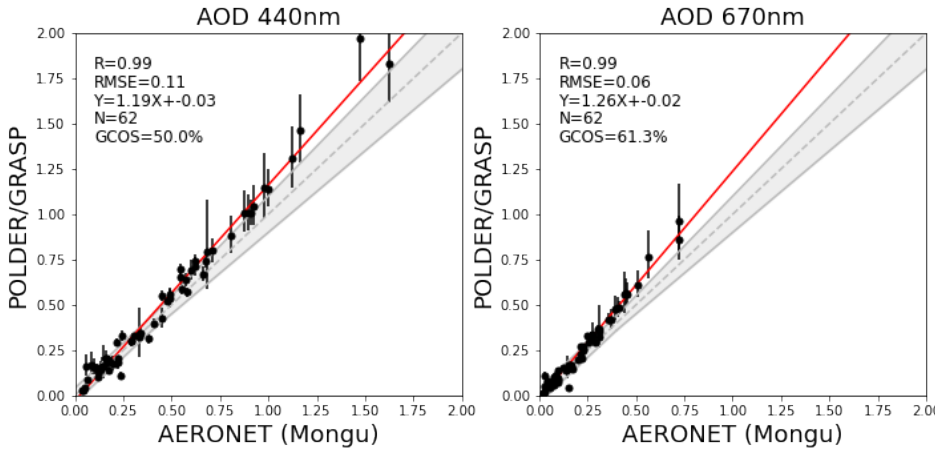


Figure 6.16: Evaluation of 2008 year of PARASOL/GRASP AOD at 440 nm and 670 nm against AERONET over Mongu using GRASP/HP configuration. The gray dashed line and the red solid line are the $1:1$ reference line and the linear regression line. The gray envelope indicates GCOS requirement: max (0.04 or 0.1 AOD).

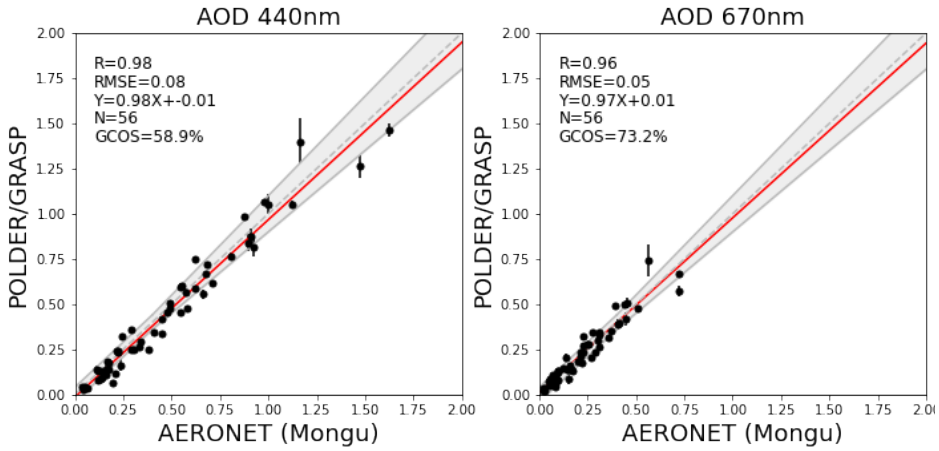


Figure 6.17: Evaluation of 2008 year of PARASOL/GRASP AOD at 440 nm and 670 nm against AERONET over Mongu using GRASP/Models configuration. The gray dashed line and the red solid line are the $1:1$ reference line and the linear regression line. The gray envelope indicates GCOS requirement: max (0.04 or 0.1 AOD).

increase for lower values of AOD. Some differences can be seen between error estimates for the different configurations. For example, more dispersion of errors can be seen in the case of GRASP/Optimized while the lower dispersions are present in the case of GRASP/Models.

In addition, GRASP/Models configuration shows the lower AOD errors provided by GRASP. For example, in Banizoumbou the AOD errors at $443\,nm$ are lower than 40% using GRASP/Models while for GRASP/Optimized and GRASP/HP they are lower than 60%. On the other hand, AOD errors of 15% can be observed at $443\,nm$ in Mongu using GRASP/Models, while errors of $\sim$ 50% are observed using GRASP/Optimized and GRASP/HP.

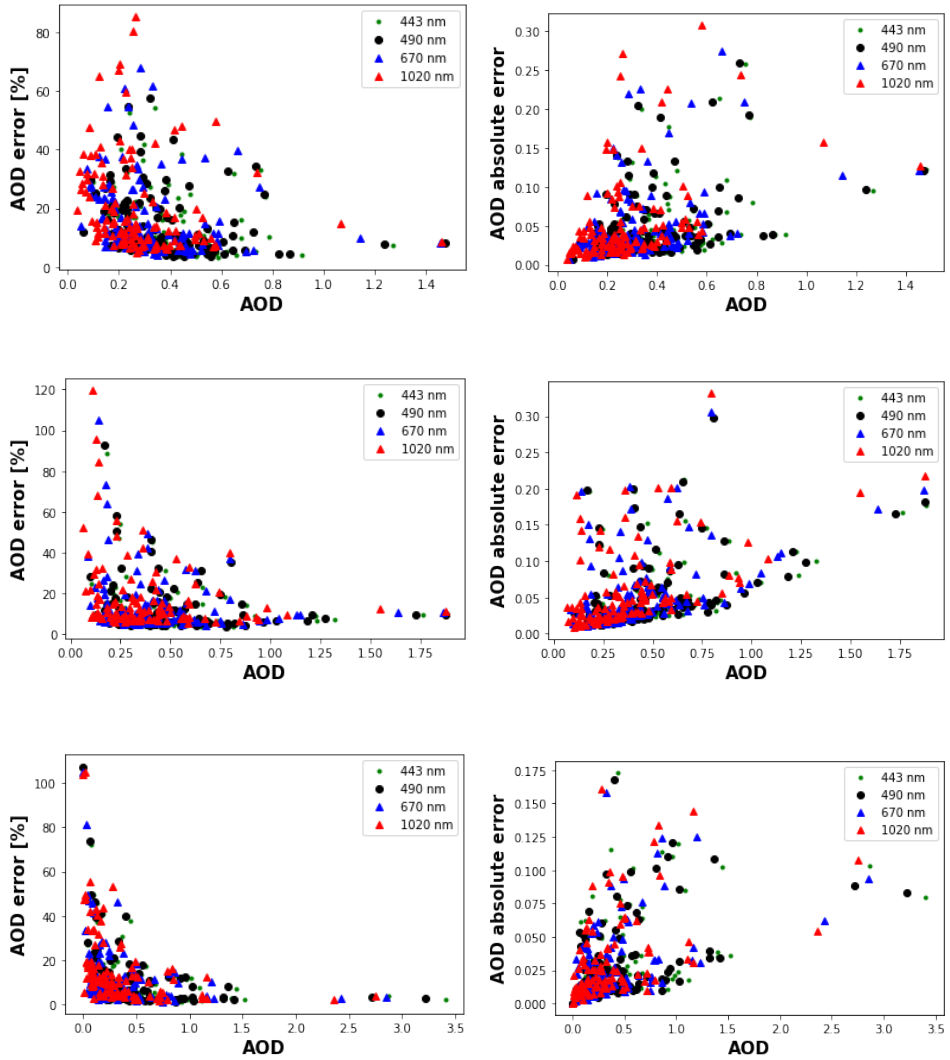


Figure 6.18: Scatter plot of 2008 year of PARASOL/GRASP AOD at 440 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red) vs their error estimates over Banizoumbou using the three different GRASP configurations: GRASP/Optimized, GRASP/HP and GRASP/Models configuration. a) relative error, b) absolute error.

6.4 Conclusions and perspectives

This chapter has presented the preliminary results for PARASOL-like retrievals over Banizoumbou and Mongu sites. The results attempt to show the performance of the GRASP error estimates over satellite retrievals for a series of numerical tests and also in real POLDER/PARASOL observations.

In the first part of this chapter, a series of tests were designed over Mongu and Banizoumbou sites. For these particular studies the synthetic measurements were perturbed by random noise and then inverted.

The PARASOL-like observations were inverted using both: the single-pixel and

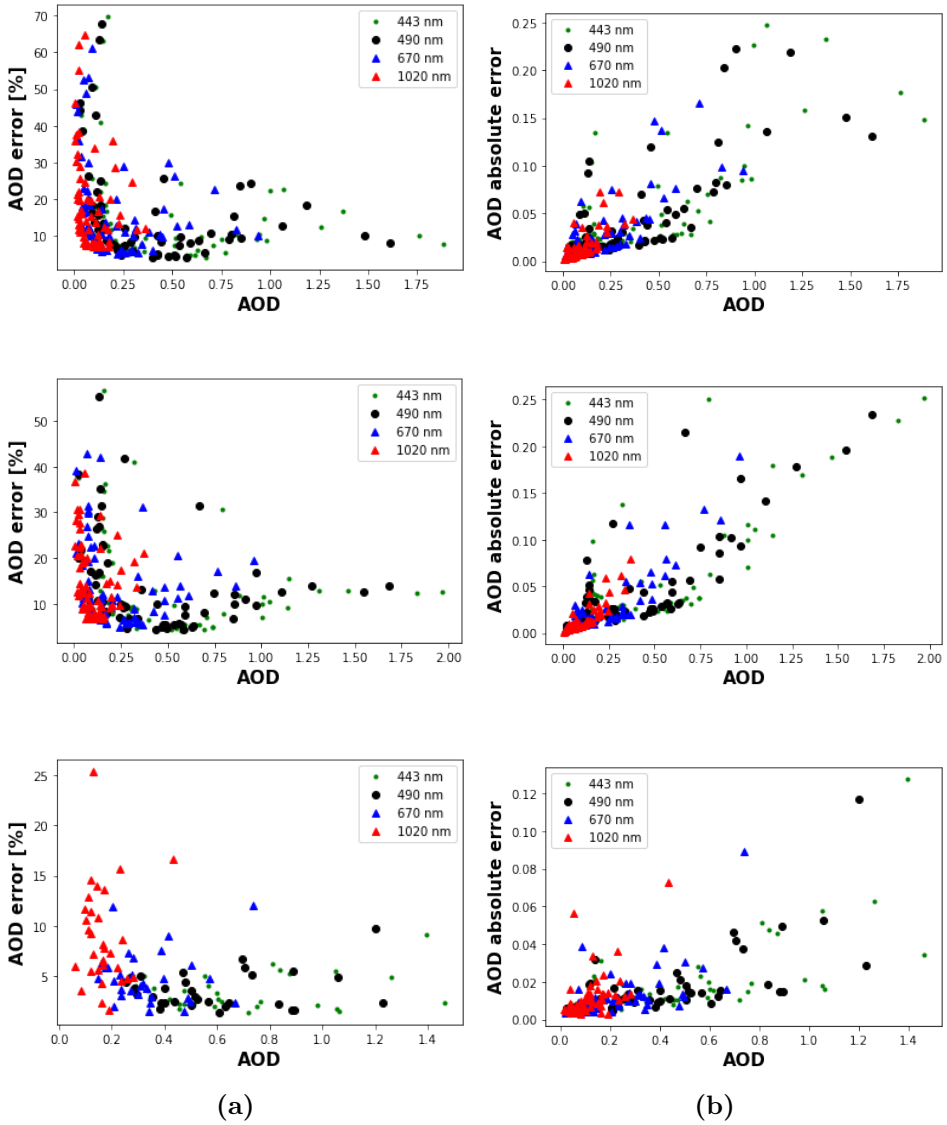


Figure 6.19: Scatter plot of 2008 year of PARASOL/GRASP AOD at 440 nm (green), 490 nm (black), 670 nm (blue) and 1020 nm (red) vs their error estimates over Mongu using the three different GRASP configurations: GRASP/Optimized, GRASP/HP and GRASP/Models configuration. a) relative error, b) absolute error.

multi-pixel approach showing the best results in the case of multi-pixel approach. Two retrieval configurations were used: GRASP/HP and GRASP/Models. The preliminary results showed good performance using both approaches and slight differences between them. Moreover, the results show high correlations for most of the illustrated properties and present good agreements with the error estimates since the errors provided by GRASP are comparable with the actual errors. In addition, a description of the correlation matrix for a particular example is also provided, this correlation matrix allows to see different tendencies. Thus, an interesting feature is that there are not too strong correlations between aerosol and surface parameters. This suggests in the

GRASP/POLDER approach, aerosol and surface can be well distinguished.

This chapter also shows the performance of the algorithm in real applications. Thus, GRASP was also applied to the full 2008 year of POLDER/PARASOL observations over Mongu and Banizoumbou AERONET sites. Comparisons against AERONET and GRASP retrievals have shown consistency of the retrievals for the 2008 year of POLDER/PARASOL observations.

This study is a first step in the analysis of the error estimates for satellite applications. Future work will be done to provide a complete study of the error estimates assuming both, random and systematic errors as it was made for the ground-based observations.

Chapter 7

General Conclusions and Perspectives

This study was devoted to one of the most challenging and largely unaddressed aspect of providing dynamic error estimates of retrieved parameters, adequately and rigorously calculated, from remote sensing observations.

This thesis presented and discussed an approach realized in the GRASP algorithm to estimate errors of the retrieved parameters from remote sensing observations. The employed approach relies on rigorous consideration of error propagation based on the concept of statistical estimations and tends to account for the propagation of both random and systematic errors. For the case of aerosol retrieval from ground-based observations the performance of the GRASP error estimates was extensively evaluated. AERONET-like retrievals from observations by Sun/sky-scanning radiometer and GRASP synergy aerosol retrieval from joint observations by radiometer and lidar were specifically considered. GRASP generates the full covariance matrices that are expected to be used to generate error bars for retrieved parameters, from the diagonal elements, and to provide an interesting insight for understanding retrieval tendencies from the non-diagonal elements. Therefore, the reliability of the obtained covariance diagonal elements was evaluated quantitatively, while the structure of correlation coefficient of covariance matrices was mostly discussed qualitatively.

Chapter 3 provides detailed description of the overall concept and specific key implementations of the error estimates in GRASP. This approach generates full covariance matrix that includes both random and systematic components. The formulation used to estimate errors account for some contribution of systematic errors that could be originated from biases in the measurement or some modifications implemented in the algorithm to improve retrieval convergence of non-linear solutions.

This approach provides extensive information about retrieval errors. The diagonal

elements of the covariance matrix are used for the error bars of the retrieved and derived parameters. Moreover, as it was shown, the errors can be estimated for any characteristics that are not directly retrieved but are a function of retrieved parameters, using the full covariance matrix of retrieved parameters. Moreover, this thesis has demonstrated the importance of non-diagonal elements analysis which is very useful to identify unobvious retrieval tendencies and that can be used to improve the observation schemes and/or retrieval setups.

Thus, Chapter 4 presents the evaluation and validation of the methodology developed to generate the error estimates. Indeed, the practical evaluation of developed error formalism and the possible tuning are necessary for a comprehensive evaluation of the approach and for gaining full confidence in its practical efficiency. Two practical situations of the aerosol retrieval were considered: i) using the measurements by ground-based sun/sky-scanning radiometer alone; and ii) using the combined observations by sun/sky-scanning radiometer and multi-wavelength lidar. Two GRASP retrieval setups were tested: i) when the retrieval assumes that aerosol is composed by homogeneous particles and the parameters of only one aerosol component are retrieved and ii) when the aerosol is assumed as external mixture of two aerosol components and the parameters of each component are retrieved separately. The analysis was performed for synthetic observations for three different aerosol types, as well as, for the mixture of them. Observations of desert Dust were modeled using AERONET retrieval climatology at Solar Village (Saudi Arabia) site. The AERONET retrieval climatologies from African savanna (Zambia) and the GSFC (Maryland, USA) were used simulating Urban and BB aerosol observations respectively.

First, the effects of random errors were evaluated. For that purpose, the synthetic proxy observations perturbed by 300 random noise generated realizations were inverted using the GRASP algorithm. Then, retrieved parameters were compared with those used for synthetic data generation, while the obtained error estimates were compared with actual deviations of the retrieved parameters from assumed values. The result of the test showed the capability of GRASP to provide dynamic error estimates for all retrieved and derived parameters. They are provided for the fine and coarse mode and also for the total components. The results of the statistical tests with randomly generated noise showed that, in most cases, GRASP error estimates are comparable or exceed the actual errors by 20% to 30% and therefore, they can be safely used to assure uncertainties of actual retrieval products. In addition, the observation of typical error values was summarized for different situations and retrieval scenarios. The study confirmed that the detailed properties of aerosol mixtures can be rather reliably retrieved from a combination of radiometer and lidar data, provided that there is sufficient amount of aerosol load in both aerosol components.

Another important contribution in this work is related to the effect of the systematic errors. In principle, each retrieval methodology assumes that there are no systematic uncertainties in measurements and neither in the used forward model. If any systematic bias is identified, it is corrected in measurements or in their interpretation. However, in practice the systematic uncertainties may remain unidentified and make significant contributions to the retrieval uncertainties. Initially, in the GRASP algorithm, the apparent misfit was used as an indicator of bias, however in real situations not all biases can be seen from the misfit. Chapter 4 presents the results considering a possible solution to this problem. Specifically, the contribution of potential bias was included in the estimation of the retrieval errors. The tests showed that the resulting estimations very well reflect all the qualitative effects of biases on the retrievals. Thus, as a result of this study, an extra term that accounts for propagation of possible bias from the measurements was added in the equation for calculation of the retrieval errors.

Thus, based on the results of the numerical tests, the GRASP approach was adjusted and tuned for generating errors that adequately characterize the retrieval uncertainties in the practice. For example, for a case when aerosol is retrieved from ground-based sun/sky photometer data, two types of potential main biases were considered: in measured AOD and sky-radiances. Both the possible effects of underestimations and overestimations of AOD and sky-radiances were considered and their quadratic means were averaged. For the cases, when lidar data are used, it is suggested to simulate biases in lidar attenuation measurements for each wavelength similarly as discussed in studies by Lopatin et al. (2013, 2021). The finalized error equations were additionally tested. The tests showed that the systematic error over all agrees well with actual manifestations of both random errors and the bias in the measurements. Also, the evaluation results clearly confirm improvements in the error estimates from adding the new term equation accounting for potential effect of the systematic errors.

The observed tendencies in the effects of biases on the retrieval are consistent with all the results already described in other studies. The results of the tests showed that the complete set of aerosol parameters for each aerosol component can be robustly derived with acceptable accuracy in almost all considered situations. The obtained results are expected to be representative for most practical situations, while some additional tests and analysis can be certainly useful.

It should be noted that even such a challenging approach as retrieval mixed bi-component aerosol was considered. Namely, the retrieval of bi-component aerosol was evaluated using radiometer only simulated measurements and then by adding lidar observations. These tests allowed us to observe that by using the synergy of two instruments, some improvements were observed in the aerosol properties retrievals of each aerosol mixture component and in the estimations of the retrieval errors. The test for selected cases with presence of different aerosol components ($\tau_f = 0.2$, $\tau_c = 0.8$

and $\tau_f = 0.8$, $\tau_c = 0.2$) showed that optical properties of the dominant mode can be retrieved significantly more accurately as can be expected. It is interesting to note that in all situations using only radiometer data or adding lidar simulated measurements, such properties as total SSA and effective refractive index can be retrieved rather accurately, even in cases where the retrieval of properties of each mode separately is questionable.

Chapter 5 has demonstrated the application of GRASP for aerosol retrieval from the synergy between lidar and solar photometer in Buenos Aires (Argentina). The GRASP retrievals and the error estimates of the columnar aerosol properties showed to be fully adequate in comparative analysis with the aerosol products available from AERONET operational retrievals. Furthermore, GRASP has demonstrated the capability to provide dynamic error estimates for retrieved parameters in all the different configurations and also for fine and coarse mode. It can be considered as notable advantage of GRASP since AERONET provides the errors of some retrieved aerosol parameter for the Level 2 (particularly, for SSA and IRI when AOD at $440\,nm$ is greater than 0.4 and SZA$>$ 50 degrees).

This approach to modeling error estimates for the derived and retrieved parameters enables its application in different configurations and with several measurements or combinations of them. In other words, this approach is not only limited to the ground-based observations presented in this work but it could also be used with any other possible combinations of existing remote sensing instruments, such as satellite observations.

Finally, the last Chapter outlines the full potential of developed error estimation methodology in satellite observations. Thus, the chapter provides preliminary illustration of the generated errors for aerosol retrievals from satellite data and describes the perspectives of the methodology evolution. For example, the evaluation of the methodology in using more measurement scenarios and considering additional situations with different aerosol mixtures and abundances. Also, while the results of the first illustration of error estimates in satellite retrievals are promising, additional extensive analysis is highly desirable, especially for complex satellite retrievals based on multi-pixel retrieval concept (Dubovik et al. (2011)).

The main perspectives of this work are focus in extend the formalism for more ground-based retrieval applications. For example, analyze the error tendency with addition of polarization measurements. In addition, is expected to extend the error estimation approach for practical characterization of satellite retrieval uncertainties. Finally, more comprehensive studies are expected in the consideration of additional error factors, e.g. biases in forward model and diverse assumptions.

Bibliography

Ackermann, J. The extinction-to-backscatter ratio of tropospheric aerosol: A numerical study. *Journal of atmospheric and oceanic technology*, 15(4):1043–1050, 1998.

Albrecht, B. A. Aerosols, cloud microphysics, and fractional cloudiness. *Science*, 245 (4923):1227–1230, 1989.

Antuña-Marrero, J. C., Landulfo, E., Estevan, R., Barja, B., Robock, A., Wolfram, E., Ristori, P., Clemesha, B., Zaratti, F., Forno, R., et al. Lalinet: The first latin american–born regional atmospheric observational network. *Bulletin of the American Meteorological Society*, 98(6):1255–1275, 2017.

Barnaba, F. and Gobbi, G. P. Lidar estimation of tropospheric aerosol extinction, surface area and volume: Maritime and desert-dust cases. *Journal of Geophysical Research: Atmospheres*, 106(D3):3005–3018, 2001.

Barnaba, F. and Gobbi, G. P. Modeling the aerosol extinction versus backscatter relationship for lidar applications: maritime and continental conditions. *Journal of Atmospheric and Oceanic Technology*, 21(3):428–442, 2004.

Bellouin, N., Quaas, J., Gryspeerdt, E., Kinne, S., Stier, P., Watson-Parris, D., Boucher, O., Carslaw, K. S., Christensen, M., Daniau, A.-L., et al. Bounding global aerosol radiative forcing of climate change. *Reviews of Geophysics*, 58(1): e2019RG000660, 2020.

Benavent-Oltra, J. A., Román, R., Granados-Muñoz, M. J., Pérez-Ramírez, D., Ortiz-Amezcua, P., Denjean, C., Lopatin, A., Lyamani, H., Torres, B., Guerrero-Rascado, J. L., Fuertes, D., Dubovik, O., Chaikovsky, A., Olmo, F. J., Mallet, M., and Alados-Arboledas, L. Comparative assessment of grasp algorithm for a dust event over granada (spain) during charmex-adrimed 2013 campaign. *Atmospheric Measurement Techniques*, 10(11):4439–4457, 2017. doi: 10.5194/amt-10-4439-2017.

Benavent-Oltra, J. A. *Synergetic use of active and passive remote sensing techniques to retrieve vertical profiles of atmospheric aerosol properties during day-and night-time.* PhD thesis, Universidad de Granada, 2019.

Benavent-Oltra, J. A., Román, R., Casquero-Vera, J. A., Pérez-Ramírez, D., Lyamani, H., Ortiz-Amezcua, P., Bedoya-Velásquez, A. E., Arruda Moreira, G. d., Barreto, Á., Lopatin, A., Fuertes, D., Herrera, M., Torres, B., Dubovik, O., Guerrero-Rascado, J. L., Goloub, P., Olmo-Reyes, F. J., and Alados-Arboledas, L. Different strategies to retrieve aerosol properties at night-time with the grasp algorithm. *Atmospheric Chemistry and Physics*, 19(22):14149–14171, 2019. doi: 10.5194/acp-19-14149-2019.

Benavent-Oltra, J. A., Casquero-Vera, J. A., Román, R., Lyamani, H., Pérez-Ramírez, D., Granados-Muñoz, M. J., Herrera, M., Cazorla, A., Titos, G., Ortiz-Amezcua, P., Bedoya-Velásquez, A. E., Arruda Moreira, G., Pérez, N., Alastuey, A., Dubovik, O., Guerrero-Rascado, J., Olmo-Reyes, F. J., and Alados-Arboledas, L. Overview of the slope i and ii campaigns: aerosol properties retrieved with lidar and sun–sky photometer measurements. *Atmospheric Chemistry and Physics*, 21(12):9269–9287, 2021. doi: 10.5194/acp-21-9269-2021.

Boucher, O. Air traffic may increase cirrus cloudiness. *Nature*, 397(6714):30–31, 1999.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., Zhang, X., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. 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*, pages 571–657. Cambridge University Press, 2013.

Bovchaliuk, V., Goloub, P., Podvin, T., Veselovskii, I., Tanre, D., Chaikovsky, A., Dubovik, O., Mortier, A., Lopatin, A., Korenskiy, M., and Victori, S. Comparison of aerosol properties retrieved using garrlic, liric, and raman algorithms applied to multi-wavelength lidar and sun/sky-photometer data. *Atmospheric Measurement Techniques*, 9(7):3391–3405, 2016. doi: 10.5194/amt-9-3391-2016.

Bréon, F.-M., Vermeulen, A., and Descloitres, J. An evaluation of satellite aerosol products against sunphotometer measurements. *Remote sensing of environment*, 115(12):3102–3111, 2011.

Carboni, E. Gome aerosol optical depth retrieval over ocean: correcting for the effects of residual cloud contamination. *Atmospheric Environment*, 40(36):6975–6987, 2006.

Cayuela, M. Á. L., Córdoba-Jabonero, C., Pérez-Ramírez, D., Herrera, M., and Guerrero-Rascado, J. L. Grasp retrievals in synergy with both polarized micro-pulse lidar and sun/sky photometer measurements to derive optical and microphysical

properties of aged smoke plumes. In *Remote Sensing of Clouds and the Atmosphere XXVI*, volume 11859, pages 79–89. SPIE, 2021.

Chaikovsky, A., Dubovik, O., Holben, B., Bril, A., Goloub, P., Tanré, D., Pappalardo, G., Wandinger, U., Chaikovskaya, L., Denisov, S., et al. Lidar-radiometer inversion code (liric) for the retrieval of vertical aerosol properties from combined lidar/radiometer data: development and distribution in earlinet. *Atmospheric Measurement Techniques*, 9(3):1181–1205, 2016. doi: 10.5194/amt-9-1181-2016.

Chen, C., Dubovik, O., Fuertes, D., Litvinov, P., Lapyonok, T., Lopatin, A., Ducos, F., Derimian, Y., Herman, M., Tanré, D., Remer, L. A., Lyapustin, A., Sayer, A. M., Levy, R. C., Hsu, C., Descloitres, J., Li, L., Torres, B., Karol, Y., Herrera, M., Herreras, M., Aspetsberger, M., Wanzenboeck, M., Bindreiter, L., Marth, D., Hangler, A., and Federspiel, C. Validation of grasp algorithm product from polder/parasol data and assessment of multi-angular polarimetry potential for aerosol monitoring. *Earth System Science Data*, 12(4):3573–3620, 2020. doi: 10.5194/essd-12-3573-20202020.

Chu, D., Kaufman, Y., Ichoku, C., Remer, L., Tanré, D., and Holben, B. Validation of modis aerosol optical depth retrieval over land. *Geophysical research letters*, 29(12): MOD2–1, 2002.

Cox, C. and Munk, W. Measurement of the roughness of the sea surface from photographs of the sun's glitter. *Josa*, 44(11):838–850, 1954.

Derimian, Y., Dubovik, O., Huang, X., Lapyonok, T., Litvinov, P., Kostinski, A. B., Dubuisson, P., and Ducos, F. Comprehensive tool for calculation of radiative fluxes: illustration of shortwave aerosol radiative effect sensitivities to the details in aerosol and underlying surface characteristics. *Atmospheric Chemistry and Physics*, 16(9): 5763–5780, 2016.

Deschamps, P.-Y., Bréon, F.-M., Leroy, M., Podaire, A., Bricaud, A., Buriez, J.-C., and Seze, G. The polder mission: Instrument characteristics and scientific objectives. *IEEE Transactions on geoscience and remote sensing*, 32(3):598–615, 1994.

Deshler, T., Johnson, B. J., and Rozier, W. R. Balloonborne measurements of pinatubo aerosol during 1991 and 1992 at 41 n: Vertical profiles, size distribution, and volatility. *Geophysical Research Letters*, 20(14):1435–1438, 1993.

Deuzé, J., Herman, M., Goloub, P., Tanré, D., and Marchand, A. Characterization of aerosols over ocean from polder/adeos-1. *Geophysical Research Letters*, 26(10): 1421–1424, 1999.

Deuzé, J., Bréon, F., Devaux, C., Goloub, P., Herman, M., Lafrance, B., Maignan, F., Marchand, A., Nadal, F., Perry, G., and Tanré, D. Remote sensing of aerosols over land surfaces from polder-adeos-1 polarized measurements. *Journal of Geophysical Research: Atmospheres*, 106(D5):4913–4926, 2001.

Diner, D. J., Beckert, J. C., Reilly, T. H., Bruegge, C. J., Conel, J. E., Kahn, R. A., Martonchik, J. V., Ackerman, T. P., Davies, R., Gerstl, S. A., et al. Multi-angle imaging spectroradiometer (misr) instrument description and experiment overview. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1072–1087, 1998.

Dubovik, O., Smirnov, A., Holben, B., King, M., Kaufman, Y., Eck, T., and Slutsker, I. Accuracy assessments of aerosol optical properties retrieved from aerosol robotic network (aeronet) sun and sky radiance measurements. *Journal of Geophysical Research: Atmospheres*, 105(D8):9791–9806, 2000. doi: 10.1029/2000jd900040.

Dubovik, O., Holben, B., Lapyonok, T., Sinyuk, A., Mishchenko, M., Yang, P., and Slutsker, I. Non-spherical aerosol retrieval method employing light scattering by spheroids. *Geophysical Research Letters*, 29(10):54–1, 2002a.

Dubovik, O., Lapyonok, T., Kaufman, Y., Chin, M., Ginoux, P., Kahn, R., and Sinyuk, A. Retrieving global aerosol sources from satellites using inverse modeling. *Atmospheric Chemistry and Physics*, 8(2):209–250, 2008. doi: 10.5194/acp-8-209-2008.

Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J., Ducos, F., Sinyuk, A., and Lopatin, A. Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. *Atmospheric Measurement Techniques*, 4(5):975–1018, 2011. doi: 10.5194/amt-4-975-2011.

Dubovik, O. Optimization of numerical inversion in photopolarimetric remote sensing. In *Photopolarimetry in remote sensing*, pages 65–106. Springer, 2004.

Dubovik, O. and King, M. D. A flexible inversion algorithm for retrieval of aerosol optical properties from sun and sky radiance measurements. *Journal of Geophysical Research: Atmospheres*, 105(D16):20673–20696, 2000. doi: 10.1029/2000jd900282.

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanré, D., and Slutsker, I. Variability of absorption and optical properties of key aerosol types observed in worldwide locations. *Journal of the atmospheric sciences*, 59(3): 590–608, 2002b. doi: 10.1175/1520-0469(2002)059<0590:voaaop>2.0.co;2.

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F., Volten, H., Munoz, O., Veihelmann, B., Van der Zande, W. J., Leon,

J., 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(D11), 2006. doi: 10.1029/2005JD006619.

Dubovik, O., Lapyonok, T., Litvinov, P., Herman, M., Fuertes, D., Ducos, F., Lopatin, A., Chaikovsky, A., Torres, B., Derimian, Y., et al. Grasp: a versatile algorithm for characterizing the atmosphere. *SPIE Newsroom*, 25(10.1117):2–1201408, 2014. doi: 10.1117/2.1201408.005558.

Dubovik, O., Li, Z., Mishchenko, M. I., Tanre, D., Karol, Y., Bojkov, B., Cairns, B., Diner, D. J., Espinosa, W. R., Goloub, P., et al. Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 224:474–511, 2019. doi: 10.1016/j.jqsrt.2018.11.024.

Dubovik, O., Fuertes, D., Lytvynov, P., Lopatin, A., Lapyonok, T., Doubovik, I., Xu, F., Ducos, F., Chen, C., Torres, B., et al. Multi-term lsm for applying multiple a priori constraints in problems of atmospheric remote sensing: Grasp algorithm-concept and applications. *Frontiers in Remote Sensing*, page 23, 2021. doi: 10.3389/frsen.2021.706851.

Dubovik, O., Lapyonok, T., and Oshchepkov, S. Improved technique for data inversion: optical sizing of multicomponent aerosols. *Applied optics*, 34(36):8422–8436, 1995. doi: 10.1364/ao.34.008422.

Eck, T. F., Holben, B., Reid, J., Dubovik, O., Smirnov, A., O'neill, N., 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(D24): 31333–31349, 1999.

Edie, W. T., Drijard, D., James, F. E., Roos, M., and Sadoulet, B. Statistical methods in experimental physics. *New York North-Holland*, 155, 1971.

Espinosa, W. R., Remer, L. A., Dubovik, O., Ziemba, L., Beyersdorf, A., Orozco, D., Schuster, G., Lapyonok, T., Fuertes, D., and Martins, J. V. Retrievals of aerosol optical and microphysical properties from imaging polar nephelometer scattering measurements. *Atmospheric measurement techniques*, 10(3):811–824, 2017. doi: 10.5194/amt-10-811-2017.

Evans, B. T. Sensitivity of the backscatter/extinction ratio to changes in aerosol properties: implications for lidar. *Applied Optics*, 27(15):3299–3305, 1988.

Fedarenka, A., Dubovik, O., Goloub, P., Li, Z., Lapyonok, T., Litvinov, P., Barel, L., Gonzalez, L., Podvin, T., and Crozel, D. Utilization of aeronet polarimetric measurements for improving retrieval of aerosol microphysics: Gsfc, beijing and dakar data analysis. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 179:72–97, 2016. doi: 10.1016/j.jqsrt.2016.03.021.

Fourgeaud, C. and Fuchs, A. Statistique, dunod. *Paris. Zbl0201*, 51501, 1967.

Giles, D. M., Holben, B. N., Eck, T. F., Sinyuk, A., Smirnov, A., Slutsker, I., Dickerson, R., Thompson, A., and Schafer, J. An analysis of aeronet aerosol absorption properties and classifications representative of aerosol source regions. *Journal of Geophysical Research: Atmospheres*, 117(D17), 2012. doi: 10.1029/2012JD018127.

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., et al. Advancements in the aerosol robotic network (aeronet) version 3 database–automated near-real-time quality control algorithm with improved cloud screening for sun photometer aerosol optical depth (aod) measurements. *Atmospheric Measurement Techniques*, 12(1): 169–209, 2019. doi: 10.5194/amt-12-169-2019.

Goloub, P., Li, Z., Dubovik, O., Blarel, L., Podvin, T., Jankowiak, I., Lecoq, R., Deroo, C., Chatenet, B., Morel, J., et al. Photons/aeronet sunphotometer network overview: description, activities, results. In *Fourteenth International Symposium on Atmospheric and Ocean Optics/Atmospheric Physics*, volume 6936, page 69360V. International Society for Optics and Photonics, 2008.

Gottwald, M., Bovensmann, H., Lichtenberg, G., Noel, S., von Bargen, A., Slijkhuis, S., Piter, A., Hoogeveen, R., von Savigny, C., Buchwitz, M., et al. Sciamachy, monitoring the changing earth's atmosphere. 2006.

Govaerts, Y., Wagner, S., Lattanzio, A., and Watts, P. Joint retrieval of surface reflectance and aerosol optical depth from msg/seviri observations with an optimal estimation approach: 1. theory. *Journal of Geophysical Research: Atmospheres*, 115 (D2), 2010.

Grey, W. M., North, P. R., Los, S. O., and Mitchell, R. M. Aerosol optical depth and land surface reflectance from multiangle aatsr measurements: global validation and intersensor comparisons. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (8):2184–2197, 2006.

Guerrero-Rascado, J. L., Landulfo, E., Antuña, J. C., Barbosa, H. d. M. J., Barja, B., Bastidas, Á. E., Bedoya, A. E., da Costa, R. F., Estevan, R., Forno, R., et al.

Latin american lidar network (lalinet) for aerosol research: Diagnosis on network instrumentation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 138:112–120, 2016.

Hansen, J. E. Multiple scattering of polarized light in planetary atmospheres. part i. the doubling method. *Journal of Atmospheric Sciences*, 28(1):120–125, 1971.

Hasekamp, O. P. and Landgraf, J. Retrieval of aerosol properties over land surfaces: capabilities of multiple-viewing-angle intensity and polarization measurements. *Applied optics*, 46(16):3332–3344, 2007.

Haywood, J. and Boucher, O. Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review. *Reviews of geophysics*, 38(4):513–543, 2000.

Herreras, M., Román, R., Cazorla, A., Toledano, C., Lyamani, H., Torres, B., Cachorro, V., Olmo, F., Alados-Arboledas, L., and de Frutos, A. Evaluation of retrieved aerosol extinction profiles using as reference the aerosol optical depth differences between various heights. *Atmospheric Research*, 230:104625, 2019. doi: 10.1016/j.atmosres.2019.104625.

Higurashi, A., Nakajima, T., Holben, B. N., Smirnov, A., Frouin, R., and Chatenet, B. A study of global aerosol optical climatology with two-channel avhrr remote sensing. *Journal of climate*, 13(12):2011–2027, 2000.

Holben, B. N., Eck, T. F., Slutsker, I. a., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y., Nakajima, T., et al. Aeronet - a federated instrument network and data archive for aerosol characterization. *Remote sensing of environment*, 66(1):1–16, 1998. doi: 10.1016/s0034-4257(98)00031-5.

Holben, B. N., Eck, T., Slutsker, I., Smirnov, A., Sinyuk, A., Schafer, J., Giles, D., and Dubovik, O. Aeronet's version 2.0 quality assurance criteria. In *Remote Sensing of the Atmosphere and Clouds*, volume 6408, page 64080Q. International Society for Optics and Photonics, 2006.

Hsu, N. C., Herman, J. R., and Weaver, C. Determination of radiative forcing of saharan dust using combined toms and erbe data. *Journal of Geophysical Research: Atmospheres*, 105(D16):20649–20661, 2000.

Hsu, N. C., Tsay, S.-C., King, M. D., and Herman, J. R. Deep blue retrievals of asian aerosol properties during ace-asia. *IEEE transactions on geoscience and remote sensing*, 44(11):3180–3195, 2006.

Hu, Q., Goloub, P., Veselovskii, I., Bravo-Aranda, J.-A., Popovici, I. E., Podvin, T., Haeffelin, M., Lopatin, A., Dubovik, O., Pietras, C., et al. Long-range-transported canadian smoke plumes in the lower stratosphere over northern france. *Atmospheric Chemistry and Physics*, 19(2):1173–1193, 2019. doi: 10.5194/acp-19-1173-2019.

Ignatov, A. and Stowe, L. Physical basis, premises, and self-consistency checks of aerosol retrievals from trmm virs. *Journal of Applied Meteorology and Climatology*, 39(12):2259–2277, 2000.

Jäger, H. and Hofmann, D. Midlatitude lidar backscatter to mass, area, and extinction conversion model based on in situ aerosol measurements from 1980 to 1987. *Applied optics*, 30(1):127–138, 1991.

Junge, C. The size distribution and aging of natural aerosols as determined from electrical and optical data on the atmosphere. *Journal of Atmospheric Sciences*, 12 (1):13–25, 1955.

Justice, C. O., Vermote, E., Townshend, J. R., Defries, R., Roy, D. P., Hall, D. K., Salomonson, V. V., Privette, J. L., Riggs, G., Strahler, A., et al. The moderate resolution imaging spectroradiometer (modis): Land remote sensing for global change research. *IEEE transactions on geoscience and remote sensing*, 36(4):1228–1249, 1998.

Kahn, R. A., Gaitley, B. J., Martonchik, J. V., Diner, D. J., Crean, K. A., and Holben, B. Multiangle imaging spectroradiometer (misr) global aerosol optical depth validation based on 2 years of coincident aerosol robotic network (aeronet) observations. *Journal of Geophysical Research: Atmospheres*, 110(D10), 2005.

Kasten, F. and Young, A. T. Revised optical air mass tables and approximation formula. *Applied optics*, 28(22):4735–4738, 1989.

Kaufman, Y. J., Tanre, D., Holben, B., Mattoo, S., Remer, L., Eck, T., Vaughan, J., and Chatenet, B. Aerosol radiative impact on spectral solar flux at the surface, derived from principal-plane sky measurements. *Journal of the atmospheric sciences*, 59(3):635–646, 2002.

Kienle, T., Dean, K., Garbeil, H., and Rose, W. Satellite surveillance of volcanic ash plumes, application to aircraft safety. *Eos, Transactions American Geophysical Union*, 71(7):266–266, 1990.

King, M. D., Kaufman, Y. J., Tanré, D., and Nakajima, T. Remote sensing of tropospheric aerosols from space: Past, present, and future. *Bulletin of the American Meteorological society*, 80(11):2229–2260, 1999.

King, M. D., Menzel, W. P., Kaufman, Y. J., Tanré, D., Gao, B.-C., Platnick, S., Ackerman, S. A., Remer, L. A., Pincus, R., and Hubanks, P. A. Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from modis. *IEEE Transactions on Geoscience and Remote Sensing*, 41(2):442–458, 2003.

Kokhanovsky, A., Davis, A., Cairns, B., Dubovik, O., Hasekamp, O., Sano, I., Mukai, S., Rozanov, V., Litvinov, P., Lapyonok, T., et al. Space-based remote sensing of atmospheric aerosols: The multi-angle spectro-polarimetric frontier. *Earth-Science Reviews*, 145:85–116, 2015.

Kokhanovsky, A. A. and Breon, F.-M. Validation of an analytical snow brdf model using parasol multi-angular and multispectral observations. *IEEE Geoscience and Remote Sensing Letters*, 9(5):928–932, 2012.

Kokhanovsky, A. A. and Leeuw, G. *Satellite aerosol remote sensing over land*, volume 111. Springer, 2009.

Kokhanovsky, A. A. and Zege, E. P. Scattering optics of snow. *Applied Optics*, 43(7): 1589–1602, 2004.

Kovalev, V. A. and Eichinger, W. E. *Elastic lidar: theory, practice, and analysis methods*. John Wiley & Sons, 2004.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 525(7569):367–371, 2015.

Lenoble, J., Herman, M., Deuzé, J., Lafrance, B., Santer, R., and Tanré, D. A successive order of scattering code for solving the vector equation of transfer in the earth's atmosphere with aerosols. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 107(3):479–507, 2007.

Lenoble, J., Remer, L., and Tanre, D. *Aerosol remote sensing*. Springer Science & Business Media, 2013.

Levy, R., Remer, L., Kleidman, R., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. Global evaluation of the collection 5 modis dark-target aerosol products over land. *Atmospheric Chemistry and Physics*, 10(21):10399–10420, 2010.

Li, J., Carlson, B. E., and Lacis, A. A. A study on the temporal and spatial variability of absorbing aerosols using total ozone mapping spectrometer and ozone monitoring instrument aerosol index data. *Journal of Geophysical Research: Atmospheres*, 114 (D9), 2009a.

Li, L., Dubovik, O., Derimian, Y., Schuster, G. L., Lapyonok, T., Litvinov, P., Ducos, F., Fuertes, D., Chen, C., Li, Z., et al. Retrieval of aerosol components directly from satellite and ground-based measurements. *Atmospheric Chemistry and Physics*, 19 (21):13409–13443, 2019. doi: 10.5194/acp-19-13409-2019.

Li, R.-R., Remer, L., Kaufman, Y. J., Mattoo, S., Gao, B.-C., and Vermote, E. Snow and ice mask for the modis aerosol products. *Ieee Geoscience and Remote Sensing Letters*, 2(3):306–310, 2005.

Li, X., Strahler, A., et al. Geometric-optical bidirectional reflectance modeling of mutual shadowing effects of crowns in a forest canopy. *IEEE Trans. Geosci. Remote Sens*, 30:276–292, 1992.

Li, Z., Goloub, P., Dubovik, O., Blarel, L., Zhang, W., Podvin, T., Sinyuk, A., Sorokin, M., Chen, H., Holben, B., et al. Improvements for ground-based remote sensing of atmospheric aerosol properties by additional polarimetric measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(17):1954–1961, 2009b. doi: 10.1016/j.jqsrt.2009.04.009.

Litvinov, P., Hasekamp, O., and Cairns, B. Models for surface reflection of radiance and polarized radiance: Comparison with airborne multi-angle photopolarimetric measurements and implications for modeling top-of-atmosphere measurements. *Remote Sensing of Environment*, 115(2):781–792, 2011a.

Litvinov, P., Hasekamp, O., Cairns, B., and Mishchenko, M. Semi-empirical brdf and bpdf models applied to the problem of aerosol retrievals over land: testing on airborne data and implications for modeling of top-of-atmosphere measurements. In *Polarimetric Detection, Characterization and Remote Sensing*, pages 313–340. Springer, 2011b.

Litvinov, P., Hasekamp, O., Dubovik, O., and Cairns, B. Model for land surface reflectance treatment: Physical derivation, application for bare soil and evaluation on airborne and satellite measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 113(16):2023–2039, 2012.

Lopatin, A., Dubovik, O., Chaikovsky, A., Goloub, P., Lapyonok, T., Tanré, D., and Litvinov, P. Enhancement of aerosol characterization using synergy of lidar and sun-photometer coincident observations: the garrlic algorithm. *Atmospheric Measurement Techniques*, 6(8):2065–2088, 2013. doi: 10.5194/amt-6-2065-2013.

Lopatin, A. *Enhanced remote sensing of atmospheric aerosol by joint inversion of active and passive remote sensing observations*. PhD thesis, Lille 1, 2013.

Lopatin, A., Dubovik, O., Fuertes, D., Stenchikov, G., Lapyonok, T., Veselovskii, I., Wienhold, F. G., Shevchenko, I., Hu, Q., and Parajuli, S. Synergy processing of diverse ground-based remote sensing and in situ data using the grasp algorithm: applications to radiometer, lidar and radiosonde observations. *Atmospheric Measurement Techniques*, 14(3):2575–2614, 2021. doi: 10.5194/amt-14-2575-2021.

Maignan, F., Bréon, F.-M., and Lacaze, R. Bidirectional reflectance of earth targets: Evaluation of analytical models using a large set of spaceborne measurements with emphasis on the hot spot. *Remote Sensing of Environment*, 90(2):210–220, 2004.

Maignan, F., Bréon, F.-M., Fédèle, E., and Bouvier, M. Polarized reflectances of natural surfaces: Spaceborne measurements and analytical modeling. *Remote Sensing of Environment*, 113(12):2642–2650, 2009.

Martins, J. V., Tanré, D., Remer, L., Kaufman, Y., Mattoo, S., and Levy, R. Modis cloud screening for remote sensing of aerosols over oceans using spatial variability. *Geophysical Research Letters*, 29(12):MOD4–1, 2002.

Martonchik, J. V., Diner, D. J., Kahn, R. A., Ackerman, T. P., Verstraete, M. M., Pinty, B., and Gordon, H. R. Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1212–1227, 1998.

Masson-Delmotte, V., Zhai, P., Priani, A., Connors, S., Péan, C., Berger, S., et al. IPCC, 2021: Climate Change 2021: The physical science basis. Contribution of working group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 2021.

Mishchenko, M. I., Travis, L. D., Kahn, R. A., and West, R. A. Modeling phase functions for dustlike tropospheric aerosols using a shape mixture of randomly oriented polydisperse spheroids. *J. Geophys. Res. Atmos.*, 102(D14):16831–16847, 1997. doi: 10.1029/96JD02110.

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(36):7325–7341, 1999.

Mishchenko, M. I., Travis, L. D., and Lacis, A. A. *Scattering, Absorption, and Emission of Light by Small Particles*. Cambridge University Press, 1 edition, June 2002. ISBN 052178252X.

Mishchenko, M. I., Cairns, B., Kopp, G., Schueler, C. F., Fafaul, B. A., Hansen, J. E., Hooker, R. J., Itchkawich, T., Maring, H. B., and Travis, L. D. Accurate

monitoring of terrestrial aerosols and total solar irradiance: introducing the glory mission. *Bulletin of the American Meteorological Society*, 88(5):677–692, 2007.

Moulin, C. and Chiapello, I. Evidence of the control of summer atmospheric transport of african dust over the atlantic by sahel sources from toms satellites (1979–2000). *Geophysical research letters*, 31(2), 2004.

Nakajima, T., Tonna, G., Rao, R., Boi, P., Kaufman, Y., and Holben, B. Use of sky brightness measurements from ground for remote sensing of particulate polydispersions. *Applied optics*, 35(15):2672–2686, 1996. doi: 10.1364/AO.35.002672.

Nakajima, T., Yoon, S.-C., Ramanathan, V., Shi, G.-Y., Takemura, T., Higurashi, A., Takamura, T., Aoki, K., Sohn, B.-J., Kim, S.-W., et al. Overview of the atmospheric brown cloud east asian regional experiment 2005 and a study of the aerosol direct radiative forcing in east asia. *Journal of Geophysical Research: Atmospheres*, 112 (D24), 2007.

Nakajima, T., Campanelli, M., Che, H., Estellés, V., Irie, H., Kim, S.-W., Kim, J., Liu, D., Nishizawa, T., Pandithurai, G., et al. An overview of and issues with sky radiometer technology and skynet. *Atmospheric Measurement Techniques*, 13(8): 4195–4218, 2020. doi: 10.5194/amt-13-4195-2020.

Olmo, F., Quirantes, A., Lara, V., Lyamani, H., and Alados-Arboledas, L. Aerosol optical properties assessed by an inversion method using the solar principal plane for non-spherical particles. *Journal of quantitative spectroscopy and radiative transfer*, 109(8):1504–1516, 2008.

Ortega, J. M. and Rheinboldt, W. C. *Iterative solution of nonlinear equations in several variables*. SIAM, 1970.

Pappalardo, G., Amodeo, A., Apituley, A., Comeron, A., Freudenthaler, V., Linné, H., Ansmann, A., Bösenberg, J., D'Amico, G., Mattis, I., et al. Earlinet: towards an advanced sustainable european aerosol lidar network. *Atmospheric Measurement Techniques*, 7(8):2389–2409, 2014.

Phillips, D. L. A technique for the numerical solution of certain integral equations of the first kind. *Journal of the ACM (JACM)*, 9(1):84–97, 1962. doi: 10.1145/321105.321114.

Popp, T., De Leeuw, G., Bingen, C., Brühl, C., Capelle, V., Chedin, A., Clarisse, L., Dubovik, O., Grainger, R., Griesfeller, J., et al. Development, production and evaluation of aerosol climate data records from european satellite observations (aerosol_cci). *Remote Sensing*, 8(5):421, 2016.

Press, W. H., Teukolsky, S. A., Flannery, B. P., and Vetterling, W. T. *Numerical recipes in Fortran 77: volume 1, volume 1 of Fortran numerical recipes: the art of scientific computing*. Cambridge university press, 1992.

Rahman, H., Verstraete, M. M., and Pinty, B. Coupled surface-atmosphere reflectance (csar) model: 1. model description and inversion on synthetic data. *Journal of Geophysical Research: Atmospheres*, 98(D11):20779–20789, 1993.

Remer, L. A., Tanré, D., Kaufman, Y., Ichoku, C., Mattoo, S., Levy, R., Chu, D., Holben, B., Dubovik, O., Smirnov, A., et al. Validation of modis aerosol retrieval over ocean. *Geophysical research letters*, 29(12):MOD3–1, 2002.

Remer, L. A., Kaufman, Y., Tanré, D., Mattoo, S., Chu, D., Martins, J. V., Li, R.-R., Ichoku, C., Levy, R., Kleidman, R., et al. The modis aerosol algorithm, products, and validation. *Journal of atmospheric sciences*, 62(4):947–973, 2005.

Rodgers, C. D. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Reviews of Geophysics*, 14(4):609–624, 1976. doi: 10.1029/rg014i004p00609.

Rodgers, C. D. Characterization and error analysis of profiles retrieved from remote sounding measurements. *Journal of Geophysical Research: Atmospheres*, 95(D5): 5587–5595, 1990. doi: 10.1029/jd095id05p05587.

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

Rolph, G., Stein, A., and Stunder, B. Real-time environmental applications and display system: Ready. *Environmental Modelling & Software*, 95:210–228, 2017. doi: 10.1016/j.envsoft.2017.06.025.

Román, R., Cazorla, A., Toledano, C., Olmo, F., Cachorro, V., de Frutos, A., and Alados-Arboledas, L. Cloud cover detection combining high dynamic range sky images and ceilometer measurements. *Atmospheric research*, 196:224–236, 2017. doi: 10.1016/j.rse.2017.05.013.

Román, R., Benavent-Oltra, J. A., Casquero-Vera, J. A., Lopatin, A., Cazorla, A., Lyamani, H., Denjean, C., Fuertes, D., Pérez-Ramírez, D., Torres, B., et al. Retrieval of aerosol profiles combining sunphotometer and ceilometer measurements in grasp code. *Atmospheric Research*, 204:161–177, 2018. doi: 10.1016/j.atmosres.2018.01.021.

Ross, J. *The radiation regime and architecture of plant stands*. Number 3. Springer Science & Business Media, 1981.

Roujean, J.-L., Leroy, M., and Deschamps, P.-Y. A bidirectional reflectance model of the earth's surface for the correction of remote sensing data. *Journal of Geophysical Research: Atmospheres*, 97(D18):20455–20468, 1992.

Sayer, A. M., Hsu, N. C., Lee, J., Kim, W. V., and Dutcher, S. T. Validation, stability, and consistency of modis collection 6.1 and viirs version 1 deep blue aerosol data over land. *Journal of Geophysical Research: Atmospheres*, 124(8):4658–4688, 2019.

Sayer, A. M., Govaerts, Y., Kolmonen, P., Lipponen, A., Luffarelli, M., Mielonen, T., Patadia, F., Popp, T., Povey, A. C., Stebel, K., et al. A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing. *Atmospheric Measurement Techniques*, 13(2):373–404, 2020. doi: 10.5194/amt-13-373-2020.

Schnaiter, M., Horvath, H., Möhler, O., Naumann, K.-H., Saathoff, H., and Schöck, O. Uv-vis-nir spectral optical properties of soot and soot-containing aerosols. *Journal of Aerosol Science*, 34(10):1421–1444, 2003.

Seinfeld, J. H. and Pandis, S. N. Atmospheric chemistry and physicsfrom air pollution to climate change. 2006.

Sinyuk, A., Holben, B. N., Eck, T. F., Giles, D. M., Slutsker, I., Korkin, S., Schafer, J. S., Smirnov, A., Sorokin, M., and Lyapustin, A. The aeronet version 3 aerosol retrieval algorithm, associated uncertainties and comparisons to version 2. *Atmospheric Measurement Techniques*, 13(6):3375–3411, 2020. doi: 10.5194/amt-13-3375-2020.

Stein, A., Draxler, R. R., Rolph, G. D., Stunder, B. J., Cohen, M., and Ngan, F. Noaa's hysplit atmospheric transport and dispersion modeling system. *Bulletin of the American Meteorological Society*, 96(12):2059–2077, 2015. doi: 10.1175/BAMS-D-14-00110.1.

Takamura, T. Overview of skynet and its activities. *Optica pura y aplicada*, 37(3): 3303–3308, 2004.

Tanré, D., Bréon, F., Deuzé, J., Dubovik, O., Ducos, F., François, P., Goloub, P., Herman, M., Lifermann, A., and Waquet, F. Remote sensing of aerosols by using polarized, directional and spectral measurements within the a-train: the parasol mission. *Atmospheric Measurement Techniques*, 4(7):1383–1395, 2011.

Tikhonov, A. N. On the solution of ill-posed problems and the method of regularization. In *Doklady Akademii Nauk*, volume 151, pages 501–504. Russian Academy of Sciences, 1963.

Torres, B. and Fuertes, D. Characterization of aerosol size properties from measurements of spectral optical depth: a global validation of the grasp-aod code using long-term aeronet data. *Atmospheric Measurement Techniques*, 14(6):4471–4506, 2021. doi: 10.5194/amt-14-4471-2021.

Torres, B., Dubovik, O., Toledano, C., Berjón, A., Cachorro, V. E., Lapyonok, T., Litvinov, P., and Goloub, P. Sensitivity of aerosol retrieval to geometrical configuration of ground-based sun/sky radiometer observations. *Atmospheric Chemistry and Physics*, 14(2):847–875, 2014. doi: 10.5194/acp-14-847-2014.

Torres, B., Dubovik, O., Fuertes, D., Schuster, G., Cachorro, V. E., Lapyonok, T., Goloub, P., Blarel, L., Barreto, A., Mallet, M., et al. Advanced characterisation of aerosol size properties from measurements of spectral optical depth using the grasp algorithm. *Atmospheric measurement techniques*, 10(10):3743–3781, 2017. doi: 10.5194/amt-10-3743-2017.

Torres, O., Bhartia, P., Herman, J., 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(3):398–413, 2002.

Torricella, F., Cattani, E., Cervino, M., Guzzi, R., and Levoni, C. Retrieval of aerosol properties over the ocean using global ozone monitoring experiment measurements: method and applications to test cases. *Journal of Geophysical Research: Atmospheres*, 104(D10):12085–12098, 1999.

Tsekeri, A., Lopatin, A., Amiridis, V., Marinou, E., Igloffstein, J., Siomos, N., Solomos, S., Kokkalis, P., Engelmann, R., Baars, H., et al. Garrlic and liric: strengths and limitations for the characterization of dust and marine particles along with their mixtures. *Atmospheric Measurement Techniques*, 10:4995–5016, 2017. doi: 10.5194/amt-10-4995-2017.

Twomey, S. Comparison of constrained linear inversion and an iterative nonlinear algorithm applied to the indirect estimation of particle size distributions. *Journal of Computational Physics*, 18(2):188–200, 1975. doi: 10.1016/0021-9991(75)90028-5.

Twomey, S. Introduction to the mathematics of inversion in remote sensing and indirect measurement(book). *Amsterdam, Elsevier Scientific Publishing Co.(Developments in Geomathematics*, (3):253, 1977.

Twomey, S. On the numerical solution of fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature. *Journal of the ACM (JACM)*, 10(1):97–101, 1963. doi: 10.1145/321150.321157.

Veefkind, J. P., de Leeuw, G., Stammes, P., and Koelemeijer, R. B. Regional distribution of aerosol over land, derived from atsr-2 and gome. *Remote Sensing of Environment*, 74(3):377–386, 2000.

Wagner, S., Govaerts, Y., and Lattanzio, A. Joint retrieval of surface reflectance and aerosol optical depth from msg/seviri observations with an optimal estimation approach: 2. implementation and evaluation. *Journal of Geophysical Research: Atmospheres*, 115(D2), 2010.

Wang, M., Bailey, S., and McClain, C. R. Seawifs provides unique global aerosol optical property data. *Eos, Transactions American Geophysical Union*, 81(18):197–202, 2000.

Wanner, W., Li, X., and Strahler, A. On the derivation of kernels for kernel-driven models of bidirectional reflectance. *Journal of Geophysical Research: Atmospheres*, 100(D10):21077–21089, 1995.

Waquet, F., Cairns, B., Knobelspiesse, K., Chowdhary, J., Travis, L. D., Schmid, B., and Mishchenko, M. Polarimetric remote sensing of aerosols over land. *Journal of Geophysical Research: Atmospheres*, 114(D1), 2009.

Welton, E. J., Campbell, J. R., Spinhirne, J. D., and Scott III, V. S. Global monitoring of clouds and aerosols using a network of micropulse lidar systems. In *Lidar Remote Sensing for Industry and Environment Monitoring*, volume 4153, pages 151–158. International Society for Optics and Photonics, 2001.

Xu, X., Wang, J., Zeng, J., Spurr, R., Liu, X., Dubovik, O., Li, L., Li, Z., Mishchenko, M. I., Siniuk, A., et al. Retrieval of aerosol microphysical properties from aeronet photopolarimetric measurements: 2. a new research algorithm and case demonstration. *Journal of Geophysical Research: Atmospheres*, 120(14):7079–7098, 2015.

Yu, H., Zhang, Y., Chin, M., Liu, Z., Omar, A., Remer, L. A., Yang, Y., Yuan, T., and Zhang, J. An integrated analysis of aerosol above clouds from a-train multi-sensor measurements. *Remote sensing of environment*, 121:125–131, 2012.

Yu, H., Remer, L. A., Kahn, R. A., Chin, M., and Zhang, Y. Satellite perspective of aerosol intercontinental transport: From qualitative tracking to quantitative characterization. *Atmospheric Research*, 124:73–100, 2013.

Zieger, P. C. *Effects of relative humidity on aerosol light scattering*. PhD thesis, ETH Zurich, 2011.