THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *28/11/2014* **par :**

BOJAN SIČ

Amélioration de la représentation des aérosols dans un modèle de chimie-transport : Modélisation et assimilation de données

JURY

JEAN-LUC ATTIÉ (UNIVERSITÉ TOULOUSE III)	Président du Jury
MATTHIAS BEEKMANN (LISA)	Rapporteur
PATRICK CHAZETTE (CEA/DSM/LSCE)	Rapporteur
LAURENT MENUT (LMD)	Rapporteur
ANGELA BENEDETTI (ECMWF, UK)	Examinateur
VIRGINIE MARÉCAL (CNRM)	Directeur de thése
LAAZIZ EL AMRAOUI (CNRM)	Codirecteur de thése
ALAIN DABAS (CNRM)	Co-encandrant de thése

École doctorale et spécialité :
SDU2E : Océan, Atmosphère et Surfaces Continentales

Unité de Recherche :
Centre National de Recherches Météorologiques – Groupe d'études de l'Atmosphère Météorologiques (UMR 3589)

Directeur(s) de Thèse :
Virginie MARÉCAL, *Laaziz EL AMRAOUI* et *Alain DABAS*

Rapporteurs :
Matthias BEEKMANN, *Patrick CHAZETTE* et *Laurent MENUT*

Improvement of aerosol representation in a chemical-transport model: Modelling and data assimilation

Summary:

The improvement of the aerosol representation in a chemical transport model (CTM) permits better understanding of aerosol properties, forecasts, and their widespread effects. The main goal of this thesis is to improve the aerosol representation in the CTM MOCAGE. The work may be divided into two approaches to achieve the main goal: the direct improvement of aerosol physical schemes and parameterizations, and the development of a data assimilation system able to assimilate aerosol optical depth (AOD) and lidar profiles into the model.

On the modelling side, the processes that underwent the important improvements were sea salt, desert dust and volcanic aerosol emissions, wet deposition and sedimentation. The ambition is related to improve the model biases compared to observations, and to implement more physically detailed schemes in the model. We evaluated the impacts of these changes and compared the modelled fields to observations. The implemented updates significantly enhanced the model agreement with the observations and the inter-model comparison data. The results also confirmed that large uncertainties in models can come from the use of different parameterizations.

The aerosol data assimilation is implemented to further reduce the model uncertainties. The set of observation operators and their tangent linear and adjoint operators for AOD and lidar profile observations are developed to link the model and the observation space. Aerosol assimilation proved to be very efficient to reduce the differences between the model and the observations. AOD observations assimilated for the periods of the extensive field campaigns over the Mediterranean basin in 2012 and 2013 and the period of volcanic ash plume from the Eyjafjöll eruption in 2010 in Iceland showed that the AOD assimilation is able to significantly improve the model performance in terms of AOD, but also other aerosol parameters such as concentrations. Assimilation of different elastic backscatter lidar profile measurements, namely of the backscatter signal, the extinction coefficient and the backscatter coefficient, also showed an efficient influence on the modelled aerosol vertical profiles.

Amélioration de la représentation des aérosols dans un modèle de chimie-transport : Modélisation et assimilation de données

Resumé:

L'amélioration de la représentation des aérosols dans un modèle de chimie-transport (en anglais, Chemical-Transport Model, CTM) permet une meilleure compréhension des propriétés des aérosols et leurs nombreux effets ainsi que leur prévision. Sur cette base, l'objectif général de cette thèse est d'améliorer la représentation des aérosols dans le CTM MOCAGE. Pour remplir les objectifs de la thèse, dans un premier temps, nous avons modifié directement la représentation des aérosols dans le CTM MOCAGE en réexaminant et améliorant les différents processus déjà présents via la prise en compte de schémas et de paramétrisations plus détaillés. Les processus ayant subi les améliorations les plus importantes sont les émissions des aérosols du type sel marin, poussière désertique et cendre volcanique, le dépôt humide et la sédimentation. Nous avons évalué les impacts de ces changements et comparé les champs modélisés avec des observations. Les modifications implémentées ont permis d'améliorer significativement l'accord entre modèle et observations et celui du modèle avec les données de la comparaison inter-modèle AeroCom.

Comme approche complémentaire pour répondre à l'objectif de cette thèse, nous avons également implémenté dans le CTM MOCAGE l'assimilation de données des aérosols. Le système d'assimilation de données est capable d'assimiler les observations de l'épaisseur optique des aérosols (en anglais, Aerosol Optical Depth, AOD), ainsi que les différentes grandeurs obtenues par les mesures lidar. Nous avons réalisé une validation rigoureuse du système d'assimilation de l'AOD en assimilant les données de l'instrument MODIS pour les périodes correspondant aux campagnes de mesure des projets TRAQA (TRAnsport à longue distance et Qualité de l'Air dans le bassin méditerranéen) et ChArMEx (CHemistry and AeRosol MEditerranean EXperiment) durant l'été 2012 et 2013, respectivement, et en comparant les champs directement modélisés et assimilés avec les observations. Lorsque l'on compare ces résultats avec des observations indépendantes d'AOD, les champs assimilés ont des indicateurs statistiques meilleurs que ceux du modèle direct. Les mesures in-situ de TRAQA et ChArMEx ont également été utilisées pour évaluer l'impact de l'assimilation des AOD sur d'autres paramètres des aérosols modélisés. Les résultats ont montré que l'assimilation des AOD était un outil particulièrement efficace pour améliorer les performances du modèle en terme d'AOD, mais également pour les autres paramètres des aérosols tel que la concentration.

Les observations lidar fournissent des informations importantes sur la répartition des aérosols sur la verticale. Nous avons implémenté un système d'assimilation des profils lidar mesurés pour des lidars à rétrodiffusion élastique. Le système est capable d'assimiler le signal de rétrodiffusion et les grandeurs inversées: les coefficients de rétrodiffusion et d'extinction. Les premiers tests effectués à partir des données d'un lidar au sol ont montré un impact cohérent sur le modèle en assimilant séparément les 3 différents types de mesures lidars implémentées. Ceci constitue un premier résultat très prometteur.

Acknowledgements:

Finishing a PhD is a long and demanding journey. Through it, many people helped me and made easier to pass all unfamiliar terrain that I encountered. I am indebted to all of them, and, here, I would like to acknowledge all their support.

First of all, I am grateful to my supervisors Virginie Marécal and Laaziz El Amraoui, for being a great support during my whole PhD pursuit. They supported me with great promptness and care, and were always encouraging and patient with me throughout the thesis. They taught how a good research is done, and many insightful suggestions and discussions with them led to a better understanding of the problems and provided a constant advancement. During the thesis, it is Laaziz with whom I worked the most closely. All his scientific advices and knowledge were a priceless help for the accomplished work. He always listened and encouraged my new ideas, guided me in good directions, and his hard work gave me an excellent example. The same appreciation goes to Virginie. Her flexibility in scheduling, gentle encouragement and relaxed behaviour helped me a lot and made me to believe that we were doing a good thing, and especially were important at the end, when all what we did during the thesis had to be transferred to the manuscript. Also, the supervisors taught me that it is not only about science that we learn during a PhD. They gave important insights that helped me to grow as a researcher, colleague and person. Thank you for all that. I extend my gratitude also to Alain Dabas whose remarkable expertise in lidars and all his help were indispensable for our lidar assimilation.

Besides supervisors, there very numerous colleagues that I have been privileged to get to know and to work with. From all in CARMA and CAIAC groups I learned a lot about different problems, techniques to solve them and in general, life. I would also like to thank to the board of CNRM/Meteo-France for believing in me and in the project from the very beginning, and for giving us all the necessary support.

I would like to express gratitude to Jean-Luc Attié, who multiple times was there for me when it was time to choose my path (before my master internship and before the PhD thesis). For sure, this had a great impact on my achievements, and without this help I would not have arrived where I am today.

Special thanks go also to my master thesis supervisor Philippe Ricaud who gave me an excellent basis how to do research well, which permitted me to start well my PhD thesis; and who, by chance, followed me and moved to CNRM soon after I started my thesis, which allowed me to have his help in many numerous ways during the PhD thesis.

At the beginning of the second year of my thesis, before we seriously tackled with the aerosol assimilation, Laaziz and I visited ECMWF in Reading. There, we met Angela Benedetti and Jean-Jacques Morcrette who generously explained us and transferred their experience in the aerosol assimilation. This helped us to better and faster identify the the main steps, problems and choices that we have for the development of the aerosols assimilation algorithm. For all the help during the visit and later, I am indebted to Angela and Jean-Jacques.

The implementation of the aerosol assimilation in the Valentina assimilation system would not have been possible without expertise of Cerfacs team. I am indebted to Andrea Piacentini and Emanuele Emili for their unselfish help. Special gratitude goes to Andrea for all the help in the implementation of the developed observation operators to Valentina. His willingness to help and unmatched expertise in Valentina and assimilation systems in general were of unparalleled significance for my work.

I would also like to thank my committee members Mathias Beekmann, Patrick Chazette, Laurent Menut, Jean-Luc Attié and Angela Benedetti for serving at my thesis committee, giving me constructive comments and suggestions, and making my defense a very enjoyable moment.

Finally, I would like to thank to my family, especially to my wife Elena, and my mother Orhidea. I met my wife during the thesis, and she has been a great support to me, especially at the end of thesis when the finish line was visible and the load was the highest. For my mum, words cannot express how grateful I am to her for all of the sacrifices that she have made for me. She knows well that this journey started well, well before beginning of the thesis and it was full of difficulties that we passed together. For all that she did for me, she knows well that all my successes are her successes, too.

Contents

Introduction

Aerosols, particles suspended in the atmosphere, have widespread effects on the Earth and people: they are an important factor in the air quality, the climate system, the Earth radiative budget, the cloud formation, the hydrological cycle, the civil aviation, the atmospheric circulation and chemistry. Aerosols in polluted areas are linked with human health problems. Another direct effect of aerosols is the scattering and absorption of the radiation, which directly lowers the visibility. Most of aerosols strongly reflect the sunlight and increase the planet albedo, but carbon-containing aerosols efficiently absorb the light. All this results in cooling of the planet surface, and heating of the atmosphere, and affects the temperature profile, evaporation, clouds, precipitation, etc. Aerosols also influence clouds more directly by acting as condensation and ice nuclei for the cloud particle formation. Still, there are many other pathways and feedbacks which make the total effect on clouds, precipitation and the climate difficult to assess. It is believed that the overall effect of aerosols is to cool the climate system, but this effect could reverse in the future (IPCC, 2013). Moreover, aerosols contribute to different processes as: the production of tropospheric ozone, destruction of stratospheric ozone, changes in circulation patterns, acid deposition, eutrophication, threat to airplane engines, etc.

Studying aerosol effects is difficult because aerosols take a variety of sizes, shapes and forms, and their quantities can vary quickly over time and space. Is necessary to better understand aerosol distribution, composition and effects, and their changes due to industrialization, population growth, or air quality standards. This requests to continuously observe them and accurately simulate them in models.

Observations help to determine aerosol properties, variability and their impact on the environment. Aerosols are observed by ground-based instruments, during dedicated field campaigns and from satellites. The ground-based instruments can provide detailed and accurate in-situ measurements, or make remotely sensed measurements by the network of stations, like the AERONET network (AErosol RObotic NETwork, Holben et al., 1998). The same applies to the field campaigns, during which in-situ and remotely sensed data is usually intensively collected by extensive instrumental resources (aircrafts, balloons, ground stations). This data generally has a good horizontal and vertical coverage, but is limited to short periods of time during campaigns. Satellite

instruments provide routine aerosol measurements on a global scale for already more than 30 years allowing to establish the aerosol climatology and long-term trends. The satellites generally measure the integrated quantities, like aerosol optical depth (AOD). One of the most used data includes MODIS (MODerate resolution Imaging Spectrometer, Remer et al., 2005) and SEVIRI (Spinning Enhanced Visible and InfraRed Imager, Thieuleux et al., 2005) data. The limb-viewing instruments or the spaceborne lidars like CALIOP (Cloud-Aerosol LIdar with Orthogonal Polarization, Winker et al., 2007) can provide the vertical profile information.

On the modelling side, the goal of aerosol models is to provide a realistic aerosol representation in the atmosphere both in time and space. The state-of-the-art models became powerful tools to study aerosols. Each subsequent IPCC (Intergovernmental Panel on Climate Change) report – the reference publications on the subject of climate change – notes considerable improvements in aerosol models (IPCC, 2001, 2007, 2013). As knowledge and tools advance, more physically complete representations of aerosols are implemented and simulated. However, the comparisons with observations and inter-comparisons between models still show significant discrepancies in the model results and still large uncertainties (Textor et al., 2007; Mann et al., 2013).

Increased efforts remain to improve the performance of the models. One of the ways is to directly improve the model components. This includes the development of more physically detailed parameterizations, the development of the explicit, particle-resolved schemes of known processes, the addition of new processes in the model, improvement of the spatial and temporal model resolution, etc. Another approach to improve the performance is to incorporate into models the information obtained from observations using data assimilation techniques. A data assimilation system takes into account the uncertainties of both the model and the observations and searches for the state of system which corresponds to the model and observations in the most optimal way. Data assimilation steadily becomes an important technique also in the field of atmospheric chemistry and aerosol modelling.

The chemical transport model (CTM) MOCAGE developed in CNRM (Centre National de Recherches Météorologiques, Toulouse, France) simulates gases and aerosols. It is used for operational purposes and atmospheric research studies. The interest in the aerosol modelling in MOCAGE is related to a number of fields. One of them is air quality forecasts by Météo-France, in the projects MACC for Europe (Monitoring Atmospheric Composition & Climate, www.gmes-atmosphere.eu) and PREV'AIR for France (www.prevair.org/en).

Also, in Toulouse is located one of the Volcanic Ash Advisory Centres (VAAC) responsible for issuing volcanic ash warnings for civil aviation.

The Toulouse VAAC responsibility covers Africa and a big part of Eurasia. MOCAGE simulations contribute to information used by the VAAC in predicting the threat from the volcanic eruptions.

Moreover, desert dust outbreaks have an important influence on an extended region around the Saharan desert, including the Mediterranean basin and the South Europe. This rises the interest for the model forecasts of such events. Besides operational purposes, the model MOCAGE is a tool well adapted and used for aerosol research studies which contribute to the better knowledge about aerosols.

Data assimilation in MOCAGE was implemented for atmospheric gases, and it gives more accurate analyses with the coverage and resolution that the model can provide. The assimilation system was already applied to studies of the polar vortex dynamics (El Amraoui et al., 2008b), the ozone destruction in the polar vortex (El Amraoui et al., 2008a), the stratospheric tropic/mid-latitude exchange (Bencherif et al., 2007), the troposphere-stratosphere transport (El Amraoui et al., 2010; Barré et al., 2012), etc. Based on this expertise on atmospheric chemistry data assimilation, at the start of this thesis it was estimated that the implementation of data assimilation for aerosols in the model could significantly add up to the forecasting and research capabilities of the aerosol module of MOCAGE.

The main objective of this thesis is to improve the aerosol representation in the CTM MOCAGE. In the first part, this consists of evaluating the performance of the MOCAGE aerosol module and reexamining and improving different schemes and parameterizations in the model. The processes that underwent the most important improvements are aerosol emissions, wet deposition and sedimentation. The ambition is related to solving the known model biases compared to observations, and to implement more physically detailed schemes in the model. This work is presented in Sič et al. (2015).

The second objective of the thesis is to implement an aerosol data assimilation system for both AOD and lidar profiles. The aim of using the aerosol data assimilation is to reduce the model uncertainties. Aerosols are observed in physical quantities that are usually different than prognostic modelled variables. This requires the development of the set of observation operators and their tangent linear and adjoint operators to link the model and the observation spaces. The observations for which we develop and implement observation operators are the aerosol optical depth, as the most available observations nowadays, and the lidar profiles with their valuable information on the vertical profile. The implemented assimilation system is validated by an extensive set of observations.

This manuscript is organised as follows. The first chapter presents the theoretical, experimental basis and the advances in the aerosol science. The second chapter deals with the principles of aerosol modelling and introduces the chemical transport model MOCAGE. The third chapter describes the basis and advances in the field of data assimilation. The chapters devoted to the results obtained during the work on this thesis are: Chapter 4 which presents the developments done of the aerosol module in the CTM MOCAGE, and contains the study of Sič et al. (2015) submitted to the journal *Geoscientific Model Development*; and Chapter 5 which presents the development and validation of the assimilation of the aerosol optical depth and lidar profiles in the CTM MOCAGE. Appendices further explain and detail some developments and results presented in the main part of the manuscript.

Introduction (en français)

Les aérosols, particules en suspension dans l'atmosphère, ont de nombreux effets sur la Terre et les gens: ils sont un facteur important dans la qualité de l'air, le système climatique, le bilan radiatif de la Terre, la formation des nuages, le cycle hydrologique, l'aviation civile, la circulation atmosphérique et la chimie atmosphérique. Les aérosols dans les zones polluées causent des problèmes de santé. Un autre effet direct des aérosols est la diffusion et l'absorption de la radiation, ce qui réduit la visibilité. La plupart des aérosols reflètent fortement la lumière du soleil et augmentent l'albédo de la planète, mais les aérosols contenant du carbone absorbent efficacement la lumière. Cela se traduit par un refroidissement de la surface de la planète, et le chauffage de l'atmosphère, et modifie le profil de la température, l'évaporation, les nuages, les précipitations, etc. Les aérosols influencent les nuages plus directement en agissant comme noyaux de condensation et noyaux glacigènes pour la formation de particules dans les nuages.

D'autres mécanismes d'interaction des aérosols avec l'atmosphère s'ajoutent, qui rendent difficile l'évaluation de leur effet sur les nuages, les précipitations et le climat. On croit que les aérosols refroidissent le système climatique, mais cela pourrait inverser dans le futur (IPCC, 2013). De plus, les aérosols contribuent à différents processus: la production de l'ozone troposphérique, la destruction de l'ozone stratosphérique, les changements de circulation atmosphérique, les dépôts acides, l'eutrophisation, la détérioration des moteurs d'avion, etc.

L'étude des effets des aérosols est difficile parce que les aérosols ont une importante variété de tailles et de formes, et leurs quantités peuvent varier rapidement dans le temps et l'espace. Une meilleure compréhension de la distribution, de la composition et des effets des aérosols,requiert des observations en continue et des modélisations réalistes des aérosols.

Les observations aident à déterminer les propriétés des aérosols, la variabilité et leur impact sur l'environnement. Les aérosols sont observés par des instruments au sol, des campagnes de mesures et des satellites. Les instruments basés au sol peuvent fournir des mesures in situ précises et détaillées, ou des mesures de télédétection par le réseau de stations, comme le réseau AERONET

(AErosol RObotic NETwork, Holben et al., 1998). Le même type d'observations s'applique aux campagnes de mesures. Les données in situ et de télédétection peuvent être mesurées par de nombreux instruments situés dans des avions, des ballons, ou des stations au sol. Ces données ont généralement une bonne couverture horizontale et verticale, mais se limitent à de courtes périodes de temps pendant les campagnes de mesures. Les instruments satellitaires fournissent régulièrement des mesures d'aérosols à l'échelle globale depuis plus de 30 ans, permettant d'établir la climatologie des aérosols et des tendances à long terme. Les satellites mesurent, en général, des quantités intégrées, comme l'épaisseur optique des aérosols (AOD). Les instruments au limbe ou les lidars spatiaux comme CALIOP (Cloud-Aerosol LIdar with Orthogonal Polarization, Winker et al., 2007) peuvent fournir les informations sur le profil vertical.

En ce qui concerne la modélisation des aérosols, le but est de fournir une représentation réaliste dans l'atmosphère à la fois dans le temps et dans l'espace. Les modèles état de l'art sont devenus des outils puissants pour étudier les aérosols. Chaque rapport successif du GIEC (Groupe d'experts intergouvernemental sur l'évolution du climat) - publications de référence sur le sujet du changement climatique - note des améliorations considérables dans les modèles d'aérosols (IPCC, 2001, 2007, 2013). Comme les connaissances et les outils avancent, des représentations plus complètes de la physique des aérosols sont intégrées et simulées. Cependant, les inter-comparaisons de modèles et les comparaisons avec les observations montrent encore des écarts importants et de grandes incertitudes (Textor et al., 2007; Mann et al., 2013).

Des efforts supplémentaires sont nécessaires afin d'améliorer les performances des modèles. L'un des moyens est d'améliorer directement les modèles. Ceci inclut le développement de paramétrisations physiques plus détaillées, le développement de schémas explicites, de schémas à particules résolues, l'ajout de nouveaux processus dans le modèle, l'amélioration de la résolution spatiale et temporelle du modèle, etc. Une autre approche pour améliorer les performances est d'intégrer les informations obtenues par des observations dans les modèles en utilisant des techniques d'assimilation de données. Un système d'assimilation des données prend en compte les incertitudes à la fois du modèle et des observations et cherche l'état du système qui correspond le mieux au modèle et aux observations. L'assimilation de données devient progressivement une technique importante dans le domaine de la chimie atmosphérique et de la modélisation des aérosols.

MOCAGE est un modèle de chimie-transport (CTM) développé au CNRM (Centre National de Recherches Météorologiques, Toulouse, France), qui simule les gaz et les aérosols. Il est utilisé à des fins opérationnelles et scientifiques. L'intérêt pour la modélisation des aérosols dans MOCAGE est lié à un certain nombre de domaines.

L'un de ces domaines est la prévision de la qualité de l'air par Météo-France, dans les projets MACC pour l'Europe (Monitoring Atmospheric Composition & Climate, www.gmes-atmosphere.eu) et PREV'AIR pour la France (www.prevair.org/en).

À Toulouse est situé l'un des centres VAAC (Volcanic Ash Advisory Centre) responsable d'avertir la présence de cendres volcaniques pour l'aviation civile. La responsabilité du VAAC de Toulouse couvre l'Afrique et une grande partie de l'Eurasie. Les simulations de MOCAGE contribuent à l'information utilisée par le VAAC pour prévoir le danger des éruptions volcaniques.

En outre, les événements de fortes émissions de poussières désertiques ont une influence importante autour du désert du Sahara, y compris le bassin méditerranéen et l'Europe du Sud. Cela augmente l'intérêt des prévisions des poussières désertiques dans le modèle.

Outre des fins opérationnelles, le modèle MOCAGE est un outil adapté pour des études de recherche sur les aérosols qui contribuent à leur meilleure connaissance.

L'assimilation de données dans MOCAGE a été implémentée pour les gaz atmosphériques. Elle donne des analyses plus précises avec la couverture et la résolution que le modèle fournit. Le système d'assimilation a déjà été appliqué à l'étude de la dynamique des vortex polaires (El Amraoui et al., 2008b), la destruction de l'ozone dans le vortex polaire (El Amraoui et al., 2008a), l'échange tropique/moyenne latitude dans la stratosphère (Bencherif et al., 2007), le transport troposphère-stratosphère (El Amraoui et al., 2010; Barré et al., 2012), etc. Basée sur cette expertise de l'assimilation des données de la chimie atmosphérique, au début de cette thèse, il a été estimé que la mise en œuvre de l'assimilation de données pour les aérosols dans le modèle MOCAGE pourrait améliorer de manière significative les capacités du modèle à prévoir et à étudier les aérosols.

L'objectif principal de cette thèse est d'améliorer la représentation des aérosols dans le CTM MOCAGE. Dans la première partie, cela consiste à évaluer les performances du module d'aérosol de MOCAGE, réexaminer et améliorer les schémas et les différentes paramétrisations dans le modèle. Des améliorations importantes sont implémentées dans les schémas d'émissions d'aérosols, les dépôts humides et la sédimentation. L'ambition est de résoudre les biais des modèles connus par rapport aux observations, et de mettre en œuvre des schémas physiques plus détaillés dans le modèle. Ce travail est présenté dans Sič et al. (2015).

Le deuxième objectif de la thèse est de mettre en place un système d'assimilation des données d'aérosol pour les AOD et les profils lidar. Le but

d'utiliser l'assimilation de données des aérosols est de réduire les incertitudes du modèle. Généralement, les observations des aérosols fournissent des grandeurs physiques différentes de celles du modèle. Cela nécessite le développement de l'ensemble des opérateurs d'observation, leur linéaires tangentes et adjointes pour lier les espaces d'observations et de modèle. Les observations pour lesquelles nous avons développé et mis en œuvre les opérateurs d'observation sont celles de l'AOD, ce sont les observations les plus disponibles de nos jours, et les profils lidar avec leurs informations sur le profil vertical. Le système d'assimilation mis en œuvre est validé par un important jeu d'observations.

Ce manuscrit est organisé comme suit. Le premier chapitre présente les bases théoriques et expérimentales ainsi que les avancées de la science sur les aérosols. Le deuxième chapitre traite des principes de modélisation des aérosols et introduit le modèle de chimie-transport MOCAGE. Le troisième chapitre décrit les bases et les avancées dans le domaine de l'assimilation de données. Les chapitres consacrés aux résultats obtenus au cours des travaux de cette thèse sont les suivants: Chapitre 4, qui présente les développements réalisés pour les aérosols dans le CTM MOCAGE, et contient l'étude de Sič et al. (2015) soumis à la revue; et le chapitre 5, qui présente le développement et la validation de l'assimilation d'AOD et des profils lidar dans le CTM MOCAGE. Les annexes expliquent et détaillent certains développements et résultats présentés dans la partie principale du manuscrit.

Part I

Theoretical Background

1

Aerosols

This chapter describes the theoretical, experimental basis and the advances in the aerosol science.

Contents

1.1 Definition

1.1.1 Introduction

In the atmosphere, there are atoms and molecules, and they present the main constituents of the atmosphere — gases. But besides gases, in the air one can find also suspended particles — *aerosol particles*. Aerosol is a suspension of particles which float in the air and which can be either, in a solid or liquid state. In this manuscript, the term *aerosols* will be used in the sense to refer to aerosol particles. The term *aerosol* appeared in 1920s, to describe clouds of microscopic particles in the air. It was coined from *aero* – "air" and *sol* – short of "solution".

Aerosols reside in the atmosphere for at least several seconds, their typical sizes vary between a few nanometres and several tens of micrometers, and they could be of either natural or anthropogenic origin. An understanding of the aerosols is of a great importance because aerosols can influence climate. This may be done in different ways: directly through scattering and absorbing radiation (see Section 1.8) and indirectly by having a role in the cloud formation and condensation, modifying the properties and lifetime of clouds and influencing the hydrological cycle. Aerosols also influence the air quality as the different types of aerosols act as air pollutants.

Research history The aerosol research advanced at the beginning of the 20th century, when the size of the aerosols was at the reach of scientific experimental observations. At that time the smaller constituents of matter (atoms) were still too small for the observations and experimentation. In the beginning, the research of aerosols rested mainly on the experiments on the mechanics of aerosols, and the goals were to explore and understand concepts like diffusion, or Brownian motion. In the second half of the 20th century, the importance of aerosol research increased as the awareness and knowledge of the aerosol effects on environment and health increased. At the time when industry had an important expansion, the anthropogenic aerosol emissions increased manifold. That fast development also produced the growth in air pollution concerns and research. As the high-technologies developed, the aerosols started to be used in many production processes, or being avoided in others (clean room environments). In the last decade of the 20th century, it became clear that the aerosol effect on the climate could be important, and the number of studies

on this effect increased. Each IPCC (Intergovernmental Panel on Climate Change) report, which are the reference publications on the subject of climate change, brings considerable progress on the knowledge of the aerosol climate influence.

Particulate matter Particles in the air are sometimes called 'particulate matter' (or abbreviated PM). They include all aerosols: dust, smoke, dirt, soot... Usually, the abbreviation PM is used for marking the aerosols in the air up to a certain size. So, we have PM_1, $PM_{2.5}$, PM_{10} designations, which indicate all aerosol particles with aerodynamical diameters smaller than the number of micrometers indicated in the name: 1 µm, 2.5 µm, 10 µm, respectively. The term aerodynamical diameter of an irregular particle refers to the diameter of the spherical particle of density of $1000\,\mathrm{kg\,m^{-3}}$ (the same as of the water) which falls with the same settling velocity as the irregular particle. These PM designations are important in the air quality in order to distinguish the sizes of the particles and their diverse effects.

1.1.2 Aerosol characterisation and classifications

Aerosols are classified based on different criteria: according to their origin, to their sources, to their sizes, to their chemical composition, or their mixing state.

Source In the division based on aerosol sources, the term *primary aerosols* refers to all particles that are directly injected to the atmosphere, and the term *secondary aerosols* refers to all particles that are formed in the atmosphere by chemical reactions between different aerosol particles, or between aerosols particles and gases.

Size One of the most obvious parameters to characterize aerosols is their size. The aerosol behavior and properties depend on their size, as well as many physical laws that govern them. Aerosols cover a wide range of different sizes. Typically from a few tens of angstroms to several hundreds of micrometers in diameter (in this manuscript, if not stated differently, the sizes of particles refer to particle diameters). In the classification based on the aerosol sizes, the term *monodisperse aerosols* refers to particles which are all the same, all of the them having the same size, physical and chemical properties. Such groups of particles are usually created in laboratories and used for the laboratory

experiments. In nature, the aerosols are usually *polydisperse*, which means that they have different sizes, and maybe similar but not exactly the same physical or chemical properties. To study them, it is necessary to observe them, and characterize statistically their properties. To describe the sizes of particles in a polydisperse aerosol distribution, it is necessary to define a particle size distribution which gives relative (or absolute) amounts of particles sorted by size (see Section 1.5). The first distinction in the size distribution is the distinction between fine and coarse particles. Fine particles are generally referred to have their diameters less than 2.5 µm, while coarse particles above 2.5 µm. Aerosols are usually arranged in different populations, which are on the size distribution characterized by peaks on certain sizes. These groups are called *modes*. In the domain of coarse particles, there is usually only one mode which is called *coarse mode*. And in the domain of fine particles are usually two modes: *nucleation* and *accumulation* modes.

While defining sizes of aerosols it is important to also consider their shapes. Aerosols can take many different shapes. Liquid aerosol particles have a spherical shape, hydrophilic particles can change their shape during time, but tend to take nearly spherical shape. Solid particles can have very complex shapes. In order to develop theories of the aerosol properties and to efficiently model them, it is often necessary to simplify and to consider aerosols as spherical particles. In this sense, the size refers to the equivalent size of a hypothetical spherical particle which has the same certain property as an irregular particle whose size we want to define. Already mentioned aerodynamical diameter is an example of equivalent diameter. Another example is *surface area equivalent diameter* which refers to the diameter of an aerosol particle equal to the diameter of a sphere with same surface area as the particle. Other equivalent diameters are also defined: volume equivalent diameter, sieve equivalent diameter, mobility equivalent diameter, etc. Particles with shapes that greatly differ from spheres (like long filaments, for example) can be considered with simplified non-spherical shapes (for example, cylinders, long ellipsoids, etc). Also, if it is known how much shapes of aerosols differ from spheres, sometimes it is possible to apply correction factors to compensate for the effects of non-sphericity of the aerosols. In the case of a group of particles clustered as aggregates (for example smoke particles), the shape can be defined by fractal dimensions (Hinds, 1999).

Composition If we classify aerosols according to their chemical composition, we distinguish *inorganic* and *organic* aerosols. Inorganic aerosols are composed of inorganic substance, while organic aerosols contain organic carbon in their composition. Aerosols in the air are usually a complex mixture.

Origin Aerosols in nature are a complex mixture of the aerosols emitted from natural sources and aerosols emitted by human activities. They dominate different environments, natural sources in rural areas, and anthropogenic sources in urban areas. But, once in the atmosphere, it is difficult to distinguish aerosols from natural and anthropogenic sources, as even individual particles can be mixed and composed of materials from natural and anthropogenic origin. (Kaufman et al., 2002).

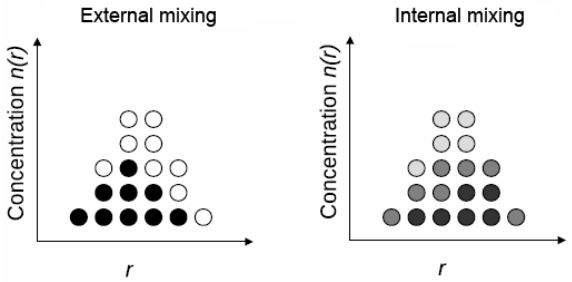

Figure 1.1: Illustration of mixing states of aerosols. The figure is from Boucher (2012).

Mixing states Aerosols can be in different mixing states. If aerosols of two different types and two different chemical compositions are together in the air, they can come into contact. They can react, condensate or coagulate, and in this case they will form new particles of these two types and their chemical composition will become mixed. This new aerosol mixture is called *internally* mixed aerosols where individual aerosol particles will be composed of different aerosol types. In the case when two different types of aerosols do not react between each other, they would stay only mixed in the air, but the individual particles are still independent and composed of only one chemical component. In that case we have *externally* mixed aerosols. In the atmosphere, aerosols are externally mixed usually only just after being emitted. Aerosols staying longer in the atmosphere will have a higher probability to react with other aerosols or gases, and change their composition. Related to the whole population, aged aerosols are usually well internally mixed.

1.2 Aerosol sources

On the global scale, natural emissions of aerosols are larger than anthropogenic. Anthropogenic emissions contribute in mass with 10 % to 50 % depending on the

Table 1.1: Estimated ranges for global annual aerosols emissions. Data come from Lamarque et al. (2010), IPCC (2013) and Lewis and Schwartz (2004b). DD stands for desert dust, SS for sea salt, BC for black carbon, and OC for organic carbon

	DD	SS	BC	OC
Emissions [$\mathrm{Tg\,yr^{-1}}$]	1000-4000	1000-30000	7-15	25-90

region (Hinds, 1999; Dentener et al., 2006). The majority of the natural sources is well distributed around the globe, with emissions coming from huge surfaces like oceans or deserts. Anthropogenic sources, on contrary, are concentrated in the smaller regions, around high density of industries and high human activities. Often, in these concentrated regions, anthropogenic sources can be larger than natural sources. Using models, satellite observations, and the knowledge about human and fire activities, one can distinguish anthropogenic from natural emissions with a good confidence (King et al., 1999; Takemura et al., 2002; Kaufman et al., 2002). An estimation of the global annual emissions of the different aerosol types is given in Table 1.1.

Sources of primary aerosols Primary aerosols have natural and anthropogenic sources. Primary aerosols that are emitted naturally are desert dust particles emitted from deserts and arid areas, sea salt particles from seawater surfaces, biomass-burning aerosols from fires, and primary biogenic aerosols from living organisms.

Also, different human and industrial activities lead to the emission of primary anthropogenic aerosols: ground and air transportation (fuel combustion), fossil fuel (coal, petroleum, natural gas) combustion, waste incineration, cement manufacturing, metallurgy, fires, etc. During these activities various anthropogenic aerosols are directly emitted into the atmosphere: industrial dust, black carbon (BC), soot, organic carbon (OC).

Sources of secondary aerosols Sources of secondary aerosols follow the same separation. Precursors of secondary aerosols that are emitted naturally are: terpenes and other biogenic volatile organic compounds (VOC) from the living organisms which are precursors of secondary organic aerosols (SOA); dimethyl sulphide (DMS, C_2H_6S) and hydrogen sulphide (H_2S) from oceans, soils and land organisms and sulphur dioxide (SO_2) from volcanoes and biomass-burning which are precursors of sulphate aerosols; mono-nitrogen oxides (NO_X) from lightning, natural soils and biomass-burning and ammonia (NH_3) from natural soils and biomass-burning which are the precursors of nitrate aerosols.

The anthropogenic sources are similar as for the primary aerosols with the addition of agriculture and cattle breeding as important additional sources. The precursors are: anthropogenic volatile organic compounds that leads to formation of secondary organic aerosols; SO_2 from fuel fosil fuel combustion, aircrafts and industry that is the gas precursor of sulphate aerosols; NO_X from fossil fuel combustion, aircrafts, agriculture, and NH_3 from biomass-burning, agriculture, livestock and wild animals, industry, fossil fuel combustion and humans as the precursors of nitrate aerosols.

1.3 Aerosol processes

Nucleation During nucleation gas molecules aggregate and form a cluster that can condense in a small liquid particle. If the particle grows to the critical size, it will become stable and it can continue growing by condensation (Jacobson, 2005). *Homogeneous* nucleation occurs in the case when vapour molecules nucleate on their own molecules. In *heterogeneous* nucleation, the nucleation happens on the surface of foreign particles. In the case of *homomolecular* nucleation, only one chemical species is involved. And, in *heteromolecular* nucleation, two or more gas species.

Homomolecular nucleation will happen only if the species in the vapour phase are in supersaturation. Unsaturated or saturated vapour can become supersaturated in the case of fast changes, like adiabatic expansion, cooling at the same pressure, compression at the constant temperature. In the case of heteromolecular nucleation, when more gas species nucleate together, the condition of supersaturation has to be achieved for the solution of these species together, and not for the species separately (Seinfeld and Pandis, 1998).

Condensation When a lot of nucleated particles are formed and supersaturation becomes low, condensation takes place instead of nucleation. There is no further formation of new particles. Instead, already existing particles start to grow. Condensation, together with coagulation, is efficient in moving particles to larger sizes. The process of condensation, and the reverse process, evaporation, are characterized by the continuous tendency of species to achieve the equilibrium. If the vapour phase of a gas species is supersaturated, it will condensate. If the vapour phase becomes unsaturated, the aerosol will start to evaporate.

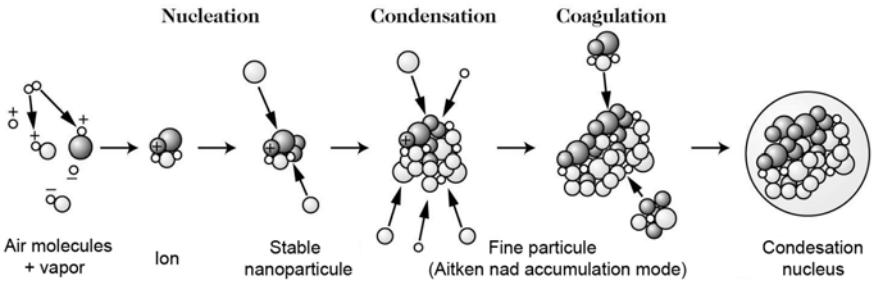

Figure 1.2: A possible path of evolution of a particle, from nucleation, to condensation and coagulation, to finish as a condensation nucleus for a cloud particle. The figure is from Delmas et al. (2005).

Coagulation Coagulation is a process when two aerosol particles come in contact, collide and stick together. Collision can happen due to Brownian motion of the particles, or phoretic, electrical, gravitational or other forces. In this process, externally mixed particles become internally mixed. A number of small particles are lost due to the formation of larger particles, but the volume of particles is preserved. Coagulation can be important for aerosol-cloud interactions, and it influences the lifetime and evolution of particles. The process and its rate depends on many factors: Brownian motion, shapes of particles, small turbulences and eddies, differences in sedimentation speeds, viscosity, van der Waals forces, phoretic forces, electric charges of particles (Jacobson, 2005).

1.4 Aerosol types

1.4.1 Black carbon

Black carbon (BC) particles are a type of carbonaceous aerosols formed during combustion of fuels and biomasses (Fig. 1.3). Particles are composed of small spherical carbon particles of the size of 10 nm to 50 nm combined together into a chain-like aggregate composed of hundreds of members (Wentzel et al., 2003; Li et al., 2003). They are composed almost entirely of carbon, with small amounts of other elements (Andreae and Gelencsér, 2006). After emission, aggregates collapse into more compact clusters, and then water, gases and

other particles can condensate and coagulate on the black carbon particles (Ramachandran and Reist, 1995). Black carbon is the material insoluble in water, chemical components of other aerosols, and organic solvents (Fung, 1990). This means that it remains chemically unchanged in the contact with other aerosols. Because of this low chemical reactivity, when it is mixed with other aerosols and materials, interactions happen on their surfaces which produces internally mixed aerosol particles. Often it can be found as a core of more complex, aged aerosols with a shell made of other materials. Such modified aerosols are called *black carbon containing particles*. Black carbon mixes with other aerosols and substances in a time range of 1 to 5 days, and this process has an important effect on optical properties of black carbon aerosols (Jacobson, 2000, 2001).

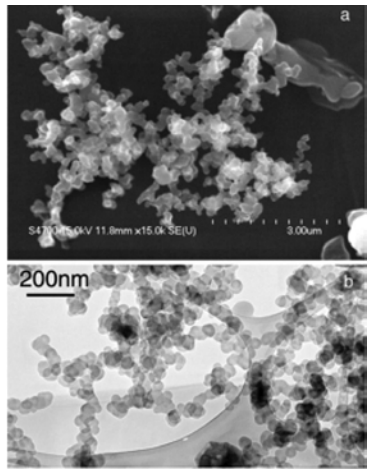

Figure 1.3: Scanning electron microscope image of BC aggregates in young smoke in South Africa, collected on 20 August 2000 **(a)**; (b) transmission electron microscope (TEM) image of chain-like BC aggregates in flaming smoke in Zambia, collected on 5 September 2000 **(b)**. From Bond et al. (2013).

Size The black carbon aerosols are mostly in the sub-micron range, which compared to other aerosol types makes them one of smaller aerosols. Their size range depends on the emission source, the size at emission and later ageing. Urban black carbon and biomass-burning black carbon aerosols have different sizes. In urban environment, freshly emitted black carbon aerosols have number median diameters in the range of 50 nm to 80 nm, while black carbon aerosols created in fires have number median diameters of ≈120 nm (Kondo et al., 2011). The size of the emitted particles depend on the combustion temperature and fuel burning efficiency; this make biomass-burning BC particles bigger than engine combustion BC particles, while jet aircraft BC particles are the smallest

in size (Petzold et al., 2005; Bond et al., 2013).

Evolution Because of its small size, the gravitational settling is not important for black carbon. Also, because of its low chemical reactivity, chemical reactions do not remove it from the atmosphere. The main removal mechanisms for black carbon aerosols are wet deposition by precipitation and dry deposition by surface. More aged black carbon aerosols, mixed and coated with other aerosols and materials, have different physical and optical properties. They become more hydrophilic, which produce more interactions with clouds, make them more often cloud condensation nuclei and ice nuclei. These processes modify their removal rate from the atmosphere by shortening their lifetime (Stier et al., 2006).

Optical properties The main particularity of the black carbon is a strong absorption of visible light in the whole range of the visible spectrum. There is no strong spectral dependency of its absorption capabilities in the visible spectrum, because of the material properties, but also because of its small size which is much smaller than the wavelength of the visible radiation. Black carbon is the most efficient absorber of the visible radiation in the atmosphere (Bond et al., 2013). This is the reason why there are many studies of its influence on the radiative budget of the atmosphere, and on its direct radiative forcing. The ageing of black carbon aerosols, change its absorption capabilities (Schnaiter et al., 2005; Lack and Cappa, 2010). Laboratory and modelling studies show that mixing can enhance its absorption capabilities by the lensing effect of the coated material or decrease it if the coating is too thick. As the result, the ageing of black carbon and BC-containing aerosols has to be taken in account in radiative transfer studies.

1.4.2 Primary organic carbon

Organic carbon (OC) refers to the type of carbonaceous aerosols which are composed of carbon-containing compounds that besides carbon contain also hydrogen and oxygen, and possibly other elements. Their composition is about 85 % to 95 % of C, 3 % to 8 % of O, and 1 % to 3 % of H (Ebert, 1990). Strictly speaking, organic carbon refers only to carbon content in the organic material (Seinfeld and Pandis, 1998), and this is the quantity that is commonly measured. Yet, often in the literature the term organic carbon also refers to organic aerosols, or the complete mass of the organic material of the particle.

That is the quantity that is relevant in the climate effect studies. In this manuscript the term organic carbon will be used in this sense.

Organic carbon aerosols are released from the same processes as black carbon particles or made in-situ by condensation of hydrocarbons. This means that they are both primary and secondary aerosols. Primary organic carbon aerosols are formed in combustion processes together with black carbon particles. Studies (Pósfai et al., 2003; Mallet et al., 2004) show that during combustion processes black carbon and organic carbon are usually emitted in different particles. Their relative quantity depends on sources and the type and properties (like temperature for example) of combustion. In sources up to 50 % could be organic carbon (Medalia and Rivin, 1982).

Optical properties Organic carbon is a less absorbing material than black carbon. At 550 nm its mass absorption efficiency is considerably smaller. However, organic carbon shows an important spectral dependency in its absorption efficiency in the visible spectrum (Kirchstetter et al., 2004; Chakrabarty et al., 2010). At lower wavelengths the absorption efficiency increases, and in the UV region it could have a significant effect (Hoffer et al., 2006). This absorption spectral dependency is used to distinguish organic carbon particles from black carbon particles (Wonaschütz et al., 2009). Sizes of organic carbon aerosols from combustion processes are similar as for black carbon (Bond et al., 2013).

1.4.3 Sea salt

The sea salt aerosol is a type of aerosol that consists of sea salt and seawater droplets. Oceans cover about 70 % of the planet's surface, and sea salt aerosols are the most emitted aerosols in the atmosphere (IPCC, 2001). They are formed at sea surface by different mechanisms which strongly depend on wind speed. All of them work by injecting sea spray droplets into the air. The main two mechanisms depend on whitecaps, which form when wind blows and form waves that break on the water surface. The breaking of the waves entrains the air into water, which forms bubbles and foamy patches on the seawater surface — whitecaps. This process happens at wind speeds above $3\,\mathrm{m\,s^{-1}}$–$4\,\mathrm{m\,s^{-1}}$ (Monahan and O'Muircheartaigh, 1986). In the first mechanism, when a whitecap bubble bursts, it injects several hundreds of sea spray droplets called film drops into the air (Fig. 1.4a). Film drops are sub-micron droplets released from the film cap which separates the air in bubble from the air outside (O'Dowd et al., 1997). The second mechanism is just the continuation of the

first mechanism. By filling the cavity created by the bursting of the bubble, water injects drops in the air (Fig. 1.4b). These drops can break into as much as ten separate particles and they are called jet drops (Fig. 1.4c). Jet drops have a typical size of about 2 µm to 5 µm, but can be significantly larger (O'Dowd et al., 1997). The third mechanism directly injects sea spray drops into the air. At wind speeds above $7\,\mathrm{m\,s^{-1}}$–$9\,\mathrm{m\,s^{-1}}$, wind can tear so-called spume drops from wave crests and inject them directly into the air (O'Dowd et al., 1997). The size of spume drops is large, with diameters from about 20 µm to several hundreds of micrometers (Andreas, 2002). These particles, if they do not split into smaller drops, fall back into the sea very quickly by sedimentation.

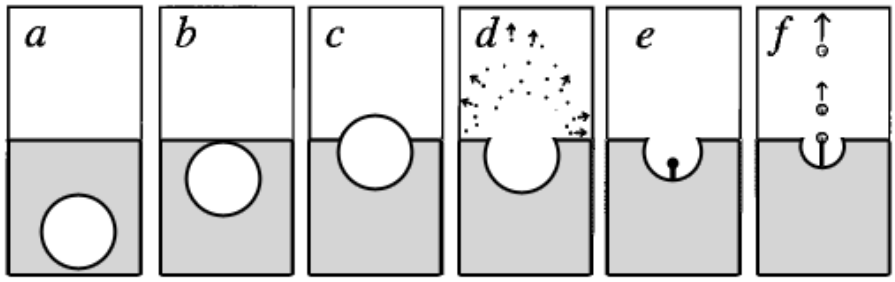

Figure 1.4: Drop formation from bubble bursting. **(a-c)** Bubble rises to the surface and forms the film or cap. **(d)** When the film bursts, film drops are ejected. **(e)** The cavity shrinks and a jet rises from its center. **(f)** The cavity gets filled by the surrounding water, and several jet drops form from breakup of the jet. The figure is from Lewis and Schwartz (2004b).

Often the air is less humid than droplets, and after injection into the air, sea spray droplets start to evaporate. They try to achieve an equilibrium with ambient relative humidity, and they shrink in size during this process. Many of sea spray droplets before reaching moisture equilibrium fall back into the sea due to gravitational settling. This is especially true for the largest spume drops. Turbulent mixing near the sea surface has an important role in pushing the large particles vertically where they could achieve moisture equilibrium. But in general, only drops with diameters smaller than ≈40 µm get into moisture equilibrium and turn into sea salt aerosols (Andreas, 2002).

Composition At the time of its emission, the composition of sea salt aerosols is very similar as of the seawater in respect to dissolved material in it, but not necessarily the same because of possible fractionation in the conversion of seawater to aerosols (Duce et al., 1972). The composition of sea salt covers a wide range of elements and compounds. It can be quite complex and it could vary significantly in time and space. Sodium chloride (NaCl) comprises

85 % of all salts in seawater, and it is the most important component in sea salt aerosols. Other chlorides ($MgCl_2$, KCl) and sulphates (Na_2SO_4, $MgSO_4$, $CaSO_4$) are also significant components of sea salt aerosols (Tang et al., 1997; Lewis and Schwartz, 2004b). Moreover, in freshly emitted sea salt aerosols one can find organic matter (O'Dowd et al., 2004), and even bacteria (Blanchard, 1983).

Hygroscopicity Sea salt aerosols are hygroscopic particles. To be in a moisture equilibrium with the ambient air, sea salt aerosols absorb water from the air and grow in size as relative humidity increases in the air (or lose water by evaporation and shrink in size in the case when relative humidity drops). The hygroscopic growth of sea salt aerosols change their physical (size, density...) and optical properties. For example, a sea salt aerosol particle at relative humidity of 80 % is about twice as big as the dry particle. The size range of sea salt particles covers from 0.05 µm to 10 µm; particles >10 µm have very short lifetimes if they are not split into smaller components.

Because of their hygroscopicity, sea salt particles are very efficient cloud condensation nuclei (CCN). Even with small concentrations, they are capable of modifying rain drop production and cloud albedo significantly (O'Dowd et al., 1997; Feingold et al., 1999). The sea salt aerosols that are the most efficient as CCN are particles with diameters of <1 µm (O'Dowd et al., 1997).

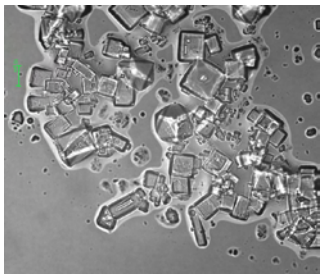

Figure 1.5: A microscope image of sea salt particles. Image courtesy of Chere Petty, University of Maryland.

Chemical evolution Because of its complex chemical composition, sea salt aerosols have a strong chemical evolution. For example, the ratios between chlorine, sodium, and sulphur observed in many sea salt aerosol particles differ from those in the normal seawater (McInnes et al., 1994). There are many possible reactions that occur inside the sea salt aerosols. When they are in the form of aqueous solution, they are a good niche for chemical reactions. In the air, sea salt aerosols go through the process of acidification. They react with

acids from the atmosphere, like sulphuric acid (H_2SO_4), nitric acid (HNO_3), or hydrochloric acid (HCl), which dissolve, dissociate and release hydrogen ions H^+. Then, these positive ions react with available negative ions and acidify the particles (Keene et al., 1998).

As sea salt aerosols become more acidic, the process of removing of chlorine from the particles, or dechlorination, starts. Nitric (HNO_3) and sulphuric (H_2SO_4) acids from the air react with sea salts, make hydrochloric acid (HCl) which is released to the atmosphere, leaving the particles enriched in nitrates and sulphates, and depleted in chlorine (Eriksson, 1959). The chlorine deficit can be also provoked by other reactions (Duce, 1969; Chameides and Stelson, 1992). Similarly, bromide (HBr) also gets removed from the sea salt aerosols (Mozurkewich, 1995). These changes with time enrich sea salt aerosols with nitrates and non sea- salt sulphates (sulphates in excess compared to freshly emitted aerosols). But, aged sea salt aerosols are still soluble and represent good CCN particles. Described chemical changes are more pronounced near to the coasts because of the big influence of polluted, continental air masses rich in acids. In remote regions of the oceans, the concentration of acids is smaller, which reflects on the small rate of sea salt aerosols chemical transformations (Murphy et al., 1998b).

Optical properties Sea salt aerosols absorb little radiation (Tang et al., 1997; Smirnov et al., 2002; Dubovik et al., 2002). They interact with radiation primary by scattering. Over the regions far from coasts and pollution, sea salt dominate aerosol scattering by 95 %, and the submicron particles have the biggest effect (Murphy et al., 1998a). The hygroscopic growth changes sea salt aerosol scattering properties due to their change of size and composition. The chemical evolution changes also the scattering properties: particles with less Cl scatter radiation less efficiently, but this effect is smaller than the influence of relative humidity (Tang et al., 1997).

1.4.4 Desert dust

Desert dust is a source of mineral aerosols and it is one of the largest source of aerosols in general. Mineral aerosol sources are mainly in deserts and dry lake beds, and sometimes in the literature the terms mineral aerosols and desert dust aerosols are equated. The other sources of mineral aerosols are semi-arid soils, regions with reduced vegetation or soils disturbed by human activities, pastures, traffic, construction.

Composition Desert dust aerosols are eroded soils and therefore, their composition resembles to the composition of the crust and crustal rock. They are composed of various oxides and carbonates, where silicon and aluminium oxides dominate. Studies show that in general about 60 % is silicon dioxide (SiO_2), and about 10 %–15 % aluminium oxide (Al_2O_3). Other abundant oxides are iron(III) oxide (Fe_2O_3), magnesium oxide (MgO) and calcium oxide (CaO), but their percentages can vary depending on the source location (Goudie and Middleton, 2001). The elemental composition of desert dust is quite constant around the globe (Gomes and Gillette, 1993), but still, in the particles we can find a variety of different types of minerals (Usher et al., 2003). Dust from desert is mainly from sand, and dust from other arid and semi-arid areas contain more silt and clay (Alfaro et al., 1998).

Formation Desert dust aerosols are, as already mentioned, mechanically disintegrated parts of soils. Its formation and emission highly depend on the wind, but also on the physical properties and conditions of soils. Particles are rarely injected into the air directly by the wind, but instead mainly by an intermediary process — saltation. When the wind achieves a high enough velocity, it starts to move particles. This wind velocity is called erosion threshold velocity, and it depends on the size of the particle which starts to move. It depends also on the soil properties: how rough it is, or how moist, or how strong are cohesive forces between particles in the soil, or is there any vegetation present. It the vegetation is present it is important to which degree is present and which type. When the erosion wind velocity is reached, particles start to move horizontally in the layer close to the surface – saltate. The biggest particles do not bounce because of their weight and gravitational drag, but just roll on the surface. Saltating particles when impacting the surface can break down, or break down the soil aggregates which they hit. The impact could provide enough of energy to break the cohesive forces which link the soil particles and release small particles of dust into the air (Marticorena et al., 1997; Laurent et al., 2006). Typically, particles 100 µm–500 µm can saltate and their size distribution resembles closely to the size distribution of the soil particles (Kok et al., 2012); particles that are sandblasted are generally 0.1 µm–10 µm in diameter (Alfaro et al., 1998).

Transport The particularity of desert dust aerosols is that they are often transported over very long distances, commonly over several thousands of kilometres (Prospero, 1999), or even more (Uno et al., 2009). They often travel in high, horizontally layered plumes. The transatlantic path of the African dust from Sahara is a well studied topic (Prospero et al., 1981; Prospero, 1999; Perry et al., 1997), and it has also a biochemical significance, because

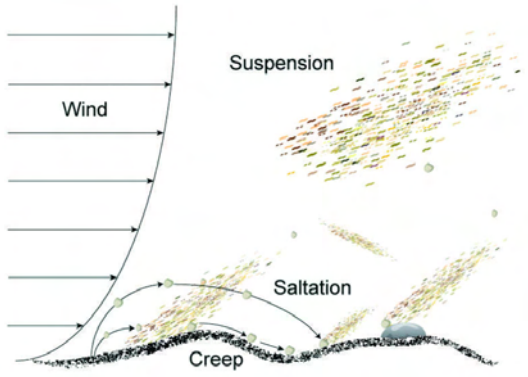

Figure 1.6: Illustration of the process emission of desert dust particles. The biggest particles roll (creep) over the surface under the influence of the wind. Smaller particles rise and then fall (saltate) hitting the soil aggregates and eject fine particles in the air. Figure from Hatfield and Sauer (2011).

desert dust has a fertilizing role for oceanic and continental ecosystems by providing micro nutrients, like phosphorus (P) or iron (Fe). This impact seams to be especially important for the Amazon forest and the equatorial Atlantic Ocean (Bristow et al., 2010). Besides westward transatlantic path to North and South America, Saharan dust is also transported northward to western Mediterranean sea (Guieu et al., 2002) and Europe up to Scandinavia (Franzen et al., 1995), or eastward towards eastern Mediterranean (Levin et al., 1996) and Middle East (Alpert and Ganor, 2001). Asian dust, whose the most importante source is the desert Gobi, usually takes eastern path and there are common examples of its detection on Hawai (Parrington et al., 1983) or continental United States (Duncan Fairlie et al., 2007). Also, the dust transport depends on the meteorological and climatic conditions. The transatlantic transport is increased during El Niño period (Prospero and Nees, 1986), and the transport to Mediterranean is correlated with the phase of the North Atlantic Oscillation (Moulin et al., 1997).

Ageing The chemical reactivity of desert dust aerosols is very low, but they can provide a surface for chemical reactions. Interactions with reactive inorganic or organic chemical species or aerosols can change the chemical nature of the desert dust particles. They can change and alter their surface or make a coating on it which can enhance or prevent other reactions. For example, desert dusts often get covered with sulphates or other soluble materials. Levin et al. (1996) found that this process is frequent and is probably surface dependant, because the amount of coating is bigger on bigger particles.

Desert dust aerosols are insoluble and hydrophobic, but they are efficient ice nuclei (IN) particles (DeMott et al., 2003; Field et al., 2006). The ageing process, where they may interact with other aerosols and compounds and become coated with hydrophilic material, could turn them also in good cloud condensation nuclei (CCN) without limiting their efficiency as IN particles (Levin et al., 2005). There are also other interactions that desert dust aerosols have on clouds (for example Karydis et al., 2011).

Optical properties Desert dust aerosols are efficient in the scattering of the radiation. Also, although less efficient than black carbon, they absorb the radiation. But because of its high mass abundance (especially compared to BC), its global total absorption is significant (Sokolik and Toon, 1996). Also, because of its abundance, it can largely influence the visibility.

1.4.5 Volcanic aerosols

Volcanic eruption can eject into the atmosphere very large quantities of aerosols. Primary volcanic aerosols are in the form of ash. Volcanic ash particles are made up of fragments of volcanic rock which are composited of rock mineral, SiO_2, and glass and have diameters smaller then 2 mm. The rock is usually of silicate type: rhyolite, andesite and basalt, and its composition is important to know because of different optical properties (Gangale et al., 2010). Volcanic ash clouds can be injected at high altitudes and transported over very long distances in layered plumes. They can have an important impact on civil airplane flight safety (Prata and Tupper, 2009).

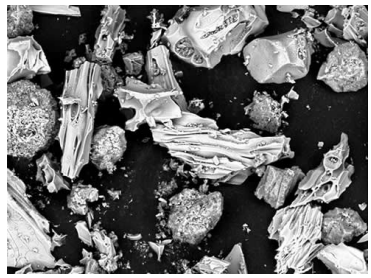

Figure 1.7: Microscopic particles of volcanic ash collected in Italy. Image courtesy of Suzanne MacLachlan/BOSCORF/National Oceanography Centre, UK.

Volcanoes can also eject large quantities of SO_2 that can form secondary sulphate aerosols with very long lifetimes in the stratosphere. These stratospheric aerosols can have important effects on climate and ozone depletion.

1.4.6 Secondary aerosols

Secondary aerosols are not emitted directly from natural or anthropogenic sources, but formed in the air by chemical reactions and transformations. Because of the additional complexity brought by different pathways and processes of formation, the concentrations of the secondary aerosols in the air do not depend linearly on primary aerosol concentrations and precursor concentrations (Seinfeld and Pandis, 1998).

A lot of aerosols are formed from the gas phase. Volatile and condensable gases can form less volatile substances that can convert to aerosols through aerosol nucleation and condensation (Seinfeld and Pandis, 1998).

1.4.6.1 Secondary inorganic aerosols

Secondary inorganic aerosols can be in aqueous and solid states. The solid aerosols are usually in the form of salts. In aqueous form, compounds can be also in the form of ions. The most frequently encountered secondary aerosols are sulphates, nitrates, ammonium. They are formed from precursors by a chain of chemical reactions. The majority of secondary inorganic aerosols in the present-day atmosphere comes from anthropogenic sources (IPCC, 2013).

These aerosols are hydrophilic and they grow when exposed to water. The growth starts (and aerosols pass from the crystalline state to hydrated state) above a certain threshold relative humidity which differs for each species (or solution in the case of the solution of more aerosols species) and also depends on the air temperature. In the reverse direction, as the relative humidity decreases, the water evaporates and the aerosol shrinks. But, also the hysteresis behaviour appears. The crystallisation happens only when the solution of aerosol in water reaches supersaturation, which happens mostly at the relative humidity which is lower than the threshold relative humidity for the growth (Seinfeld and Pandis, 1998). For example, ammonium-sulphate changes from the crystalline state to the hydrated state at the relative humidity of ≈80 %, but returns to the crystalline state at the relative humidity of ≈40 % Taylor et al. (2011). Because of their hygroscopicity, secondary inorganic aerosols make good cloud-condensation nuclei (IPCC, 2013).

The most common sulphate aerosols are pure sulphuric acid (H_2SO_4) which is the most abundant in very acidic conditions, ammonium hydrogen sulphate ($(NH_4)HSO_4$) in acidic conditions, and ammonium sulphate ($(NH_4)_2SO_4$)

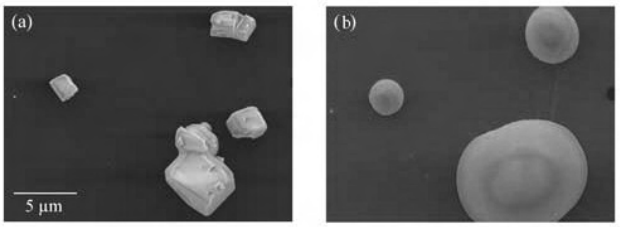

Figure 1.8: Scanning electron microscope images of calcium carbonate particles before and after reaction with nitric acid where the particles convert to calcium nitrate. Images are from Krueger et al. (2003).

which is the preferred form of sulphates if enough of ammonia neutralizes sulphuric acid. The most common nitrate aerosol is ammonium nitrate (NH_4NO_3) which prefers low sulphate concentrations. In aqueous phase all these compounds can be also in the form of ions: NH_4^+, H^+, HSO_4^-, SO_4^{2-}, NO_3^-.

Optical properties of secondary inorganic aerosols are well studied and they show very low amount of absorption in the visible part of spectrum. But, they can enhance the absorbing properties of primary aerosols when coated on their surfaces.

1.4.6.2 Secondary organic aerosols

Secondary organic aerosols (SOA) are formed from volatile organic compounds (VOC) as precursors. VOCs are a class of organic pollutants that are volatile at ambient air conditions. In the atmosphere there are tens of thousands of different organic compounds (Goldstein and Galbally, 2007). The capability of VOCs to form SOA depends on their atmospheric concentration, chemical reactivity, and volatility of their products (Hallquist et al., 2009), but many formation processes still remain uncertain. VOCs as gases can undergo a wide range of oxidation reactions by species such as the hydroxyl radical (OH) in the day-time, ozone (O_3) or nitrate radical (NO_3) in the night-time (Seinfeld and Pandis, 1998). VOC oxidation can lead to the production of organic oxidized products (Goldstein and Galbally, 2007) which can condensate into aerosol phase onto other aerosols.

The products that have a low vapour pressure or that are more highly soluble can partition from gas phase to particulate phase and form secondary organic aerosols. If one oxidized species is above its saturation concentration

(supersaturated), it will transform into aerosol phase by condensating onto any available aerosol particle and lead to formation of secondary organic aerosols.

The biggest source of VOCs are biogenic sources (Hallquist et al., 2009). Biogenic emissions come usually from terrestrial ecosystems: plant chloroplasts, plant specialized tissues, microbes, and animals. Also, dimethylsulfide from the oceans is an important biogenic VOC source (Guenther et al., 1995). Other VOC sources are biomass burning, fossil fuel combustion and industrial activities.

1.5 Size distribution

Aerosols cover a very wide range of sizes. Therefore, it is necessary to develop a way to represent what sizes of particles are present in which quantities. This will help to distinguish aerosols and different effects that are size-dependent. This representation is called *particle size distribution*.

1.5.1 Representations of size distributions

Discrete distribution Particle size distribution can be defined in different ways. One way is to describe it as a discrete distribution that is divided in discrete size ranges – bins. Bins have to be chosen in the way to cover all possible sizes of a group of particles of interest. The size distribution counts the amount (relative or absolute) of particles in each bin. The discrete distribution can be represented in the form of a histogram (Fig. 1.9).

If bins do not have equal widths, it is important that the size distribution is normalized by dividing the quantity of particles in each bin by its width. In that way, each bin will count the quantity of particles per unit of size, for example, number of particles per micrometer.

Continuous distribution The number of particles of a certain size can also be described by a mathematical function. In this case, we talk about continuous particle size distributions. This approach allows us to calculate the number of particles of any size at a given point, at a given time by the function.

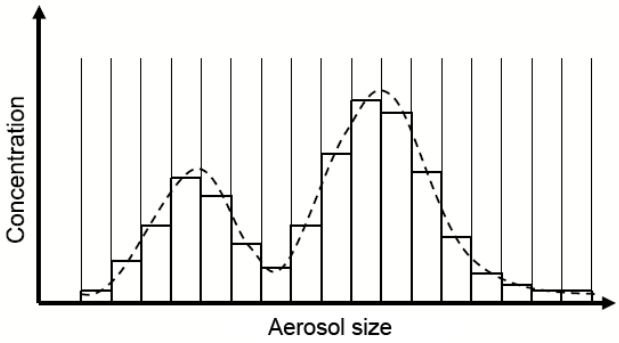

Figure 1.9: Illustration of continuous and discrete aerosol size distribution. The figure is adapted from Boucher (2012).

1.5.2 Number and mass distributions

The size distribution is most often described as the number distribution, i.e. defined in terms of the particle number concentration. But it could be described in terms of many other quantities, such as particle mass, surface, volume, velocity, etc. If a certain particle property depends on some quantity other than the number of particles, like for example the particle surface area or volume, it is useful to represent the size distribution as the surface area distribution, or volume distribution.

Number distribution The particle size distribution $f_N(D_p)$ is defined in the way that $f_N(D_p)dD_p$ presents the number concentration of particles in the diameter size range $[D_p, D_p + dD_p]$. The total number concentration of particles over all sizes n_{tot} is obtained by integrating $f(D_p)$ over the whole size domain:

$$n_{tot} = \int_0^\infty f_N(D_p)dD_p \tag{1.1}$$

where D_p is the diameter of the particle.

The particle size distribution $f_N(D_p)$ is expressed in [#particles/m^3/m], and n_{tot} in [#particles/m^3]. If we are only interested in the number concentration between two specific diameters D_{p_1} and D_{p_2} (i.e. truncated concentration), then the limits in the integral in Eq. (1.1) will be the sizes D_{p_1} and D_{p_2}:

$$n(D_{p_1}, D_{p_2}) = \int_{D_{p_1}}^{D_{p_2}} f_N(D_p)dD_p \tag{1.2}$$

Mass distribution The mass distribution $f_m(D_p)$, as an example of distributions other than the number distribution, is defined in the way that $f_m(D_p)dD_p$ presents the mass of particles per unit of volume of air in the size range $[D_p, D_p + dD_p]$. In the case of the spherical particles, it is connected with the number distribution $f_N(D_p)$ by

$$f_m(D_p) = \rho_p V_p f_N(D_p) = \rho_p \frac{\pi}{6} D_p^3 f_N(D_p) \tag{1.3}$$

where V_p is the volume of the particle, and ρ_p is the density of the particle $[\mathrm{kg\,m^{-2}}]$.

Then, if the ρ_p is considered constant for all particle sizes, the total mass of aerosols per unit of volume of air will be:

$$\begin{aligned} m_{tot} &= \int_0^\infty f_m(D_p) dD_p \\ &= \rho_p \frac{\pi}{6} \int_0^\infty D_p^3 f_N(D_p) dD_p \end{aligned} \tag{1.4}$$

The mass particle size distribution $f_m(D_p)$ is expressed in $[\mathrm{kg\,m^{-3}\,m^{-1}}]$, and m_{tot} in $[\mathrm{kg\,m^{-3}}]$.

1.5.3 Size distribution properties

Most of the distribution functions are described by two parameters: one that identifies the center of the distribution or a characteristic location on it, and the other one that defines the width or spread of the distribution. To define the center of the distribution it is possible to use different kinds of *averages*. That could be the mean, the median or the mode of different distributions (mass, number, volume, etc.). Below, we define these values for the number distribution.

Mean The mean is the arithmetic average. If used for the number distribution, the mean is the sum of all particle sizes divided by the total number of particles. With the discrete distribution of a group of N particles with S size ranges and with midpoint particle diameters D_{p_i} and number concentrations n_{p_i}, the mean is:

$$\overline{D}_p = \frac{\sum_{i=0}^{S} n_{p_i} D_{p_i}}{\sum_{i=0}^{S} n_{p_i}} = \frac{1}{n_{tot}} \sum_{i=0}^{S} n_{p_i} D_{p_i} \tag{1.5}$$

If the size distribution is in continuous form, the mean becomes

$$\overline{D}_p = \frac{\int_0^\infty D_p f_N(D_p) dD_p}{\int_0^\infty f_N(D_p) dD_p} = \frac{1}{n_{tot}} \int_0^\infty D_p f_N(D_p) dD_p \tag{1.6}$$

Mode The mode represents the most frequent size in the distribution, and corresponds to the highest point on the distribution (Fig. 1.10). It can be determined by setting the derivative of the continuous distribution function to zero:

$$D_{mode} = \left(\frac{df_N(D_p)}{dD_p} \right) = 0 \tag{1.7}$$

Median The median of the number distribution corresponds to the diameter for which one-half of the particles are smaller, and one-half of the particles are bigger. In the continuous size distribution it can be defined as

$$\int_0^{D_{med}} df_N(D_p) dD_p = \frac{1}{2} n_{tot} \tag{1.8}$$

Geometric mean Geometric mean of sizes of n_{tot} particles is $n_{tot}{}^{\text{th}}$ root of the products of all N_{tot} sizes:

$$D_{gm} = (D_1 D_2 D_3 \dots D_{N_{tot}})^{1/N_{tot}} \tag{1.9}$$

For the case of discrete distribution with S size ranges:

$$D_{gm} = (D_1^{N_1} D_2^{N_2} D_3^{N_3} \dots D_S^{N_S})^{1/N_{tot}} \tag{1.10}$$

where $N_1, N_2, N_3 \dots N_S$ are the number of particles in the respective size ranges, and $D_1, D_2, D_3 \dots D_S$ are the characteristic particle diameters (midpoints) in the respective size ranges. This equation can be expressed in logarithmic form:

$$\ln D_{gm} = \frac{1}{n_{tot}} \sum_{i=0}^{S} n_{p_i} \ln D_{p_i} \tag{1.11}$$

The value of the geometric mean for polydisperse aerosols is smaller than the value of the mean, $D_g < \overline{D}_p$.

Variance The variance measures the deviation of distribution from its mean diameter $\overline{D}_p$. It is defined as

$$\sigma^2 = \frac{\sum_{i=0}^{S} n_{p_i}(D_{p_i} - \overline{D}_p)^2}{\sum_{i=0}^{S} n_{p_i}} = \frac{1}{n_{tot}} \sum_{i=0}^{S} n_{p_i}(D_{p_i} - \overline{D}_p)^2 \tag{1.12}$$

For the continuous distribution the variance becomes

$$\begin{aligned} \sigma^2 &= \frac{\int_0^\infty (D_p - \overline{D}_p)^2 f_N(D_p) dD_p}{\int_0^\infty f_N(D_p) dD_p} \\ &= \frac{1}{n_{tot}} \int_0^\infty (D_p - \overline{D}_p)^2 f_N(D_p) dD_p \end{aligned} \tag{1.13}$$

The variance that equals to zero means that aerosols are monodisperse. As the variance increases, it indicates that the spread of the distribution around its mean size increases. The unit of the variance is [m^2].

Standard deviation The standard deviation is defined as the square root of the variance:

$$\sigma = \sqrt{\sigma^2} \tag{1.14}$$

The unit of the standard deviation is [m]. In the literature it is often the standard deviation that is used as the parameter which describes the spread of the distribution.

Averages among different distributions Mean values, as the mean, mode, median and geometric mean, are defined for other distributions (surface, volume, mass, etc.) in the similar way as for the number distribution. But, their values will not be the same between distributions. The number mean diameter and the mass mean diameter will be different. The same applies for other averages, and for other distributions (volume, surface, etc.). This is shown in Fig. 1.10, where the different averages are presented for the log-normal distribution.

1.5.4 Types of aerosol size distributions

The ambient aerosol size distributions can vary significantly. And to represent a size distribution accurately, it is necessary to measure it, and then to represent it in the discrete form as a table with values for many measured diameters. Still, it is more convenient to represent it in the continuous form, describing it

with a mathematical function which match as good as possible the measured distribution. Many functions are proposed and used for this application: normal, power-law, exponential, log-normal distribution, etc. Here we will describe the normal and log-normal distributions.

Normal distribution The normal distribution is one of the most used statistical distributions, and it can be also used to describe the distribution of the aerosol particles in the atmosphere. The normal number distribution is defined as

$$n_n(D_p) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{D_p - \overline{D_p}}{\sigma}\right)^2\right] \tag{1.15}$$

This distribution is symmetrical, and the mean, mode, median values fall exactly in the same place and have the same value. The standard deviation σ quantifies the spread of the function, and 68% of particles will fall in the range $\overline{D_p} \pm \sigma$. To represent the distribution of the aerosol particles, the normal distribution has certain disadvantages. With bigger values of the standard deviation, the normal distribution is wide, and it can show negative values on its long tail, which is physically impossible for the aerosol particles. Another inconvenience is its symmetricity. Most of aerosols in the air show a distribution that has a longer tail at larger sizes. Only some spores, pollens, and laboratory prepared aerosols can be accurately be presented by the normal distribution (Hinds, 1999).

1.5.5 Log-normal distribution

The log-normal distribution is the most frequently used distribution in aerosols studies. Any quantity is log-normally distributed if its logarithm is normally distributed. When the particle size is log-normally distributed, logarithmic scale of the x axis will reveal a characteristic bell shape of the normal distribution. With the x axis in the linear scale, the log-normal distribution show that it is skewed, with a longer tail for larger sizes (Fig. 1.10). The distribution function is defined as

$$n_n(D_p) = \frac{1}{\sqrt{2\pi} D_p \ln \sigma_g} \exp\left[-\frac{1}{2}\left(\frac{\ln D_p - \ln D_{med}}{\ln \sigma_g}\right)^2\right] \tag{1.16}$$

where D_{med} is the median diameter, and σ_g is the geometric standard deviation and it is dimensionless. The geometric standard deviation is the standard

deviation of the logarithms and it is defined as (Hinds, 1999)

$$\ln \sigma_g = \frac{\sum_i n_i (D_{p_i} - D_{gm})}{N - 1} \tag{1.17}$$

There is no closed form algebraic relationship between the standard deviation and the geometric standard deviation (Zender, 2010), but the similarities of the two can be underlined. 68% of particles will fall in the range between $D_{med} \cdot \sigma_g$ and $\frac{D_{med}}{\sigma_g}$ (Seinfeld and Pandis, 1998), which is identical to the fraction of particles enclosed within one standard deviation of the mean of the normal distribution ($\overline{D_p} \pm \sigma$).

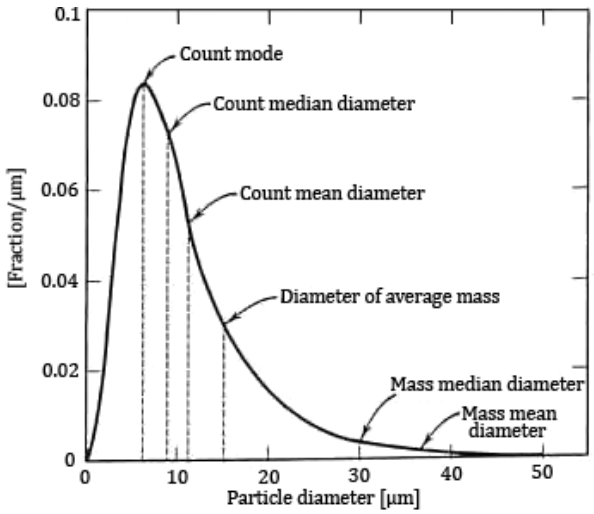

Figure 1.10: A number log-normal distribution in the linear scale and their typical *averages*. The figure is from Hinds (1999).

The fraction of the particles enclosed within any of two diameters D_{p_1} and D_{p_2} can be calculated by the equation:

$$f_N = \text{abs}\left[\frac{1}{2}\left[\text{erf}\left(\frac{\ln D_{p_1} - \ln D_{med}}{\sqrt{2}\ln \sigma_g}\right) - \text{erf}\left(\frac{\ln D_{p_2} - \ln D_{med}}{\sqrt{2}\ln \sigma_g}\right)\right]\right] \tag{1.18}$$

It is worth to point out that Eq. (1.18) is valid for all different log-normal distributions: number, mass, surface, volume distributions... This reveals an important property of the log-normal distribution. If a distribution of particles is log-normal, then all different so-called moment distributions (number, mass, surface, volume...) will also be log-normal. This property is unique for the log-normal distribution (Hinds, 1999). Also, the geometric standard deviation among them will have the same value. This means that if plotted on a logarithmic scale, they would have the same shape, only they would have a

different position on the x axis. The position is defined by the number median diameter for the number log-normal distribution, the mass median diameter for the mass log-normal distribution, etc.

As the log-normal distribution is asymmetrical and skewed, these and other characteristic diameters (mode, mean, etc.) fall on the different locations on the distribution. For any skewed distribution their order is

$$mode < median < mean.$$

This is shown in Fig. 1.10. The mass mode, median and mean diameters are bigger than their respective number diameters. The values of all characteristic diameters are determined by the value of the geometric standard deviation, and the relations between them are described by Hatch-Choate equations ((Hinds, 1999)).

1.5.6 Modes of ambient aerosol distributions

The previous discussion considered the particle size distribution with only one mode. The ambient aerosol size distributions are usually composed of a number of modes which can be presented as a sum of overlapping distributions. By looking the properties of the particles that constitute them, we can divide atmospheric aerosols in four different modes (Fig. 1.11).

Nucleation mode The nucleation mode is formed by particles that have the smallest diameters, below 20 nm. It is constituted by aerosols that are freshly formed by the process of homogeneous nucleation. Aerosol particles in this mode are usually sulphuric and hydrocarbon compounds (Petzold and Kärcher, 2012). Depending on the atmospheric conditions, it is possible sometimes that this mode is not present in the ambient aerosol distribution.

Aitken mode The Aitken mode comprises particles with diameters ≈20 nm to ≈100 nm. The particles in this mode can have different origins. They could be the particles from the nucleation mode that underwent coagulation, or the particles which grew by condensation, or the primary particles directly emitted into the atmosphere (Petzold and Kärcher, 2012).

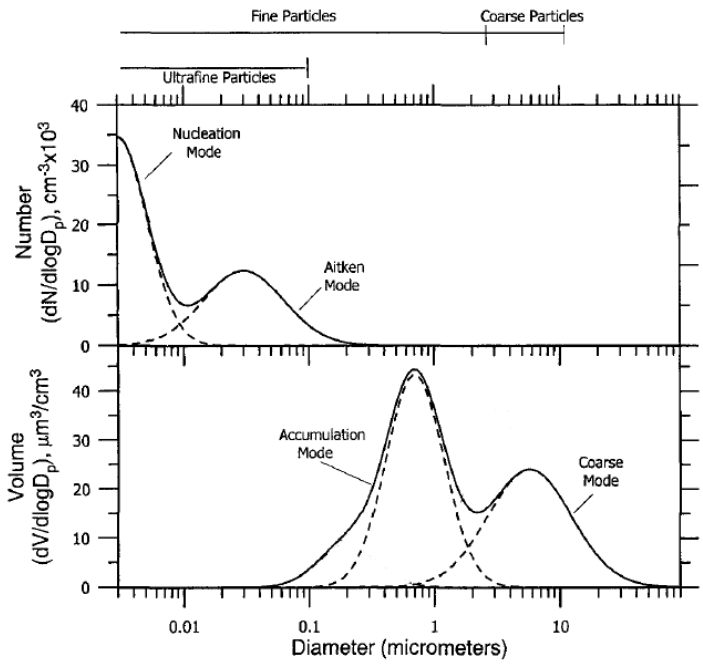

Figure 1.11: Typical number and volume distribution of atmospheric aerosols, and their characteristic modes: nucleation, Aitken, accumulation and coarse mode. The figure is adapted from Seinfeld and Pandis (1998).

Accumulation mode The accumulation mode includes the particles with sizes ≈0.1 µm to ≈2 µm. A big part of the particles in this mode are directly emitted into the air and they are often formed by the process of combustion. These, or smaller particles from the nucleation mode can coagulate, or secondary material can condensate on them. In this mode the deposition mechanisms and coagulation losses are less efficient than in other modes and the particles "accumulate" in this mode (as the name suggests) during their evolution (Hinds, 1999).

Coarse mode The coarse mode contains large particles (greater than ≈2 µm) formed mainly by mechanical processes. In this mode are located mainly primary aerosols – a big part of dust and sea salt particles is in this size range. Biogenic aerosols also mainly fall into this size range. Chemical composition of coarse particles is usually different from other modes due to different formation mechanisms. Also, the mass exchange between the coarse and other modes is not important (Seinfeld and Pandis, 1998).

1.5.7 Distributions of ambient aerosols

Urban aerosols The number size distribution of urban aerosols is dominated by very small particles from the nucleation and Aitken mode. The surface distribution is dominated by the accumulation mode, and the gas-to-particle conversion happens usually on their surfaces. The mass distribution has two peaks, one in the accumulation mode and another in the coarse mode. The aerosol size distribution in urban conditions can be very variable (Zhu et al., 2002).

Marine aerosols The aerosols in the marine environment mostly have a marine origin. There are usually three modes present in the marine aerosol distribution (Fitzgerald, 1991). The number distribution is shared between the Aitken and accumulation mode, and the mass distribution is dominated by the coarse mode (Smith et al., 1989). Under clean marine environment and moderate wind speeds, the sea cruises measurements where sea salt aerosols were separated from other aerosol types have shown that sea salt dominates the number size distribution (O'Dowd and Smith, 1993).

Continental aerosols For aerosols far from urban areas, the number distribution is dominated by two modes at about 0.02 µm and 0.08 µm, and mass is concentrated in the coarse mode at about 7 microm (Jaenicke, 1993). In the areas where the influence of anthropogenic aerosols is negligible, there is no nucleation mode and the accumulation mode is small (Seinfeld and Pandis, 1998).

Desert aerosols The size distribution of desert aerosols is comparable with the distribution of continental aerosols with a low impact of urban aerosols (Jaenicke, 1993). But, it can be very variable. Each soil has its initial size distribution which determines which particles will be emitted into the atmosphere. The soil physical properties, mineralogy and erosion conditions have a role in this process. Different desert dust sources give the size distribution with different characteristics (Alfaro et al., 1998). Also, meteorological conditions influence the size distribution of particles emitted from the same source, even from one to another dust event (Usher et al., 2003). During transport, the size distribution changes as well. Larger, heavier particles will settle more quickly, making a shift in the size distribution to the smaller particles and making the accumulation mode relatively more important.

Free tropospheric aerosols Above the planetary boundary layer, the number of aerosols decreases considerably. The influence of direct emissions at these heights are small, and the size distribution are close to the so-called background aerosol size-distribution. The number distribution is dominated by the nucleation and Aitken mode, and the accumulation mode is relatively more important compared to the aerosol distribution in the lower troposphere. The reason is that the wet deposition, weak in this size range, is stronger in the lower troposphere.

1.6 Lifetime and sinks

Once in the atmosphere, aerosols have different residence times that depends on many factors: type, physical properties, size, altitude range... Residence times of aerosols vary significantly, from a few seconds for very large particles that soon after emission fall back on the ground, to years for sulphate aerosols stable at high altitudes in the stratosphere (e.g. Chazette et al., 1995).

The removal mechanisms can be divided into dry and wet removals. Dry mechanisms are the dry deposition at surface and the gravitational sedimentation, and wet mechanisms are the in-cloud scavenging and the below-cloud scavenging. The contribution and efficiency of these mechanisms is complex and depends on location, extent of these processes, physical and chemical properties of the aerosol particles and some other properties particular to each mechanism.

1.6.1 Dry sinks

Surface dry deposition The dry deposition at surface is an aerosol deposition process in which particles are removed from the atmosphere by the interaction with surface, or more precisely with the atmospheric surface layer and a thin layer of air next to the surface, so-called *quasi-laminar sublayer*. The dry deposition flux directly depends on the aerosol concentration:

$$F_{dd} = -v_{dd} n \tag{1.19}$$

where n is the aerosol concentration and v_{dd} is the deposition velocity [$\mathrm{m\,s^{-1}}$]. The deposition velocity depends on size, shape, density of particles, properties of the surface, and the turbulence in the surface layer. In this process, the

particle first falls through the surface layer. After, it is transported in the quasi-laminar sublayer until it collides with an obstacle on the surface.

Particles near the surface are in the mean flow in the quasi-laminar sublayer. The smallest particles (<0.05 µm) are also subject to the Brownian motion and because of it they collide with the surface (Slinn, 1982a). Larger particles, that are too big for the Brownian diffusion, flow following air streamlines. When they encounter an obstacle (where air streamlines become denser) they can approach too close due to their size and collide with it. This process is called interception. The bigger particles (>2 µm) are not able to follow air streamlines close to the obstacle due to their inertia they leave the flow and collide with the surface (Slinn, 1982a). This process is called impaction. The Brownian motion is effective for particles in the nucleation and accumulation mode. The impaction and interception are effective for the coarse mode. But, for particles in the accumulation mode the surface dry deposition is the least effective (Fig. 1.12).

Gravitational Sedimentation Large particles are also the subject to the gravitational settling. On a falling particle acts two forces, the gravitational force which makes the particle falling, and the drag force which slows down its fall. The sedimentation velocity, which determines the flux of particles that

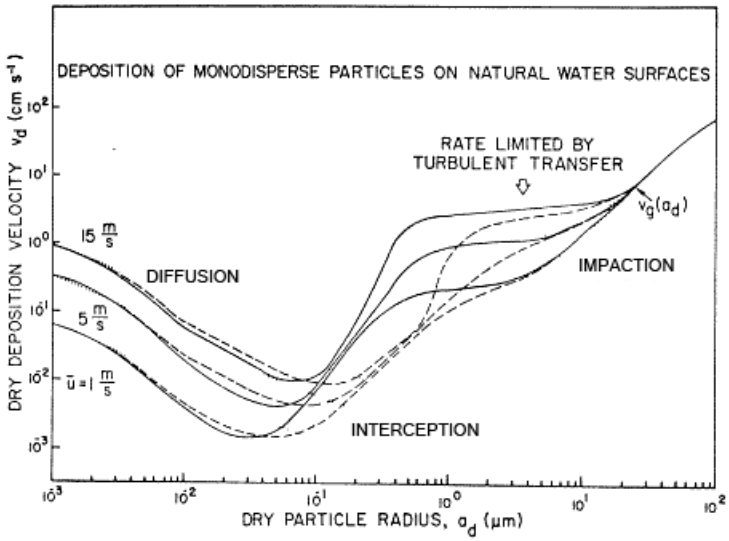

Figure 1.12: Particle surface dry deposition as the function of particle size for deposition on a water surface. Figure adapted from Slinn and Slinn (1980).

settle, depends directly on the mass of the particle. This means that for large particles the sedimentation is the dominant removal mechanism: particles with diameters >10 µm have sedimentation velocities >10 m h^{-1}. This makes their atmospheric lifetime very short.

1.6.2 Wet sinks

Wet removal mechanisms are processes that act on aerosols via atmospheric hydrometeors (cloud droplets, rain, snow, fog) and depose them to the surface. Aerosols can be scavenged when precipitation (cloud droplets) forms – *in-cloud scavenging*, or when precipitation fall – *below-cloud scavenging*. Both mechanisms can be efficient in the aerosol removing. Their efficiency is characterized by the scavenging coefficient Λ, and the change of the aerosol concentration due to the wet deposition is

$$\frac{\partial n}{\partial t} = -\Lambda n \tag{1.20}$$

The scavenging coefficient is a complex parameter that depends on the process of wet deposition involved, the properties of hydrometeors, the properties of aerosols and meteorological conditions. Wet deposition processes are reversible, because all hydrometeros that scavenged aerosols can also evaporate, releasing aerosols back into the air.

A number of synonyms are used for the wet deposition. Wet deposition is sometimes referred as wet removal, wet scavenging, or precipitation scavenging. In-cloud scavenging is also known as rainout, while below-cloud scavenging is known as washout.

In-cloud scavenging Aerosols act as cloud condensation nuclei (CCN) and make the starting point of the formation of cloud droplet. This part of the in-cloud scavenging is known as *nucleation scavenging*. But, not all aerosols act as CCN. Their activation as CCN depends on: their type because more hydrofillic aerosols are more easily activated as CNN, their size because aerosols below a certain size cannot make CCN, and the state of supersaturation in the cloud because the aerosol threshold size depends on the magnitude of supersaturation. The nucleation scavenging can scavenge a large part of the aerosol mass in a cloud.

Besides the nucleation scavenging, aerosols can be scavenged inside of non-raining clouds by direct collisions with cloud droplets. Collision efficiencies depend on the size of the aerosols, and only the smallest aerosols can be

efficiently scavenged by this process. The lifetime of an aerosol particle larger than ≈0.1 µm in a cloud due to this process is longer than the lifetime of clouds (Seinfeld and Pandis, 1998). In this way only a small part of the aerosol mass gets incorporated inside the cloud droplets.

Below-cloud scavenging A hydrometeor that is falling (raindrop or snowdrop) can collide with an aerosol particle and collect it. A raindrop while falling, perturbs the air around it. As raindrops are usually significantly larger, aerosol particles follow these flow streamlines when approaching a raindrop. Similarly to the surface dry deposition, the interaction between an aerosol and a raindrop depends on their sizes. The smallest particles are subject to the Brownian motion, and larger particles to the interception and impaction (Slinn, 1982a). These processes are collectively the least effective for the particles in the accumulation mode, for the size range of 0.1 µm to 1 µm.

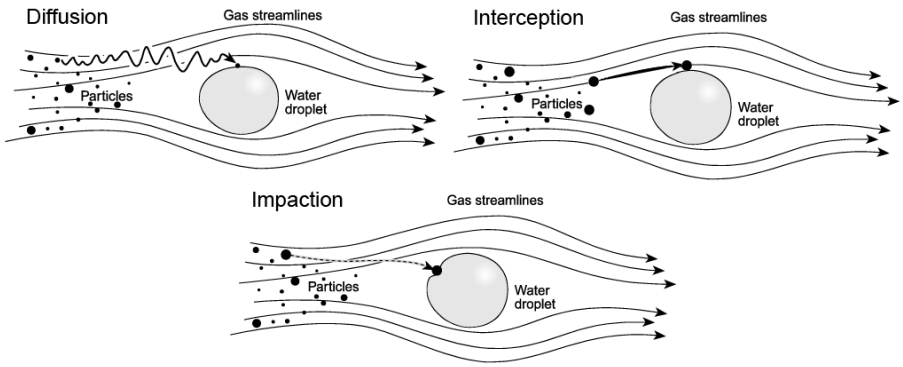

Figure 1.13: Interaction mechanisms of an aerosol particle with water droplet in the below-cloud scavenging process.

1.6.3 Other processes

Secondary aerosols can evaporate in response to the lost equilibrium between the gas and aerosol phase. Also, the coagulation is not strictly a removal mechanism, but it lowers the number of the particles and generally shortens the lifetime of particles.

1.7 Climatology

Spatial and temporal distribution Most of the natural aerosols are located in the troposphere. There, aerosol mass and number concentrations are very variable, and different species and different regions have their own patterns of the geographical and temporal distribution.

The sea salt production rapidly changes with the change of meteorological conditions, especially the wind speed. High concentrations are more typical over southern and northern oceans where strong winds are frequent. Seasonal variation can be strong, especially in the high latitude regions, with a maximum in the winter time (Gras, 2003). In the North Atlantic, the seasonal relative amplitude is even greater than 8:1. Sea salt contributes with 50 % to 70 % of aerosol mass where the influence of anthropogenic sources is weak (IPCC, 2013).

The desert dust production is very variable. It depends on the location and the season. In the northern hemisphere, the maximum in mineral aerosol emissions and transport is in the spring. Saharan dust is transported across the Central Atlantic to North and South America with a maximum in June – August. The dust from East Asia is transported to the Western Pacific with a maximum in March – May. Even in the urban areas of East Asia the desert dust dominates the aerosol mass, with about 35 % (Zhang et al., 2012b). Other important dust regions are the Arabian peninsula and sea and Australia whose dust is transported across the south-west Pacific usually during spring and summer.

Biomass burning aerosol distribution is also highly variable, but regional and seasonal patterns exist. The most important fire regions are: Africa (10°– 20°S) with a maximum in August – October, Brazil with a maximum in July – September, Indonesia during the dry season which is typically in June – October, northern Australia with a maximum in May – October, and the fires of northern boreal forests (Siberia or Canada) with peaks in spring and autumn. These aerosols are usually transported in layered plumes below 5 km.

In urban areas, organic carbon contributes to the aerosol mass with ≈15 % or even more in the urban North America and South America. Black carbon makes about 10 % to 15 %. Sulphates usually contribute with 10 % to 30 %, and nitrates and ammonium about 6 % and 4 %, respectively (IPCC, 2013).

In polar regions, aerosol concentrations are usually low because of the lack of strong sources. Nevertheless, in the winter and early spring, the Arctic is influenced by anthropogenic aerosols, mainly carbonaceous aerosols and sulphates from the mid latitudes. This phenomenon is known as the Arctic haze (Barrie, 1986).

Vertical distribution/repartion Aerosols mainly have their sources in the lower troposphere, near the surface. At higher altitudes, concentrations decrease by a rate of about two-thirds for every kilometer up to ≈5 km above land, and up to ≈2 km above oceans (Jaenicke, 1993). Above this height, aerosols concentrations are fairly constant with very weak vertical structures and are so-called background aerosol concentration. The number background aerosol concentration is about $700\,\mathrm{cm}^{-3}$. In the stratosphere, aerosol concentrations further decrease, but the large particle concentrations increase. Aerosols in the stratosphere are injected directly during volcanic eruptions or from upper troposphere. They return to troposphere by sedimentation for the largest particles, or by tropopause foldings in the mid latitudes (Gras, 2003).

Primary aerosols are emitted mostly from the low altitude and they show stronger vertical gradient than aerosols in total (Petzold and Kärcher, 2012).

Trends Aerosols depend on the state of the atmosphere and respond to any changes in the system. The changes can be physical (changes of temperature, precipitation, radiation, wind, soil properties, etc.), chemical (changes in abundance of reactive species) or biological (changes in vegetation cover and properties). Observational sets are long enough to deduce aerosol response to the past changes and to conceive the future trends.

Sea salt concentrations are sensitive to wind speed and sea surface temperature, but the changes in sea salt concentrations are uncertain mainly due to uncertainties in possible wind changes over ocean. Arid and semi-arid regions grow and become more numerous (Woodward et al., 2005) which leads to increased dust emissions, but still the confidence of these predictions is fairly low. Sulphates and nitrates are sensitive to the changes in temperature and in precipitation, and studies suggest a small to moderate reduction in these secondary inorganic aerosols. Increase of biomass-burning events leads to a moderate increase in emissions of biomass-burning carbonaceous aerosols (IPCC, 2013).

Regarding anthropogenic sources, studies based on remote and in-situ measurements demonstrate that aerosol concentrations in Europe and North America shows a decrease since 1990s, while in eastern and southern Asia an

increase since 2000s (De Meij et al., 2012; Asmi et al., 2013). Models suggest that anthropogenic emissions will show a moderate decrease in the future (IPCC, 2013).

1.8 Radiative Transfer

Aerosols in the atmosphere interact with the radiation. This interaction is important for the climate and for the aerosol detection and remote sensing. Besides photochemical reactions, aerosols interact with the radiation in two ways: they can scatter or absorb radiation (Fig. 1.14). Scattering is the deviation of radiation from its original direction. It happens because of different characteristics of medium that the radiation encounters during its propagation. Refraction, reflection, diffraction can be considered as different forms of the scattering. Absorption is the uptake of the energy of photon by absorber, and its transformation into thermal energy. The combined effect of scattering and absorption is called extinction.

Here, we consider only the elastic scattering, in which the wavelength before and after the scattering stays the same.

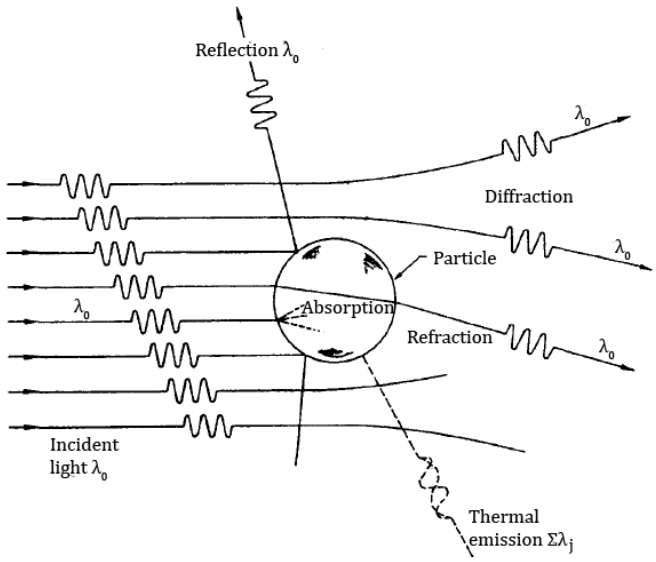

Figure 1.14: Interaction mechanisms between incident radiation and an aerosol particle. The figure is adapted from Seinfeld and Pandis (1998).

Lambert-Beer law Aerosols reflect one part of the incoming solar radiation back to the space, one part is scattered in the atmosphere, and the last part is absorbed and converted into heat. The interaction radiation-aerosols is described by Maxwell equations for electromagnetic radiation. Using the Maxwell theory, the Lambert-Beer law can be derived. It describes how the light beam is reduced due to extinction by particles:

$$\frac{dI(\lambda)}{dz} = -b_{ext} I_0(\lambda) \tag{1.21}$$

where I_0 and I are intensities of the incoming and exiting radiation, and b_{ext} is the aerosol extinction coefficient [m^{-1}] that describes the rate of extinction, and which is the sum of the scattering and absorption components ($b_{ext} = b_{scat} + b_{abs}$). Equation (1.21) shows that the extinction along the path of radiation through the aerosol medium is linear with the intensity of the incoming radiation.

1.8.1 Scattering

Scattering regimes The behavior of the scattering of a photon by a scatterer (molecule, particle, etc.) depends on the ratio of the size of scatterer and the wavelength of photon. If the size of the scatterer is much smaller than the wavelength of the light ($D \ll \lambda$), we have the Rayleigh scattering regime. In this regime the scattering is usually by gases, and there is a strong dependency on the wavelength, the scattering strongly favours shorter wavelengths $I \sim \lambda^{-4}$.

If the size of scatterer is comparable to the wavelength of the light ($D \sim \lambda$), we have the Mie scattering regime. The scattering by aerosols usually falls into this scattering regime. In the Mie scattering regime, the scatterer is represented as an isotropic, homogeneous, dielectric sphere. The Mie scattering does not depend as strongly on the wavelength of the light as the Rayleigh scattering. The scattering from particles is much stronger than that from molecules, and the size of a particle also has a role. Bigger particles scatter more light.

Particles significantly bigger than the wavelength of the light ($D \gg \lambda$) fall into the so-called geometric scattering regime where scattering is determined by the laws of geometric optics, and which does not depend on the wavelength of the light. This regime applies only to the biggest aerosols which have very short lifetimes.

Scattering phase function The scattering of the light is angle dependant, and it is the scattering phase function which describes this dependency at a given wavelength. For the Rayleigh scattering, the angle dependency is not very strong. The scattering phase function is symmetric in the forward and backward directions, and at right angles it has a half of the forward intensity. For the Mie scattering, the angle dependency is strong and it is the strongest in the forward direction. Also, the bigger the particle is – the more dominant the forward scattering is. The forward scattering by particles can be explained in the theory by considering it with the scattering by a dipole array (Bohren, 2001). The more dipoles in the array, the more they will collectively scatter in the forward direction.

The scattering in the backward direction is called the backscattering, and it is in the basis of the atmospheric sounding by lidar systems.

1.8.2 Aerosol optical properties

Refractive index As already seen in Section 1.4 aerosols have different optical properties. These properties are described by the complex refractive index ñ:

$$\tilde{n} = n_{re} + in_{im} \tag{1.22}$$

The real part of the index n_{re} describes the scattering by particles, while the imaginary part n_{im} describes the absorption. The refractive index in atmospheric calculations has to be considered as relative to the surrounding air. The refractive index of the vacuum is $\tilde{n}_0 = 1 + 0i$. For the air it is very close to this value, and practically they are considered identical. The refractive index depends on the physical properties of the material and the radiation wavelength.

1.8.2.1 Extinction properties of a single particle

The extinction by an aerosol particle depends of its composition, size, shape and the wavelength of the light. The size is usually expressed as the dimensionless *size parameter*, x:

$$x = \frac{\pi D_p}{\lambda} \tag{1.23}$$

For spherical particles, extinction properties can be calculated by Mie theory. This theory enables us to determine the particle extinction cross-section, C_{ext} [m^2], which represents a hypothetical area that describes the likelihood that a photon will interact with the particle. The extinction cross-section is a function of the size parameter and the refractive index. From it, the extinction efficiency of a single particle (or a group of monodisperse aerosols), Q_{ext}, is calculated as

$$Q_{ext}(D_p, \tilde{n}, \lambda) = \frac{C_{ext}(D_p, \tilde{n}, \lambda)}{S} \tag{1.24}$$

where S is the geometric surface of the particle [m^2], and the extinction efficiency is dimensionless.

The extinction efficiency has two components, the scattering and absorption efficiency, Q_{scat} and Q_{abs}. It can also be calculated for a population of aerosols. The ratio between the scattering and extinction efficiency represents the single scattering albedo:

$$\omega = \frac{Q_{scat}}{Q_{ext}} \tag{1.25}$$

The single scattering albedo for a non-absorbing particle would be equal to 1, but usually it takes values from 0.95 to 1.0. In more polluted areas with a lot of carbonaceous aerosols its values are much lower (De Leeuw et al., 2011).

1.8.2.2 Extinction properties of an ensemble of particles

If we take into account the scattering by an ensemble of particles, we can assume that the total scattered light intensity is just the sum of intensities scattered by individual particles. In this case it is considered that the exiting light rays are scattered at most only once. This is called a single-scattering approximation and it is true if the average distance between particles is much larger that the size of particles. In the atmosphere this is true even for large aerosol concentrations (Seinfeld and Pandis, 1998). Here we only consider the case of single scattering approximation.

Optical depth If we consider the layer of aerosols, their summed extinction effect can be expressed by the extinction coefficient α_{aer} [m^{-1}] which is for monodisperse aerosols equal to the product of the particle number concentration and the extinction cross-section:

$$\alpha_{aer} = C_{ext} n \tag{1.26}$$

if we consider polydisperse aerosols the extinction coefficient depends on the size distribution

$$\alpha_{aer}(\lambda) = \int_0^\infty f_N(D_p) C_{ext}(D_p, \tilde{n}, \lambda) dD_p \tag{1.27}$$

To calculate how much light will pass through the layer of aerosols at the height z we have to integrate Eq. (1.21):

$$I(\lambda) = I_0(\lambda) \exp\left(-\int_0^z \alpha_{aer}(\lambda, z) dz\right) \tag{1.28}$$

The term $\int_0^z \alpha_{aer} dz$ is equal to the aerosol optical depth (AOD) of the layer, τ:

$$\tau = \int_0^z \alpha_{aer} dz \tag{1.29}$$

The AOD is the parameter which is frequently used to represent the extinction of the light by aerosols, or even the hint of aerosol quantities and it is the primary quantity observed and retrieved by satellites (De Leeuw et al., 2011). From Eq. (1.28) AOD can be expressed as a negative logarithm of the fraction of the light that passes through the extinction layer. The fraction of the light attenuated in an aerosol layer is

$$\frac{I_{ext}}{I_0} = 1 - e^{-\tau}$$

For example, the aerosol layer of AOD of 1.0 will extinct $1 - e^{-1} = 63\%$ of light and only 37% will pass through. The AOD in the atmosphere usually takes values from 0.05 in remote, clear environments to 2.0 or more in locations with high aerosol concentrations, like during desert dust outbreaks or in forest fire plumes.

Angstrom exponent The aerosol optical depth is the function of the wavelength of the light. This dependency is described by the Angstrom exponent α, where

$$\tau \approx \lambda^{-\alpha}. \tag{1.30}$$

The Angstrom exponent can be calculated from two optical depths at two different wavelengths

$$\alpha = -\frac{\log \frac{\tau_{\lambda_1}}{\tau_{\lambda_2}}}{\log \frac{\lambda_1}{\lambda_2}}, \tag{1.31}$$

and it can be used to calculate the aerosol optical depth at another wavelength (under assumption that it stays constant for the whole considered spectral

domain)

$$\tau_\lambda = \tau_{\lambda_1} \left(\frac{\lambda}{\lambda_0} \right)^{-\alpha}. \tag{1.32}$$

The Angstrom exponent can also hint about the typical size of the size distribution of aerosols. Smaller particles give bigger values of α. The angstrom exponent has values in the range of about 0 to 4. The value $\alpha \approx 4$ corresponds to the case where particles are so small that they are between the Mie and Rayleigh scattering regime. And values of the exponent around $\alpha \approx 0$ correspond to the case where particles are so big that they are at the limit of geometric optics scattering regimes. This also tells us that AOD depends on the light wavelength more strongly for small particles than for big particles (Van de Hulst, 1981).

1.9 Observations

In-situ measurements It is necessary to make measurements of aerosols in order to determine their properties and their impact on the environment. The first aerosol measurements were ground-based in-situ measurements. This type of measurements gives very detailed and accurate information about the size distribution, the chemical composition, the optical properties, etc. They are performed by collecting aerosols, filtering them according to their size and analysing the collected representative sample. Sampled aerosols can be treated with many different methods and instruments to acquire their properties. But ground based in-situ measurements are sparse and limited in locations and time. It is quite expensive to build an extensive network, demanding a continuous effort and funds in order to acquire data for extended periods of time.

In-situ measurements nowadays are not only limited to ground based measurements. The ambient aerosol properties can be obtained by instrument on balloons or aircrafts. These flights usually have a good horizontal and vertical coverage, but they are limited to short periods of time during extensive observation campaigns (e.g. Dulac and Chazette, 2003; Heese and Wiegner, 2008).

Remote sensing measurements From the ground, aerosols can be also measured by remote-sensing methods. These measurements, obtained by instruments like photometers or lidars, can provide a variety of aerosol properties. They can be performed many times per day, and give a detailed information about aerosols on the local scale.

The remote-sensing methods show their full potential for aerosol measurements from space (Gras, 2003). Satellites provide routine measurements on a global scale. The time resolution ranges from almost continuously to at worst once every 2 to 3 days. Observations done by a single instrument cover all continents and for at least several years ensure the uniformity in the observation technique. This allows us to compare aerosols on the planet for longer periods of time and to establish the aerosol climatology and long-term trends. Aerosol satellite observations have been continuously performed for more than 30 years. Satellites instruments can be nadir-viewing which measure column-integrated quantities, or limb-viewing which provide aerosol profile information and much longer path through the atmosphere (Kaufman and Tanré, 2003). Nowadays, nadir-viewing instruments are a lot more common.

1.9.1 AOD satellite retrieval

Aerosol optical depth is the most frequently retrieved aerosol quantity from space. It was also the first one (Griggs, 1975). The radiation reflected and emitted by Earth also picks the effects of the surface and the atmosphere. The AOD retrieval is based on measuring the radiances at the top of the atmosphere and extracting the aerosol effect out of it. Aerosols can change the intensity of light, its polarization or angular properties (Kaufman and Tanré, 2003). Radiances are measured by spectrometers or radiometers that can cover a wide range of the spectrum. The retrieval is usually done for the visible and near-infrared parts of the spectrum because in these ranges the scattering is the dominant process of interaction of radiation with matter in the atmosphere. Still, for bigger aerosols particles the retrieval can be done in the infrared part of the spectrum.

Algorithms Aerosol properties can be retrieved only when there are no clouds in the instrument line of sight. When there are clouds in the field of view, their radiance dominates the signal that arrives to the instrument and the aerosol retrieval cannot be performed. For nadir-viewing instruments, besides clouds, the surface reflectance also has an effect on the signal that has to be removed. To retrieve AOD over the ocean and over the land, it is usually necessary to use two different algorithms. Over the ocean, surface reflectance characteristics are relatively well known. But over the land, the surface reflectance can be very variable. It varies spatially and temporally: it depends on the part of day, seasons, surface properties and precipitation, etc. It is necessary to make assumptions on the characteristics of the surface, its reflectance and how it

changes for different wavelengths. If the instrument is capable to observe with different viewing angles or at short-wavelength infrared (SWIR) wavelengths where the aerosol effect is negligible, the surface reflectance can be retrieved directly (De Leeuw et al., 2011; Kaufman and Tanré, 2003), and this information can be used in the aerosol retrieval. Observations over highly reflective surfaces like deserts, are particularly difficult for aerosol retrieval. In these regions it is necessary to make algorithms that use shorter wavelengths in which the reflectance is low (Hsu et al., 2004).

The next step is to calculate the top of the atmosphere radiances by a radiative transfer model (RTM) to include the effect of gases in the atmosphere. The RTM calculations are done for the viewing geometry (viewing angle of instrument, positions of satellite and Sun) using different probable values of surface characteristics and a wide range of aerosol atmospheric compositions obtained by aerosol models. These results are compared to the observed atmospheric path radiance (radiance without cloud and surface effects) and, by minimizing the difference between them, the AOD is estimated (De Leeuw et al., 2011).

Advances in the retrieval techniques also enable the retrieval of other aerosol properties than AOD: aerosol types, partitioning between fine and coarse particles, single scattering albedo, angstrom exponent, effective radius of the dominant mode, etc. (Lee et al., 2009).

Nowadays, aerosol retrieving techniques and algorithms are very sophisticated and they give results of very good quality. However, there are sometimes still large differences between retrievals from different (or even the same) instruments. They come from differences and uncertainties in calibration, sampling, cloud screening, surface reflectivity algorithms, aerosol models, wavelengths, viewing geometries and different instrument characteristics (Li et al., 2009; Kokhanovsky et al., 2010).

1.9.2 AOD observations

Ground based The aerosol optical depth from the ground is measured by the extensive AERONET (AErosol RObotic NETwork) network (Holben et al., 1998). It includes a few hundreds of stations, many of them permanently measuring aerosols by sunphotometers. It already provides data collected long enough for climatological aerosol studies on the local scale. The AOD measurements are very accurate: $\Delta\text{AOD} = \pm 0.01$. They cover different wavelengths, and besides AOD stations provide other aerosol properties like the

single scattering albedo, refractive indices and the size distribution. Besides AERONET, there other aerosol ground network like GAW (Collaud Coen et al., 2013), SKYNET (Sano et al., 2003), etc.

Spaceborne Some spaceborne instruments that measure aerosols were specially designed for this purpose, while others are not directly designed for it, but have been used for the retrieval of aerosols. An instrument dedicated to the aerosol measurements would desirably have the spectral sensitivity from ultra-violet to thermal infra-red, with multiple views and polarization sensitivity. With these characteristics it is possible to retrieve a wide range of aerosol properties.

The first operational aerosol products were from AVHRR (Advanced Very High Resolution Radiometer) (Stowe et al., 1997) and TOMS (Total Ozone Mapping Spectrometer) (Torres et al., 2002) launched at the end of 1970s and both primarily used for the retrieval of other parameters than aerosols. Instruments that are dedicated for the retrieval of aerosols are, for example, MISR (Multiangle Imaging Spectro-Radiometer), MODIS (Moderate Resolution Imaging Spectrometer), POLDER (Polarization and Directionality of the Earth's Reflectance). MISR (Kahn et al., 2005) and MODIS (Remer et al., 2005) are both capable of retrieving aerosols over ocean, land and highly reflective surfaces. MODIS is a set of instruments on-board of two sun-synchronous satellites, Terra and Aqua, observing at multiple wavelengths. MODIS AOD products are the most used aerosol observations from space (De Leeuw et al., 2011); over ocean it can separate the fine and coarse particles, and over the land the aerosol type. The AODs are retrieved with an error of

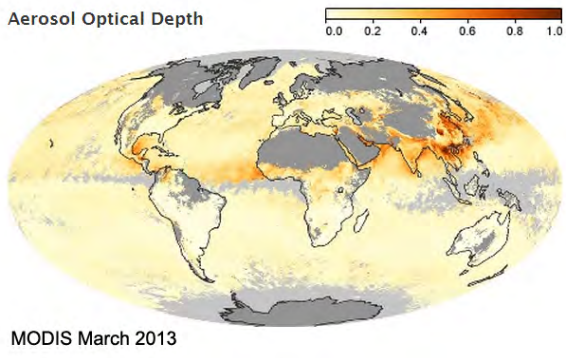

Figure 1.15: An example of aerosol optical depth measured by the satellite instrument MODIS.

$\pm 0.05 \pm 0.20$AOD over the land (Chu et al., 2002) and $\pm 0.03 \pm 0.05$AOD over the ocean (Remer et al., 2002).

MISR observes at multiple wavelengths and at multiple angles, while POLDER is the most complete instrument by observing also the polarization characteristics of the light. Of the limb-viewing instruments, we can mention SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) (Bovensmann et al., 1999) which was primarily designed for the trace gases observations and it was operational until 2012. At the geostationary orbit is SEVIRI (Spinning Enhanced Visible and Infrared Imager) which provides observations of whole planet's disk every 15 minutes. Its shortest wavelength is 0.6 µm and it lacks SWIR channel useful for the surface reflectance determination over land.

1.10 Lidar and aerosols

1.10.1 Lidar principles

Lidar is an acronym of LIght Detection And Ranging. It is a system of instruments used for the atmospheric sounding. With a lidar it is possible to obtain properties and atmospheric profiles of temperature, wind, concentration of atmospheric gases, clouds or aerosols (Argall and Sica, 2003). The main idea behind it is to transmit a beam of light and to detect the radiation that is scattered back to the instrument. The received radiation will have signatures of the air through which it travelled, and is analysed after the detection. Every lidar system is composed of the transmitter, the receiver and the detector. Lidars use wavelengths in ultraviolet, visible and infrared part of the spectrum, and the radiation in lidar systems is generated by lasers.

The light is transmitted in short, strong pulses, which have a narrow spectral width and low divergence. The signal that is backscattered is collected by the receiver, usually a telescope. The detector detects and records the signal by converting it to an electronic signal. The signal before entering the detector can be further filtered to improve its quality. This could be a narrow band filtering to reduce the background or the shutter which will limit too strong signals from near-field returns. The detector records the signal intensity as a function of height. The height at which the signal is backscattered is possible to determine by timing the returned signal. Because the exiting pulse is very short

by knowing the time and the velocity of the signal the height is determined by

$$\Delta Z = \frac{t_{tot} c}{2}, \tag{1.33}$$

where c is the speed of light, t_{tot} is the time after which the signal is detected and during which it travelled the distance $2\Delta Z$.

The length of the pulse will determine the vertical resolution of the lidar. Although very short, during the time t the pulse will travel the distance ct, which is the spatial length of the pulse. In the same moment the leading edge of the signal will be at the height z_1, and the trailing edge at the height z_2, and $z_1 - z_2 = ct$. The backscattered radiation of the leading edge will encounter the trailing edge of the pulse at the height $z_2 + \frac{z_1 - z_2}{2}$ because the two edges will travel toward each other in this short period of time. If in this moment the part of the trailing edge is backscattered, the two signals will arrive at the detector at the same time, but backscattered at two different heights. The distance

$$\frac{z_1 - z_2}{2} = \frac{tc}{2}, \tag{1.34}$$

is the vertical resolution of the lidar.

1.10.2 Aerosol detection

Using lidar observations it is possible to detect aerosols and to determine their vertical distribution and properties. The particle scattering is usually measured by the elastic-backscatter lidar. It is a classic form of lidar where the radiation emitted from the lidar is elastically scattered back, where the wavelength remains unchanged. The quantity which describes how much light is reflected back is called the backscatter coefficient. It is the specific value of the scattering coefficient for the scattering angle of 180°. The unit is $[\mathrm{m}^{-1}\,\mathrm{sr}^{-1}]$.

In the atmosphere, the lidar pulse is scattered by aerosols and molecules, so the backscatter coefficient will have two components, molecular and aerosol backscatter coefficient

$$\beta = \beta_{aer} + \beta_{mol}. \tag{1.35}$$

The molecular scattering results from the Rayleigh scattering mainly on nitrogen and oxygen molecules and it decreases with height because of the decreasing air density. For certain wavelengths, other gases can also have an influence because of their absorption and re-emission which can coincide with the direction of backscattered light. The particle scattering is very variable and

depends on the aerosol concentration and properties, which are the quantities we want to determine.

One part of the lidar pulse gets lost on the way to the scattering medium and back. The transmitted part T_r can be calculated by Eq. (1.28), only considering the two-direction path

$$T_r = \exp\left(-2\int_0^z \alpha dz\right) \tag{1.36}$$

The values of transmission can be between 0 and 1. The extinction coefficient α contains the scattering and absorption of molecules and aerosols, which has the four components

$$\alpha = \alpha_{aer} + \alpha_{mol} = \alpha_{aer_{abs}} + \alpha_{aer_{sca}} + \alpha_{mol_{abs}} + \alpha_{mol_{sca}}. \tag{1.37}$$

As already discussed, the backscattering and transmission depend on the wavelength of the light and the height.

Lidar equation The lidar equation defines the intensity of the backscattered signal that the system will detect. In the case of the single-scattering approximation it has the form:

$$P = \frac{P_0 K}{Z^2}\beta \exp\left(-2\int_0^z \alpha dz\right). \tag{1.38}$$

The received signal intensity P [W] depends on: the strength of the exiting signal P_0 [W] which decreases with the height Z [m], the degree of backscattering defined by the backscattering coefficient β and its extinction on the path defined by the transmission term. K is the parameter which described the characteristics of the instrument (resolution, efficiency, the geometry and surface of the receiver, etc.).

The received signal will contain also the background signal which during the day comes mostly from the Sun, and during the night from the Moon, stars and the artificial light. Undesired signal includes also the detector noise. These additional sources of the signal have to be removed during the signal processing.

1.10.3 Retrieval of the aerosol properties

Elastic-backscatter lidar The elastic backscatter lidar is the most widely used lidar system. Equation (1.38) shows that the aerosol properties which

can be retrieved are the aerosol backscatter and extinction coefficients profiles. These two quantities have to be estimated from the information of only one measured quantity, the lidar backscatter profile. Also, the molecular and aerosol contributions are mixed together in the signal, and it is necessary to separate them. One of the most used techniques for the retrieval is the Klett method (Klett, 1981). Klett reformulated the lidar equation in the form of the Bernoulli equation and developed the backward integration scheme to solve it. This method requires the assumption of the backscatter coefficient at the reference height and of the ratio of extinction and backscatter coefficients for aerosols, the so-called lidar ratio L_{aer}

$$L_{aer} = \frac{\alpha_{aer}}{\beta_{aer}}. \tag{1.39}$$

The lidar ratio depends on the physical and chemical properties of the aerosols, but also on the meteorological conditions, i.e. relative humidity (Ackermann, 1998). It varies significantly with height even in the same lidar profile. The largest uncertainty in the retrieval comes from the estimation of the lidar ratio (Böckmann et al., 2004). If the lidar measurements are accompanied with measurements of optical depth by sunphotometer, it is possible to improve the estimate of the lidar ratio of the whole atmospheric column by comparing the optical depth inferred from lidar and photometer observations.

Other lidar systems With more advanced lidar systems it is possible to divide the molecular part from the aerosol part of the signal more directly, and consequently, to retrieve the aerosol extinction coefficient and the aerosol backscattering coefficient independently. The High spectral resolution lidar (HSRL) can measure separately the signal from molecular Rayleigh scattering and the signal from the aerosol scattering. It relies on differences of the spectral width of the light backscattered by molecules and particles. Fast moving molecules will produce a Doppler spectral broadening because of their strong thermal motion, while slow moving particles will produce a negligible spectral broadening. Because of this difference, with a very narrow filter it is possible to isolate into separate signals the thin aerosol peak and the wider molecular signal (Ansmann and Müller, 2005).

Raman lidars can be mainly used during the night-time, and they measure separately the elastically backscattered light from molecules and particles as the normal elastic-backscatter lidar, and the inelastically Raman backscattered radiation by nitrogen and/or oxygen molecules (Argall and Sica, 2003).

The received backscattered signal that is produced only by molecules (Rayleigh backscattered signal for HSRL and Raman backscattered signal for Raman lidar) is not influenced by the aerosol backscattering, but it is by the

aerosol extinction. By knowing the molecular number density it is possible to retrieve the aerosol extinction coefficient profile (Ansmann and Müller, 2005). After it, from the ratio of total (aerosols + molecules) and molecular only backscatter signals, the aerosol backscattering coefficient can be inferred (Melfi, 1972). In this case, the assumption of the lidar ratio is not needed.

From lidar measurements it is also possible to estimate other aerosol properties. For example, by using more wavelengths at the same time we can get information about the aerosol size distribution, shape, type or optical properties.

1.10.4 Lidar aerosol observations

Ground-based observations On the ground, lidars provide a point source aerosol profile measurements. To improve the horizontal representativity, networks of lidar systems are established. The EARLINET (European Aerosol Research LIdar NETwork) is the European lidar network established in 2000 that now consists of 27 lidar systems across Europe and it grows into an important data set of the aerosol vertical distributions (Matthais et al., 2004). A similar network grew in Asia - the Asian Dust Network (ADNET, Murayama et al. 2001) and one worldwide - the Micro-Pulse Lidar Network (MPLNET, Welton et al. 2001).

Spaceborne observations From the orbit, lidars provided the first global views of the aerosol vertical distribution. The first lidar in the orbit was LITE (Lidar In-space Technology Experiment) that flew in 1994 on-board the space shuttle Discovery (Winker et al., 1996). It provided data of 53 hours of observations of the distribution of desert dust, smoke, and other aerosols (Berthier et al., 2006).

CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) is a dedicated aerosol and cloud lidar in the Earth's orbit launched in 2006 (Winker et al., 2010). It is in the satellite constellation A-Train, together with POLDER and MODIS. This constellation gives the most complete aerosol observations with different complementary instruments. CALIOP is a two-wavelength lidar that has the ability to differentiate different types of aerosols. It provides the processed backscatter signal, and the retrieved backscattering and extinction coefficients.

The relatively new field of spaceborn lidar observations proved to be very useful. Although with the coverage that is considerably smaller than the

coverage of AOD satellite measurements, the global observations of aerosol vertical distribution improved the quality of data obtained from space, and the coverage of the ground-based lidar networks (IPCC, 2013). In the near future the new spaceborn lidars are already planned, notably ADM-Aeolus and EartCARE missions.

2

Modeling of aerosols and the CTM MOCAGE

This chapter describes the principles of aerosol modelling and the chemical transport model MOCAGE.

Contents

2.1 Aerosol modelling

Simulating all the aerosols and atmospheric processes mentioned before is a continuous and complex process: it is necessary to express physical laws in mathematical terms, develop numerical methods for solving them, explicitly resolve the concerned processes, or if not possible, parameterize them into suitable schemes which then can be incorporated as the model components. The final goal is to make an accurate mathematical representation of processes that govern the atmosphere and its evolution.

The aerosol diversity and the complexity of the aerosol processes, presented in Chapter 1, is not possible to achieve in such details in the model. Models have limited resources which demands many simplifications in the aerosol representation. One of the approaches is to focus on a specific problem which can be considered in more details.

By determining and understanding a problem of interest we define what kind of model we have to use. Models have many components which have different importance and relevance for different scientific problems. Aerosols are usually simulated either in dispersion models in the case where no important interactions between different types of aerosols or between aerosols and gases exist, or in chemical transport models (CTM).

To be able to respond to the problem of interest, it is necessary to define the appropriate scales. Different processes happen over different time and spatial scales that require different minimal timesteps. Moreover, the scales determine which processes should be included in the model and to which details. Also, for the chosen timescale one process can be considered as dynamical, but for longer timescales it can be considered constant, which will demand simpler parameterizations.

In models that cover the largest spatial and time scales (for example global CTM or climate models) it is necessary to implement a larger degree of simplifications and hypotheses in order to achieve results relevant for these largest scales.

Modeling advances Present-day's state-of-the-art aerosol chemical transport models are powerful tools to study the evolution and the spatial and temporal distribution of aerosols. Their complexity and precision closely relate with the development of state-of-the-art supercomputers. The decisions about the spatial and temporal resolution of the model, aerosol representation with different number of bins/modes and how much details can be implemented into the code,

directly depend on the availability of computational power. With advances of computational tools, more explicit, particle-resolved schemes are developed instead of simpler parameterized schemes improving the level of accuracy and complexity of the models.

More computational power enables also adding of new model components or a *coupling* of models, in order to study different complex problems and to achieve further improvements of results. Model can be "coupled", not only with other models, but also with available observations. This enables the improvement of the model predictions and results. This topic, dealing with aerosol data assimilation, is separately covered in Chapter 3.

2.1.1 Uncertanties in models

A lot of processes are responsible for the evolution of the aerosols. Some aerosol processes are still not fully understood, or some of their pieces are unknown, like their efficiency, quantities, etc.

Often, uncertainties in one process influence directly other processes. For example, uncertainties in emissions lead directly to uncertainties in the composition and mixing state of aerosols, which is then crucial for cloud processes. Many processes are complex, and it is necessary to describe them via parameterizations. Each parameterization relies on different hypothesis and approximations, and each operator brings some uncertainty. The cumulative effect of uncertainties lead to a diversity in results of different state-of-the-art models (Textor et al., 2006).

There are various degrees of uncertainties for different processes in models due to a specific focus of each model to certain aerosol aspects or just because of the different model structures, modelling approaches, or different parameterization choices. Often, there are not enough observations to do complete sensitivity studies of uncertainties coming from different parameterized processes. Although knowing that the effect of some processes in different models can vary considerably, often it is not clear which would be the best modelling choice (Textor et al., 2007). Besides these uncertainties, limiting computational power also contributes to an incertitude of results. Due to this, simplifications are necessary to be introduced to a model. This can be limiting the number of bins/modes in the aerosol representation, or choosing computationally less expensive schemes. It is necessary to find a compromise between the computational cost and reliability of the schemes. Usually, the type of introduced simplifications depends on the problem of interest.

Although there are still discrepancy between models and observations, each subsequent IPCC (Intergovernmental Panel on Climate Change) climate change report notes considerable improvements in aerosol models. As knowledge and tools advances, more physically complete representations of aerosols are incorporated and simulated (IPCC, 2001, 2007, 2013).

2.2 CTM MOCAGE

MOCAGE (*fr*: Modèle de Chimie Atmosphérique à Grande Echelle, *eng*: Model of atmospheric chemistry at the large scale) is a global chemical transport model of the atmosphere that is used in this thesis. It simulates air composition from the surface up to the mid-stratosphere and it includes both gases (Josse et al., 2004; Dufour et al., 2005) and aerosols (Martet et al., 2009). It is developed by Météo-France and it is used in operational applications and atmospheric research studies. Its applications extend to: air-quality forecasts, climate-chemistry interactions (Teyssèdre et al., 2007; Lamarque et al., 2013), desert aerosol studies (Martet et al., 2009), long-range transport pollution studies (Bousserez et al., 2007), "chemical weather" (Dufour et al., 2005), data assimilation of chemical species (e.g. El Amraoui et al., 2010), troposphere-stratosphere transport (Ricaud et al., 2009; Barré et al., 2012), etc.

2.2.1 General description

Coordinate systems For the horizontal grid, MOCAGE uses a regular normal longitude-latitude grid where gridboxes are aligned in a fixed geometric pattern in a form of rectangles. For the vertical grid, MOCAGE uses a hybrid sigma-pressure coordinate system (Fig. 2.1). This system solves the problems of the constant pressure or the constant altitude vertical coordinate systems, which appear when the planes of constant vertical coordinate are intercepted by the surface topography. This is achieved by following the so-called sigma (σ) coordinate near to the surface. The sigma coordinate follows the features of terrain. It is defined as (Phillips, 1957; Eckermann, 2009):

$$\sigma = \frac{p - p_{top}}{p_s - p_{top}}, \tag{2.1}$$

where p is the layer pressure $[Pa]$, p_{top} is the top layer pressure $[Pa]$ and p_s is the pressure at the surface $[Pa]$. Going upward, the hybrid sigma-pressure

coordinate system smoothly transits from the terrain following system, to the system that follows lines of constant pressure. Vertical levels are defined by two coefficients A and B which have an influence on pressure and terrain-following properties (Simmons and Strüfing, 1983). The coefficient A influences the isobaric properties, and the coefficient B the terrain-following properties. The pressure function is defined as:

$$p(\sigma, p_s) = A(\sigma) + B(\sigma)(p_s - p_{top}). \tag{2.2}$$

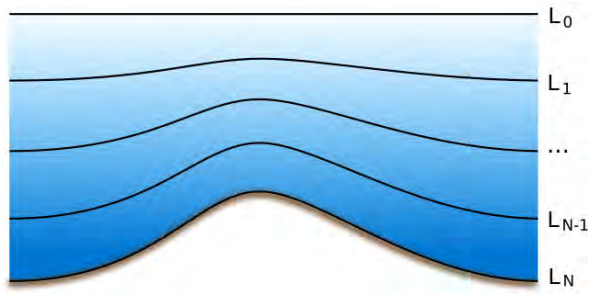

Figure 2.1: Illustration of the hybrid sigma-pressure vertical coordinate system. Close to the surface levels follow the terrain. High in the atmosphere levels follow isobaric lines.

There are 47 levels in MOCAGE from the surface up to about 5 hPa. Vertical resolution is not uniform. The levels are more packed closer to the surface, with a resolution of 40 m to 400 m in the planetary boundary layer and about 700–800 meters in the upper troposphere and lower stratosphere. This gives approximately 7 levels in the planetary boundary layer, 20 levels in the free troposphere, and 20 levels in the stratosphere. The vertical grid is the same in all possible domains that MOCAGE uses.

Nesting and spatial resolution MOCAGE is a global model with a possibility of *nesting* smaller domains inside larger domains. Smaller domains over the region of interest have a finer resolution than the larger domain, and the larger domain provides the necessary boundary conditions. Nesting in MOCAGE works in a two-way, where all domains influence each other. The global domain has a spatial resolution generally of 2°longitude×2°latitude, and several possible regional domains can have a resolution up to $0.2° \times 0.2°$ (for example over the Mediterranean sea, or the western and central Europe), and the domain over France, used for operational purposes, a resolution of $0.1° \times 0.1°$.

Meteorological forcing fields are not calculated in the model, but they come from the numerical prediction models (NWP) ARPEGE or IFS. ARPEGE

is the operational global (NWP) model of Météo-France, and IFS is the NWP model of ECMWF (European Centre for Medium-Range Weather Forecasts). ARPEGE is used for all global and large regional domains, while meteorological fields for the domain over France are provided by AROME – the high-resolution operational regional NWP model of Météo-France. New meteorological fields are introduced every three or six hours, and then they are interpolated for each hour. Although smaller timesteps are used in MOCAGE for physical processes, the meteorological variables are kept constant over the whole hour.

Lateral boundary conditions are interpolated and relaxed when passing from the larger to the smaller nested domain with a finer resolution. It means that variables in the nested domain gradually approach the values in the larger domain, to avoid strong gradients.

Initial conditions In MOCAGE, as in all CTM models, initial concentrations are necessary to replicate a realistic situation. Model configurations used in this thesis had initial concentrations that included only primary aerosols, without the chemistry scheme and without any interaction with gases. If initial concentrations from a previous MOCAGE simulation are not available, the climatological data can be used. Simulations without aerosol initial concentrations are possible, but in this case it is necessary to run the model sufficiently long to achieve the equilibrium. This period of time, necessary to achieve the equilibrium, is called a *spin-up* period. The spin-up period depends on the lifetimes of the considered species, and it is necessary if used the climatological data.

Initial conditions have more important role in short-term simulations to make an accurate prediction, than in long-term simulations. The dynamical processes reduce the effects of initial conditions during the time, and species with short lifetimes do not depend much on initial conditions if their lifetimes are much shorter than the simulation time. The sensitivity of a model to initial conditions decreases exponentially during the time, and with the same law the importance of the dynamical processes and emissions increases (Seinfeld and Pandis, 1998).

Advection For the large-scale transport of atmospheric species in the CTM MOCAGE, a semi-lagrangian advection scheme is used. Semi-Lagrangian scheme is a hybrid of two different approaches in the advection modelling: eulerian and lagrangian (Staniforth and Côté, 1991). To solve the advection equation

$$\frac{\partial n}{\partial t} + \nabla \cdot (\mathbf{v}n) = 0, \tag{2.3}$$

where n is the concentration of an atmospheric species, and $\mathbf{v}$ is the wind vector, the semi-lagrangian scheme uses a fixed grid like in the Eulerian method, and tracks back from where the particles came during the previous timestep like in the Lagrangian method (Fig. 2.2). By assuming that the concentration does not change during the advection, the concentration of species is obtained by calculating the position of a gridbox one timestep back. Using the known wind field, semi-lagrangian schemes track backward the particles from time t to time $t - \Delta t$ to find their departure points. The departure points are usually located somewhere in between gridboxes, and to find the concentration at that point it is necessary to do a spatial interpolation of concentrations of adjacent gridbox centers. The advected concentration at time t at point x will be equal to the concentration at time $t - \Delta t$ at the departure point:

$$n_{x,t} = n_{x-v\Delta t, t-\Delta t}. \tag{2.4}$$

Unlike explicit Eulerian schemes, semi-Lagrangian schemes are stable regardless of the size of the timestep and the spatial resolution (Jacobson, 2005). Still, long timesteps would produce large truncation errors. Another advantage is that the computational time is not proportional to the number of transported species because back-trajectories do not depend on the considered species and they are calculated only once for each timestep.

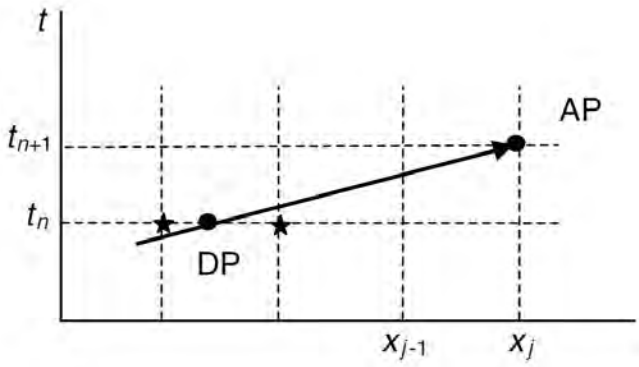

Figure 2.2: Illustration of the semi-lagrangian advection scheme. The particles are at point AP at time t_{n+1}. The advection scheme searches the point DP where the particles were at time t_n. The concentration at the point AP is equal to the concentration at DP, which is obtained by interpolation of neighboring gridpoints. The figure is adapted from Kalnay (2003).

Semi-Lagrangian schemes are not mass conservative. MOCAGE ensures the mass conservation between advection timesteps by normalization, so that the total mass that the scheme finds as a solution at the end of the timestep is equal to the total mass in the system at the beginning of the timestep. In this

way, the correction factor that is applied guarantees the conservation of the mass.

Compared to Eulerian schemes, semi-lagrangian schemes have a bigger numerical diffusion, where peak values are artificially spread. MOCAGE uses a third-order interpolation, which reduces the diffusion better than the interpolation methods of the smaller order (Randall, 2004).

The advection scheme in MOCAGE is dynamically forced by the wind field from the ARPEGE meteorological analysis or forecast, or from the IFS meteorological forecast or analyses.

Subgrid transport Two other transport processes, diffusion and convection, are subgrid-scale transport mechanisms. Such processes cannot be represented explicitly, and it is necessary to parameterize them.

Advection dominates in horizontal planes and for the model diffusion, horizontal planes are considered homogeneous. In this way, diffusion is applied only in the vertical axis:

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial z}\left(K_z \frac{\partial n}{\partial z}\right), \tag{2.5}$$

where K_z is the diffusion coefficient. Vertical diffusion is implemented following the parameterization of Louis (1979). In this scheme, the diffusion coefficients are calculated according to the estimation of the atmosphere stability using the Richardson number.

Convection is a powerful, but very localized process. Usually, only a small part of the gridbox is affected by convection. It is characterized by a strong upward motion inside clouds, a slow downward motion around them, and air entrainment and detrainment on the sides. In MOCAGE, transport by convection is simulated by the parameterization proposed by Bechtold et al. (2001), which is a unidimensional parameterization triggered only if convective clouds are detected in an atmospheric column.

2.2.2 Aerosols modelling in MOCAGE

To simulate aerosols the model has to solve the following equation:

$$\frac{\partial n}{\partial t} = \left(\frac{\partial n}{\partial t}\right)_{advection} + \left(\frac{\partial n}{\partial t}\right)_{diffusion} + \left(\frac{\partial n}{\partial t}\right)_{convection} + \left(\frac{\partial n}{\partial t}\right)_{sedimentation} + \left(\frac{\partial n}{\partial t}\right)_{dry\_deposition} + \left(\frac{\partial n}{\partial t}\right)_{wet\_deposition} + \left(\frac{\partial n}{\partial t}\right)_{coagulation} + \left(\frac{\partial n}{\partial t}\right)_{nucleation} + E + S, \quad (2.6)$$

where $\frac{\partial n}{\partial t}$ is the change of the aerosol concentration with time: the total change on the left-hand side of the equation, and changes due to different aerosol processes on the right-hand side of the equation; E is the emission rate; and S is the production and loss of the secondary aerosols.

To solve this equation, MOCAGE separates the different processes and solves them sequentially with a unique numerical scheme for each of them. This approach is called *operator splitting*. From Eq. (2.6) it is clear that an aerosol model have a lot of different operators and schemes. Each of the operator can have different timesteps and work independently during one time interval which is a common timestep when results of all processes are combined. In MOCAGE, the timestep of 15 min is used for the operators of emission, sedimentation and surface dry deposition, the timestep of 30 min is used for the operators of cloud cover, convection and wet deposition, and the timestep of 1 h, which is the common timestep, is used for the advection scheme.

Representation and types Aerosols in MOCAGE are represented by the sectional (discrete) approach (see Section 1.5.1). The size distribution of each aerosol type in the model is divided into 6 bins. This discretization is done for each size distribution of each aerosol type. Equation (2.6) has to be solved for each bin. The discretization is done in the way to have an optimal representation of the different sizes of the particles and not to be too computationally expensive. Tests with different numbers of bins were performed (6, 10 and 20 bins per aerosol type) and it was concluded that six bins per type are the optimal choice and compromise (Martet, 2008). With the sectional approach, different types of aerosols are represented adequately, but the mixing between different type of aerosols is not simple to manage.

In MOCAGE, there are five aerosol types, all of them primary: desert dust, sea salt, black carbon (BC), organic carbon (OC), and volcanic dust

aerosols. All these aerosols are considered as homogeneous, and they only externally mix in the model. In other words, they are considered as passive tracers. In total there are $6 \cdot 5 = 30$ bins which are treated as separate variables. This means that aerosols in each bin are emitted, transported and removed from the atmosphere, and they do not chemically or physically interact with other bins or gases. In MOCAGE, the schemes of coagulation and chemical reactions which form secondary aerosols are in the phase of implementation and tests, but they are not considered in this thesis (this is the development from the ongoing PhD thesis of Jonathan Guth at CNRM/Météo-France).

Bins are chosen considering the size range of each type of aerosol and the subsequent evolution of their size distribution. Also, two of the bins have their limits at 2.5 µm and 10 µm for practical air quality purposes in order to easily integrate the sum of $PM_{2.5}$ and PM_{10} particles. The exact values of the bin limits and their mean values are given in Table 4.4. Initial size distributions are considered to be log-normal. In the model, the initial desert dust distribution is represented as a superposition of 3 modes, while sea salt, black carbon and organic carbon initial distributions as a superposition of 2 modes. Initial size distribution are distributed into the defined model bins.

Emissions Emissions of primary aerosols in MOCAGE are implemented in two possible ways. For all aerosol types present in MOCAGE it is possible to describe emissions by various fixed inventories. Emission rates in inventories are derived from models, satellite, ground and field campaign observations and are estimated for specific time with specific spatial and temporal resolution. They are issued by various international projects like AeroCom, IPCC, ACCMIP. These inventories usually describe monthly emitted quantities with a horizontal resolution up to $0.5° \times 0.5°$. Biomass burning aerosol emissions are prepared separately into inventories by projects like GFED or GFAS, which have better spatial and temporal resolution, up to daily quantities with a resolution up to $0.1° \times 0.1°$. In some nested domains of MOCAGE, some emissions may have a coarser resolution than the model. This can introduce uncertainty of the representation of a small-scale features for such aerosols, but it does not influence the representation of large features.

For emissions that depend on mechanical and dynamical processes, like saltation for dust and bubble bursting for sea salt aerosols, special dynamical emission parameterizations exist. These dynamical parameterizations in MOCAGE are implemented for sea salt and desert dust aerosols. They use meteorological variables and surface data to calculate emission rates for specific points and times, and their accuracy exceeds accuracy of emission rates from

inventories. Their quality largely depends on the quality of the used meteorological and surface data. In MOCAGE, the ARPEGE or IFS meteorological fields are used as the input fields for the schemes of dynamical aerosol emissions.

Dynamical emission parameterizations deal better with initial size distribution estimation than fixed inventory data. A significant, both spatial and temporal, variability of the emitted distribution can be taken into account, especially if only one parameter affects it such as the wind speed in the case of dynamical sea salt emission parameterization.

In order to help the model to emit aerosols from the surface, at the time of emission aerosols are injected above the turbulent surface layer and distributed along the first five model levels. In this way, also strong gradients in aerosol concentration are avoided, which ensures the stability of the advection scheme.

Deposition and transformation schemes Aerosol deposition processes are represented in MOCAGE with: an explicit sedimentation scheme, and parameterized schemes for the dry deposition at surface, the in-cloud and below-cloud scavenging. Regarding the physical transformations of aerosols, only the hygroscopic particle growth is presented and described with a parameterized scheme. Coagulation and nucleation are not considered, since it is estimated that these processes are lot less important for primary than secondary aerosols.

Many of these schemes and parameterizations were subject of revision, development, sensitivity tests and validation during this PhD thesis. They are presented in more details in Chapter 4. The specific focus was given to the improvement of sedimentation, emission and wet deposition parameterizations.

Uncertainties Uncertainties in MOCAGE are uncertainties that are linked to all aerosol models (Section 2.1.1). Uncertainties among aerosol-related processes in state-of-the-art models are probably largest for emissions and wet deposition processes (Textor et al., 2006, 2007; Lee et al., 2011; Vignati et al., 2010; IPCC, 2013). Other big sources of uncertanties can be: the quality of meteorological and input data which can influence transport, dynamical emissions and other schemes; subgrid parameterizations (transport and cloud processes, Mahowald et al., 1995); discretization and modelling approach; numerical errors and bugs, etc. (Sportisse, 2008).

3

Data Assimilation

This chapter describes the theoretical basis and developments in the field of data assimilation.

Contents

3.1 Definition

Data assimilation is a technique with which it is possible to reduce uncertainties and to improve the model output by incorporating the information of real measurements related to the modelled variables (Kalnay, 2003). A data assimilation system searches for the state of the system which corresponds to the model and observations in the best optimal way. It takes into account the uncertainties of both the model and the observations. Then, the corrected assimilated state of the system can be used as the input of the new assimilation process. This process is schematically presented in Fig. 3.1.

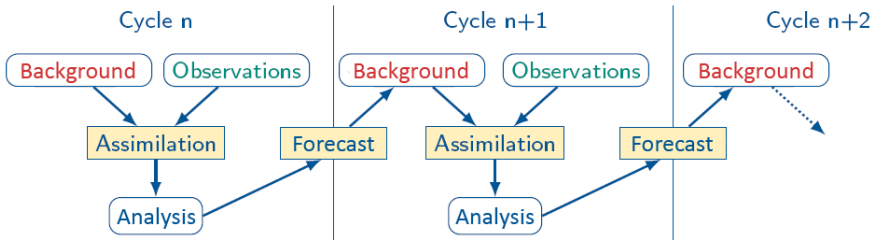

Figure 3.1: Illustration of data assimilation process. The assimilation system uses the background and the observations to find an optimal analysis. The model then calculated the improved forecast, which is used as the background in a subsequent assimilation cycle.

Using a mathematical formulation, we try to find the state of the system x which should be the best estimate of the true state x^t. The best estimate of the true state x^t is called the analysis x^a and it is the result of the data assimilation process. To find the analysis, it is necessary to consider an *a priori* estimation of the state of the system x^b, which is called *background*. The background x^b is usually the forecast of a model. It is also known as *first guess*, or *a priori*. The error of the background is defined as the difference between the background and true state:

$$\varepsilon^b = x^t - x^b. \tag{3.1}$$

The observations are described in the observation vector y. The number and spatial distribution of the observations can be very variable. This influences the quality of the analysis which directly depends of the quantity of the available information. Of the same importance is the quality of the observations. The error of observations is, analogously to Eq. (3.1):

$$\varepsilon^o = y^t - y, \tag{3.2}$$

where y^t is the true state. The observation error describes the instrumental errors, precision, possible retrieval errors, etc.

In the data assimilation process, it is supposed that errors of the background and the observations are not biased. This means that the average of the background errors and the observation errors should be zero, i.e. that there are no systematic errors:

$$E(\varepsilon^b) = 0$$
$$E(\varepsilon^o) = 0.$$

where E denotes expectation operator.

If the bias exists between the observations and the background, techniques of the bias correction should be employed (Dee and Uppala, 2008).

Also, background and observation errors should be mutually uncorrelated:

$$E(\varepsilon^b(\varepsilon^o)^{\mathrm{T}}) = 0.$$

Instrumental errors are usually not correlated, but correlation could appear during the observation retrieval and/or pre-processing, or when some features are observed which cannot be represented in the model because of too coarse resolution.

3.2 Approaches

The goal of data assimilation is to find an analysis which is the best estimation of the true state in terms of minimizing squares of background and observation errors. There are two approaches to data assimilation to achieve this: sequential and variational methods. In sequential methods, such as Kalman filter, all observations that come from the present and past can be used ("direct observer"). The states of the system are studied in order to find the state that statistically corresponds the best to the available observations, taking into account the errors of the model and the observations. In this way, every time when an observation is available it is injected into the assimilation system to estimate the corrected state of the system. This approach is adapted for real-time data assimilation systems.

In variational methods, all observations measured in the same assimilation cycle can be used. This can include observations measured before and after the exact moment at which the analysis is searched ("dynamical observer").

The optimal state of the system is found by minimizing the so-called *cost function*, which describes the misfit between the model and the observed data. The minimization can be described as an optimal control problem (Talagrand, 2010), and not a statistical one, but it requires the knowledge of the statistics of model errors and observation errors.

3.3 Variational methods

The most frequently used variational methods are 3D-Var, 4D-Var and 3D-FGAT.

3.3.1 3D-Var

3D-Var (3Dimensional VARiational assimilation) is a variational method of assimilation where the goal of estimating the most plausible true state of the system is achieved from the known statistical characteristics of the errors of the observations and the background. This method, in each assimilation cycle, collects all observations which are measured in the time period of the same cycle and compares them with the model at one given moment (usually at the beginning of the cycle) and not at the time of the observations (Courtier et al. (1998), Fig. 3.2).

The least square errors between the estimated true state (the model prediction) and the available informations (the observations) are reduced by the minimization of the cost function J:

$$\begin{aligned} J(x) &= \frac{1}{2}(x - x^b)^{\mathrm{T}}\mathbf{B}^{-1}(x - x^b) + \frac{1}{2}\sum_{i=0}^{N}(y_i - H_i x)^{\mathrm{T}}\mathbf{R}_i^{-1}(y_i - H_i x) \\ &= J_b(x) + J_o(x) \end{aligned} \tag{3.3}$$

where x is the state of the system at a given time; x^b is the background state of the model; y_i is the observation at time t_i where N is the total number of timesteps; H_i is the observation operator. It allows to pass from the model space to the observation space, in the way that the model prognostic variable x is transformed by $H_i x$ to the variable equivalent to the observation y_i and then it can be directly compared with it. If all observations are of the same type, then H will be the same for all of them. $\mathbf{B}$ is the background covariance matrix; and $\mathbf{R}_i$ is the observation covariance matrix at time t_i. The first term

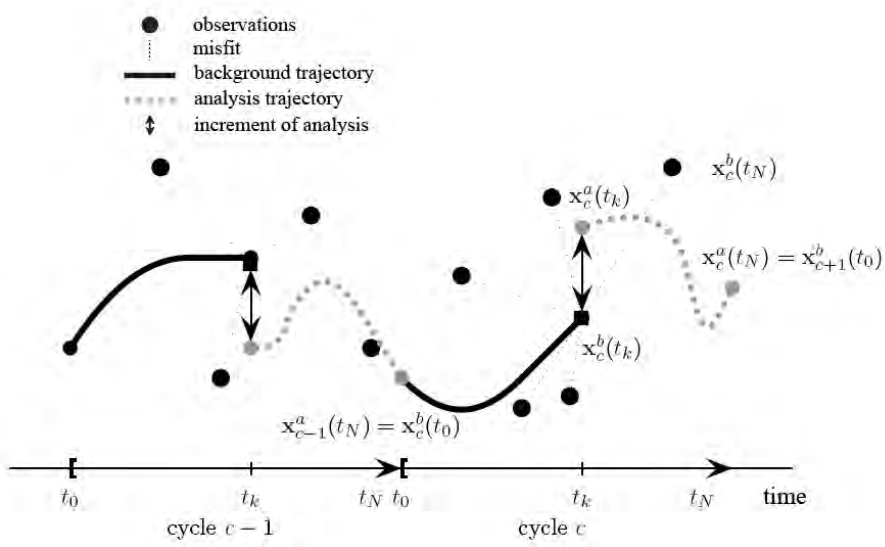

Figure 3.2: Illustration of 3D-Var data assimilation method. In each cycle, the model propagates from t_0 to t_N to calculate the background trajectory. Misfits between the background and the observations are calculated taking into account only one value of the background at time t_k (which can be located anywhere in the cycle). The assimilation system searches for the optimal analysis at time t_k from where the model propagates to calculate the forecast for the time t_N at the end of the cycle. The forecast $x_{c-1}^a(t_N)$ is used as the initial background $x_c^b(t_0)$ of the new cycle. The figure is from Daget (2008).

on the right-hand-side of Eq. (3.3) (J_b) represents the misfit to the background and the second term (J_o) represents the misfit to observations. The matrices $\mathbf{B}$ and $\mathbf{R}_i$ influence the weighting of these two terms, because the errors of the background and observations are represented by $\mathbf{B}$ and $\mathbf{R}_i$ respectively. $\mathbf{B}$ and $\mathbf{R}_i$ are positive-definite and symmetric matrices; on their diagonals are located variances, and off-diagonals covariances of all model gridpoints and the observations, respectively.

To minimize the cost function it is necessary to find its gradient, which in the case of 3D-Var is:

$$\nabla J(x) = \mathbf{B}^{-1}(x - x^b) + \sum_{i=0}^{N} \mathbf{H}_i^{\mathrm{T}} \mathbf{R}_i^{-1}(y_i - H_i x), \tag{3.4}$$

where the operator $\mathbf{H}$ is the linearised version (tangent-linear) of the operator H, and $\mathbf{H}^{\mathrm{T}}$ is its transposed matrix. In order to define a linearized operator, we can suppose that small variations of the background x^b in its vicinity produce linear variations of Hx, i.e. if x is sufficiently close to the background x^b then

it is true that

$$Hx - Hx^b \approx \mathbf{H}(x - x^b). \tag{3.5}$$

The optimal state of the model is obtained by finding the value of x for which:

$$\nabla J(x) = 0. \tag{3.6}$$

Note that, there is a large number of variables in the model, and the minimization of the cost function is not performed in the model space, but in the control space. Control variables which resides in this space are variables that are actually minimized, and they are not necessary prognostic variables of the model. The choice of the control variable directly influences the numerical efficiency of the system and its versatility (how can handle and which different types of observations).

Incremental 3D-Var In the incremental approach, it is not the variable x which is minimized, but instead, the misfit, δx, between the background and the state of the system:

$$\delta x = x - x^b. \tag{3.7}$$

The advantage of this formulation is the possibility to use the tangent-linear operator $\mathbf{H}$ within the cost function J. This is possible because by definition, δx should be small and near to the background state. This is the computationally more efficient approach, because linearised operators are often simpler to implement (Courtier et al., 1994). Also, if the observation operator is linear, the cost function is quadratic and is guaranteed to have a unique minimum (Lahoz et al., 2011).

The cost function has a form:

$$\begin{aligned} J(\delta x) &= J_b(\delta x) + J_o(\delta x) \\ &= \frac{1}{2}(\delta x)^{\mathrm{T}}\mathbf{B}^{-1}(\delta x) + \frac{1}{2}\sum_{i=0}^{N}(d_i - \mathbf{H}_i\delta x)^{\mathrm{T}}\mathbf{R}^{-1}(d_i - \mathbf{H}_i\delta x), \end{aligned} \tag{3.8}$$

where $d_i = y_i - H_i x^b$ is called the innovation and represents the distance of the observation y_i from the background x^b. The cost function will have the form:

$$\nabla J(\delta x) = \mathbf{B}^{-1}(\delta x) + \sum_{i=0}^{N}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}(d_i - \mathbf{H}_i\delta x). \tag{3.9}$$

After the estimation of the analysis increment δx^a at one specific moment in the assimilation cycle, the analysis is calculated by

$$x^a = x^b + \delta x^a. \tag{3.10}$$

After the estimation of the optimal state of the model at the specific moment, the model can find a new trajectory during this cycle. The departure point of a new cycle is usually the result of this new model run started from the analysis.

The 3D-Var method has a limitation: it does not take the observations at their proper time of measurement and it looses information of the series of subsequently taken observations during a single assimilation cycle. This is treated in the 4D-Var method.

3.3.2 4D-Var

The 4D-Var method is an extension of 3D-Var. Instead of finding an optimal state of the system at the specific time in the assimilation cycle, in 4D-Var we try to find an optimal trajectory of the system during the whole assimilation cycle. All observations are taken into consideration at their proper time of measurement. Then, the information given by every observation is propagated in time. In this way, an information of the system evolution, given by a series of observations measured at different times of the assimilation cycle, is also included in the estimation of analysis (Courtier et al., 1994).

The cost function of 4D-Var is:

$$\begin{aligned} J(x) =& \frac{1}{2}(x - x^b)^{\mathrm{T}}\mathbf{B}^{-1}(x - x^b) + \\ & \frac{1}{2}\sum_{i=0}^{N}(y_i - H_i x(t_i))^{\mathrm{T}}\mathbf{R}_i^{-1}(y_i - H_i x(t_i)). \end{aligned} \tag{3.11}$$

The state of model $x(t_i)$ at the time t_i is calculated by the model operator M (which is the model itself) which propagates the state of system x from time t_0 to time t_i:

$$x(t_i) = M_{0\to i}x(t_0). \tag{3.12}$$

In this way, the cost function becomes:

$$\begin{aligned} J(x) =& \frac{1}{2}(x - x^b)^{\mathrm{T}}\mathbf{B}^{-1}(x - x^b) + \\ & \frac{1}{2}\sum_{i=0}^{N}(y_i - H_i M_{0\to i}x)^{\mathrm{T}}\mathbf{R}_i^{-1}(y_i - H_i M_{0\to i}x). \end{aligned} \tag{3.13}$$

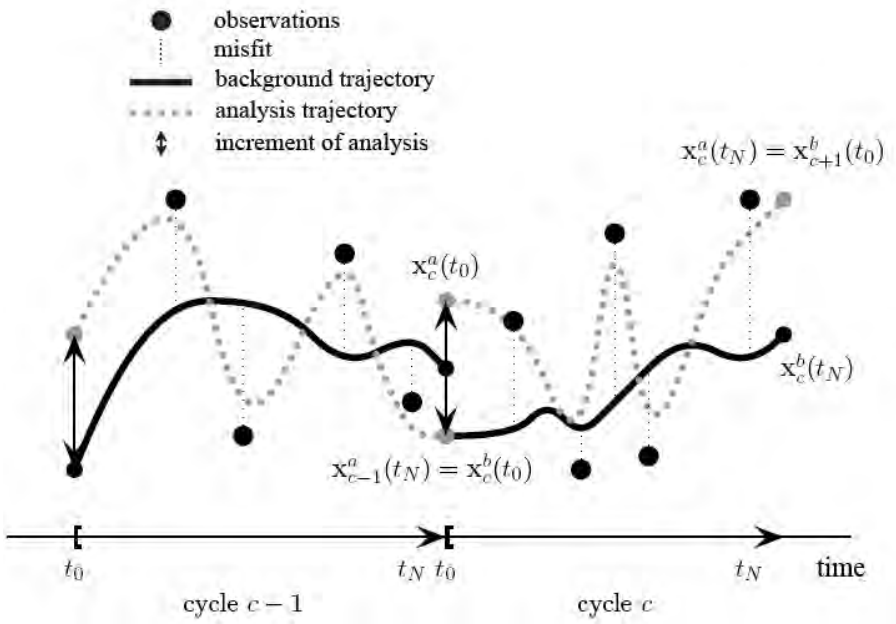

Figure 3.3: Illustration of 4D-Var data assimilation method. In each cycle, the model propagates from t_0 to t_N to calculate the background trajectory. Misfits between the background and the observations are calculated at their appropriate times of observation. The assimilation system searches for the optimal increment along the whole model trajectory by using the tangent-linear and adjoint operators of the model. The forecast $x_{c-1}^a(t_N)$ is used as the initial background $x_c^b(t_0)$ of the new cycle. The figure from Daget (2008).

Incremental 4D-Var In the incremental form of 4D-Var, analogue as in 3D-Var, the cost function is defined in terms of the increment $\delta x = x - x^b$. The model operator M is linearised ($\mathbf{M}$) for the propagation of the increment δx, but the non-linear version (the model itself) is still used to propagate the background:

$$\begin{aligned} M_{0\to i}x &= M_{0\to i}(x^b + \delta x) \\ &\approx M_{0\to i}(x^b) + \mathbf{M}_{0\to i}\delta x \end{aligned} \tag{3.14}$$

Using this relation, the cost function of the incremental 4D-Var method is:

$$\begin{aligned} J(\delta x) =& \frac{1}{2}\delta x^{\mathrm{T}}\mathbf{B}^{-1}\delta x + \\ & \frac{1}{2}\sum_{i=0}^{N}(d_i - \mathbf{H}_i\mathbf{M}_{0\to i}\delta x)^{\mathrm{T}}\mathbf{R}_i^{-1}(d_i - \mathbf{H}_i\mathbf{M}_{0\to i}\delta x), \end{aligned} \tag{3.15}$$

where $d_i = y_i - H_i x^b(t_i) = y_i - H_i M_{0\to i} x^b$ is the innovation.

The gradient of the cost function is:

$$\nabla J(\delta x) = \mathbf{B}^{-1}(\delta x) + \sum_{i=0}^{N} \mathbf{M}_{0\to i}^{\mathrm{T}} \mathbf{H}_i^{\mathrm{T}} \mathbf{R}_i^{-1} (d_i - \mathbf{H}_i \mathbf{M}_{0\to i} \delta x). \tag{3.16}$$

In the 4D-Var method, the increment is propagated using the tangent-linear versions of the model and the observation operator. The gradient is calculated using the adjoint of these two operators. At the end of the minimization, the calculated increment is added to the background, and the model is run for the whole assimilation cycle taking into account the calculated increment in order to get the assimilated trajectory.

Usually, the development of the tangent-linear and adjoint versions of a model are the most difficult points in the development of the 4D-Var system. The reason is that the atmospheric models are very complex and non-linear. One of the possible solutions is the 3D-FGAT assimilation method, which takes observations at their proper times.

3.3.3 3D-FGAT

3D-FGAT (3-Dimensional First Guess at Appropriate Time) is a compromise between 3D-Var and 4D-Var (Fisher and Andersson, 2001; Massart et al., 2010). Observations are taken at their appropriate times when they are measured, i.e. every measurement is compared with the background at the time of measurement, like in 4D-Var. But, the optimal analysis is estimated still only for a specified moment in the assimilation cycle, like in 3D-Var, and not for whole trajectory as in 4D-Var. In this way, during the assimilation, we do not use the linearised operator of the model evolution and its adjoint. This is defined by the cost function, whose incremental form in 3D-FGAT is:

$$\begin{aligned} J(\delta x) =& \frac{1}{2} \delta x^{\mathrm{T}} \mathbf{B}^{-1} \delta x + \\ & \frac{1}{2} \sum_{i=0}^{N} (d_i - \mathbf{H}_i \delta x)^{\mathrm{T}} \mathbf{R}_i^{-1} (d_i - \mathbf{H}_i \delta x), \end{aligned} \tag{3.17}$$

where $d_i = y_i - H_i x^b(t_i)$ is the innovation, and to calculate the innovation, it is necessary to find the state of background at the time t_i: $x^b(t_i) = M_{0\to i} x^b$.

The gradient of the cost function is calculated with the equation:

$$\nabla J(\delta x) = \mathbf{B}^{-1}\delta x + \sum_{i=0}^{N} \mathbf{H}_i^{\mathrm{T}} \mathbf{R}_i^{-1}(d_i - \mathbf{H}_i \delta x). \tag{3.18}$$

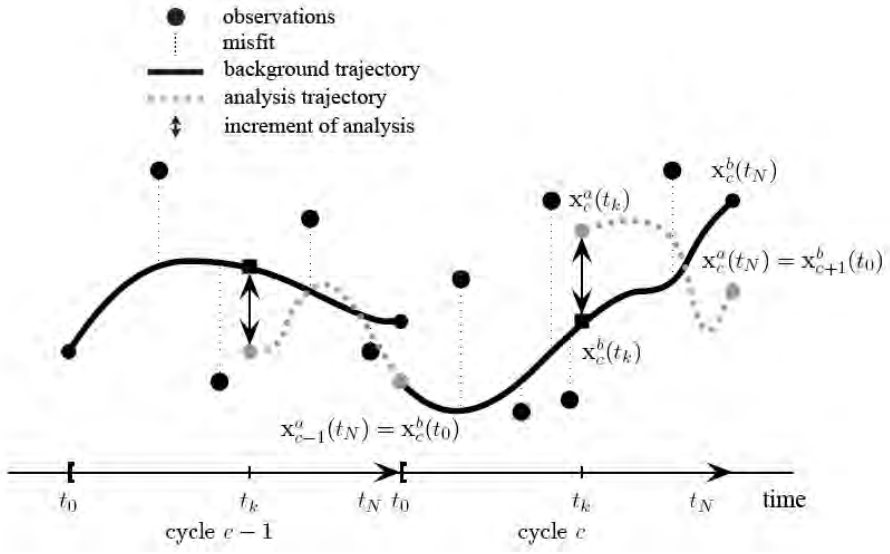

Figure 3.4: Illustration of 3D-FGAT data assimilation method. In each cycle, the model propagates from t_0 to t_N to calculate the background trajectory. Misfits between the background and the observations are calculated at their appropriate times of observation. The assimilation system searches for the optimal analysis at time t_k (which can be located anywhere in the cycle) from where the model propagates to calculate the forecast for the time t_N at the end of the cycle. The forecast $x_{c-1}^a(t_N)$ is used as the initial background $x_c^b(t_0)$ of the new cycle. The figure from Daget (2008).

After the estimation of the gradient, the optimal analysis increment δx^a is added to the background at the specific point of time, usually at the beginning of the cycle (in Fig. 3.4 is added in the middle of the cycle). From there the model is run to obtain the analysed trajectory. Its endpoint is used as a departure point for the background of the next cycle.

3.4 Observational operator

As seen before, the observation operator H, generally non-linear, enables a model output to be compared with the observations. It is done by computing

a model equivalent of the observations in the observation space (in terms of measured physical quantity).

Observational operators do a variety of operations of different complexity. In the simplest case, where the observed quantity is the same as the model control variable, the observation operator does only interpolation (spatial interpolation in the case of 3D-VAR, and spatial and temporal interpolation in the case of 3D-FGAT and 4D-Var). In more complex cases, when the observed and model control variables are not the same, the observation operator does additional transformations to go from the model space to the observation space.

Observational operators can be considered as the sequence and the multiplication of a number of sub-operators

$$H = H_1 H_2 H_3 \dots H_U, \tag{3.19}$$

where each of sub-operators performs a part of the transformation from the background to the observation.

Observations compared with a model output can be quantities that are directly measured (radiances from satellites, backscatter signal from lidars, light intensity from photometers, etc.). Another possibility is that observations are already transformed before the assimilation to a physical quantity that is closer to the model space. In this case the observation operator will be simpler. This can be a important advantage because, as seen in Section 1.9.1, the retrieval of the aerosol quantities can be a difficult task. The use of already retrieved products from data providers in data assimilation brings also some drawbacks. We directly introduce in the assimilation system all the hypotheses and assumptions made in the retrieval process without the possibility of evaluating them, and the use of other aerosol or NWP models in the retrieval can introduce correlated errors (Migliorini, 2012).

These issues are not present if assimilating directly the observed quantity. On one hand, this approach is well grounded especially because the retrieval can be seen as a part of the variational estimation problem. On the other hand, it is computationally expensive and difficult. It is necessary to implement relatively complicated radiative transfer models and the characteristics of measuring instruments in the observation operators.

The calculation of the model equivalent by the observation operator can introduce some errors in the assimilation process. For example, the observation operator will not give exactly the same output if we use different interpolation methods, or interpolate before or after applying the other (physical) part of the observation operator, or use the nearest neighbor method before or after

applying the operator. This kind of error adds up, and represents the so-called "representativeness" error in the assimilation. It is considered in the observation covariance matrix $\mathbf{R}$. The main cause is limiting model resolution, and from this it takes its name.

3.5 Tangent-Linear and Adjoint operators

The tangent-linear is a linearised version of the non-linear observation operator. It gives a first-order approximation of the difference between the unperturbed (Hx) and the perturbed results ($H(x+\Delta x)$) of the non-linear operator H. It is developed by linearising the non-linear operator around the state of the system x. The tangent-linear operator consists of partial derivatives of the non-linear observation operator with respect to all U variables which are used as input to it:

$$\delta H = \sum_{i}^{U} \frac{\partial H}{\partial x_i} \delta x_i, \tag{3.20}$$

where δx_i are perturbations of all input variables. Equation (3.20) gives a first-order approximation of the output perturbation. It is valid if H is differentiable, if all derivatives exist around the state of system x, and if all input perturbations are small.

The tangent-linear operator can be derived by constructing the explicit matrix composed of all partial derivatives. This requires at least $U + 1$ total operator runs by using the finite-difference method for each derivative. But, also the tangent-linear operator can be considered as a sequence of linearised sub-operators of the non-linear observation operator:

$$\mathbf{H} = \mathbf{H}_1 \mathbf{H}_2 \mathbf{H}_3 \ldots \mathbf{H}_U. \tag{3.21}$$

In this way, the tangent-linear operator is built piece by piece, and it consists of simple and elementary calculations. This approach is convenient for testing as well, because it allows us to test the parts of the code separately. The non-linear operator is linearised by differentiation of the each line of code or loop separately.

A similar approach applies to the adjoint operator. This operator is the transpose of the tangent-linear operator, and for every adjoint operator, the subsequent expression has to be true:

$$\langle \mathbf{H}x, y \rangle = \left\langle x, \mathbf{H}^{\mathrm{T}} y \right\rangle, \tag{3.22}$$

where $\langle x, y \rangle$ represents the inner product of x and y. Also, the adjoint operator can be considered as a sequence of operators:

$$\mathbf{H}^{\mathrm{T}} = \mathbf{H}_U^{\mathrm{T}} \mathbf{H}_{U-1}^{\mathrm{T}} \dots \mathbf{H}_2^{\mathrm{T}} \mathbf{H}_1^{\mathrm{T}}. \tag{3.23}$$

Every discrete operation in the tangent linear operator (and non-linear forward operator) has a corresponding operation in the adjoint operator, but the order of execution is reversed. Adjoint operators are used in data assimilation in the calculation of the gradient of the cost function. But, they have other applications as well, for example studying the sensitivity of the model output with respect to its input.

3.6 Application to aerosols

In the field of air quality, assimilation of satellite and ground-based observations confine and improve forecasts of the models. It can make better understanding of related atmospheric processes which overall improve models and their estimates of the state of the atmosphere and its constituents. All this, when applied to the aerosol modelling can bring important benefits. Data assimilation has been applied for a long time in numerical weather forecasting, and later introduced for atmospheric chemistry. It is now emerging for aerosols as well. Aerosol data assimilation efforts are mainly focused on assimilating satellite data, usually aerosol optical depth (AOD). Many of AOD assimilation studies use variational data assimilation techniques. For example, 3D-VAR system for assimilating the AOD data was built by Zhang et al. (2008) in the NLR model, by Niu et al. (2008) in the CUACE/Dust model, by Liu et al. (2011) in the NCEP WRF-Chem/GOCART system. Benedetti et al. (2009) described the assimilation of AOD in the ECMWF Numerical Weather Prediction model using 4D-VAR method. There are also other assimilation approaches which, for example, use an Ensemble Kalman filter (Schutgens et al., 2010).

Another type of aerosol observations which is interesting for aerosol data assimilation is lidar profiles. The spaceborne CALIPSO lidar does not have a large spatial coverage as AOD measurements, but the assimilation of its data can provide valuable information on the vertical distribution of aerosols. In recent years there are several studies on lidar assimilation: regarding CALIPSO data assimilation (Sekiyama et al., 2010; Zhang et al., 2014) or the impact of assimilation of a ground lidar network data on the air quality modelling (Wang et al., 2014).

Part II

Results

4

Developments in the CTM MOCAGE

This chapter describes the developments done of the aerosol module in the CTM MOCAGE.

Contents

4.1 Developments

It is shown in Chapter 2 that aerosol CTM models undergo constant improvements due to better understanding of aerosol processes and better availability of computational power. This chapter describes the developments in CTM MOCAGE done during this PhD thesis.

The aerosol module of MOCAGE was first developed at Météo-France during Maud Martet's PhD thesis (Martet, 2008). The emphasis was given to the studies of desert dust aerosols, with a first validation study of the same subject (Martet et al., 2009). At that time, sea salt and anthropogenic primary aerosols were also included in the model.

As the first scope of this thesis, we made an evaluation of the performance of the aerosol module of MOCAGE including all aerosol species so far implemented. During this process, a number of the model schemes underwent improvements. The further developments aim to reduce known biases compared with observations and to obtain a more physically detailed representation of the processes. These developments include:

- implementing a different dynamical sea salt emission scheme that includes the dependence on the sea surface temperature
- improving the dynamical desert dust emission scheme in terms of improved input data and more realistic initial size distributions
- improving the sedimentation scheme to ensure its mass conservation
- improving, implementing and refining various components of the wet deposition scheme
- adding volcanic dust aerosols as an additional primary aerosol type in the model

Also, a different scheme was implemented to calculate the optical properties of aerosols using the Mie theory. This scheme will be used in the validation of the new developments for the comparison with the observations. Later, it also served in the aerosol data assimilation which is explained in the Chapter 5.

From now on, we will often make a reference to the state of the model before the developments described in this thesis. That version of the model will be referred as the **standard** configuration. Similarly, the model state that follows the new developments as they are described we will refer as the **updated** configuration.

4.2 Volcanic aerosols

Volcanic ash aerosols can strongly affect civil aviation and economical activities. The eruption of Eyjafjöll volcano in 2010 provoked significant economical effects and increased the interest for the volcanic ash modelling and observation. We included the representation of volcanic ash aerosols in the updated configuration of MOCAGE. The main difference compared to other modelled aerosol types in the model is that volcanic ash is released from point sources, which needs to be differently represented in the model.

At the beginning of the work on this thesis, the volcanic ash transport was already present in the special version of MOCAGE – *MOCAGE-Accident*, which works as the operational dispersion model at Météo-France. MOCAGE-Accident is used for the dispersion of atmospheric accidental releases of gases and particles, for example radioactive gases or volcanic ash. Also, it is optimized for the performance because of the necessity of fast response and making alerts: it excludes chemistry, uses simpler parameterizations for some physical processes, optimized interpolation for advection, etc.

Point-source definition In the updated configuration of MOCAGE, the volcanic ash transport was implemented following MOCAGE-Accident. The point source of aerosols is defined by the location and the time of emission, the emission duration, the emitted mass and size distribution, the volcanic ash mass density and the height of the base and the top of plume. The emitted mass and the height of the top of the plume depend of each other. With known plume heights obtained from observations, the emitted dry rock mass can be estimated by empirical relationships (Sparks et al., 1997; Mastin et al., 2009). We use the relation by Mastin et al. (2009) where the emitted mass rate [$\mathrm{kg\,s^{-1}}$] is converted to the emitted dry rock volume rate V [$\mathrm{m^3\,s^{-1}}$] and connected with the plume height H [km] as

$$H = 2V^{0.241}. \tag{4.1}$$

The plume height can be obtained by: radar measurements (Lacasse et al., 2004); directly by visual observations from the ground or an airplane; satellite observations of cloud top brightness temperature (Woods and Self, 1992) or lidar profiles; isopleth patterns (Carey and Sparks, 1986). For important eruptions, the plume heights are reported by VAAC (Volcanic Ash Advisory Centers).

To estimate the size distribution of the volcanic ash in the plume it is necessary to use data from sampled ash deposits near the eruption. But, only

the largest particles fall down quickly after the emission. It is necessary to estimate the size distribution in the model which will reproduce the sampled deposed distribution near the source (Stohl et al., 2011).

In MOCAGE, the mass of the plume over the vertical levels is distributed uniformly. Yet, in the horizontal plane, the source is not represented in a single gridbox, but in the square of 4×4 gridboxes. In this way, the advection scheme will not likely encounter strong gradients which can provoke numerical instabilities.

It is also considered that the plume is influenced by wind. In order to calculate weights for the mass repartition in the 4×4 gridbox square, we use the normal distribution perpendicularly to the wind direction with the peak of the distribution that is at the source of the eruption. In the direction of the wind, we consider the skewed normal distribution with the skewed part located downwind. The degree of skeweness depends on the wind speed in each level.

Uncertainties in the volcanic ash modelling lies mainly in the source term. One of the possible ways of improving the representation of volcanic ash aerosols in models is the introduction of aerosol data assimilation in the model. This can considerably lower uncertainties in the extent and concentrations within volcanic plumes, especially because it is possible to 'follow' the plume by observations over multiple days.

The advantages of having volcanic ash aerosols represented in MOCAGE, besides MOCAGE-Accident, concerns various applications. First, in MOCAGE, volcanic aerosols are represented physically in more details because there is less constrain on computation time of the simulations. Second, in MOCAGE are also represented other aerosol types, besides the volcanic ash. This is important for studies of impacts of the volcanic aerosols on, for example, air quality. Also, the aerosol observations usually include all aerosol types mixed, without easy possibility to distinguish between them. For comparison with observations it is important to have all aerosol types in the model. Third, in MOCAGE it is possible to assimilate volcanic aerosols with the developed assimilation module which is not possible in a fast-response system such as the MOCAGE-Accident. All this suggest that the combined usage of MOCAGE and MOCAGE-Accident could have beneficial effects.

The evaluation of the modelled volcanic plumes for the case of 2010 eruption of Eyjafjöll on Iceland is presented in Section 5.3.4. The impact of the data assimilation to the quality of the modelled volcanic ash field is also evaluated.

4.3 Mie code

To calculate the aerosol optical depth in the model, first it is necessary to obtain the extinction efficiency Q_{ext} of all different aerosol particles (bins) in the model (Section 1.8.2.1). For that, it is necessary to calculate the complex Mie coefficients which depend on x, and the complex refractive index $\tilde{n}$ (Van de Hulst, 1981). The coefficients are defined in terms of Ricatti-Besell functions and can be only solved numerically. For these time-consuming computations, many schemes are developed, which often differ by the structure and shape of particles to which they apply.

In the standard MOCAGE configuration, the Mie code scheme that is used is the scheme for layered spherical particles developed by N. Voshchinnikov (www.astro.spbu.ru/staff/ilin2/SOFTWARE/nmie0.html). Since MOCAGE works with the homogeneous aerosol particles, the scheme had been adjusted for such particles. But, the comparison with observations indicated that this could have altered the proper functionality of the scheme.

For these reasons, in the updated MOCAGE configuration, another scheme was implemented. The scheme that was chosen and implemented in MOCAGE is Wiscombe's Mie code scheme (Wiscombe, 1980, 1979, revised 1996) which is developed for spherical homogeneous particles. To calculate the extinction efficiency in this scheme, it is necessary to define refractive indices of all aerosol types in the model. Their values are taken from the Global Aerosol Data Set (GADS, Köpke et al., 1997) and Kirchstetter et al. (2004).

Figure 4.1 highlights the impact due to the change of the Mie code. Comparison between the output of the Mie code of the standard configuration and the updated configuration is illustrated for a one day AOD field over Europe. The two schemes give different results, in particular for sea salt aerosols (with using the identical refractive indices). MODIS Terra observations showed two primary aerosol features over the field: sea salt aerosols on the east of Great Britain and the desert dust outbreak in the Middle East. The updated configuration has a better agreement on the spatial extent and the optical depth compared with observations, and has a better consistency between modelled concentrations and total columns. Desert dust AOD values were not as strongly affected as the AOD values of sea salt aerosols.

The model results with the Wiscombe Mie code will be extensively compared with various AOD observations afterwards, in Section 4.5.

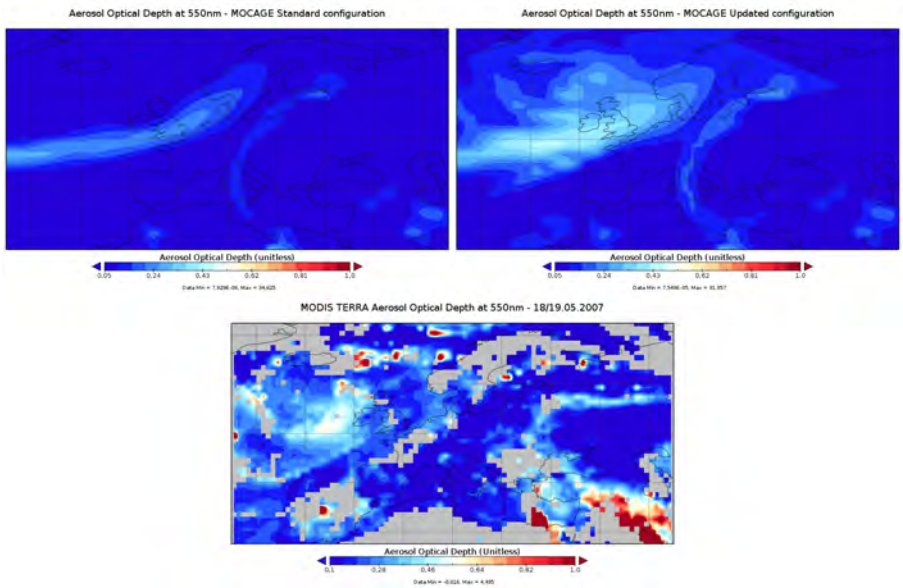

Figure 4.1: The aerosol optical depth over Europe as simulated in MOCAGE for 19.05.2007 at 0h with: **(left)** the standard configuration and **(right)** the updated configuration. For the same field, AOD from MODIS Terra is shown **(down)**.

4.3.1 AOD sensitivity to the sectional representation

The impact of the different number of bins on the aerosol mass concentration in MOCAGE was already explored in the case of desert dust aerosols (Martet, 2008). Aerosol optical depth does not linearly depend on the mass concentration. We want to estimate what uncertainty we introduce in results when using AOD with the bin repartition of MOCAGE.

With a simple setup, we calculate the aerosol optical depth at 550 nm of the light passing through the box of 10 m^3 containing 0.1 kg of aerosol particles. The aerosols are partitioned in the different ways:

- in total 6 bins distributed according to Table 4.4 (column 6*bin* in Table 4.1).
- in total 6 bins regularly partitioned in the logarithmic scale (column 6*bin_reg*).
- in total 10 bins regularly partitioned in the logarithmic scale (column 10*bin_reg*).

Table 4.1: Aerosol optical depth at 550 nm of 0.1 kg of aerosol of different types situated in the cube of 10 m^3 and partitioned in bins differently.

	6bin	6bin_reg	10bin_reg	20bin_reg	100bin_reg
desert dust	0.162	0.154	0.163	0.189	0.190
sea salt	0.766	0.605	0.706	0.760	0.763
black carbon	9.271	10.13	10.17	10.40	10.48
organic carbon	3.237	3.305	2.891	2.618	2.610

- in total 20 bins regularly partitioned in the logarithmic scale (column $20bin\_reg$).

- in total 100 bins regularly partitioned in the logarithmic scale (column $100bin\_reg$).

The size distribution is considered as log-normal, with initial size distribution properties for desert dust, black carbon and organic carbon described in Section 4.5.3.4. For the size distribution of the sea salt aerosols, we used values observed by O'Dowd et al. (1997) which were the basis for the parameterized emission scheme used in MOCAGE (Gong, 2003, further described in Section 4.5.3.4). The distribution contains two modes with number diameters $r_1 = 0.2\mu$m and $r_2 = 2\mu$m and geometric deviations $\sigma_1 = 1.9$ and $\sigma_1 = 2$.

As in the model, for each bin, the truncated number size distribution of the bin is considered to be uniform in log-scale. In this case the mean diameter of each bin is the arithmetic mean:

$$D_{bin_i} = \exp\left(\frac{\ln D_{bin_{i-1,i}} + \ln D_{bin_{i,i+1}}}{2}\right), \tag{4.2}$$

where $D_{bin_{i-1,i}}$ and $D_{bin_{i,i+1}}$ are the limits of the bin i $[m]$. This approximation introduces an error that decreases as the number of bins increases and the distribution gets closer to the continuous form. Also, the choice of bin ranges has an effect on AOD uncertainties. This can be examined with the data in Table 4.1 and Table 4.2.

The AOD from aerosols partitioned into 100 bins are considered to be sufficiently close to the continuous form of the size distribution and are considered as the reference values. Partitioning into 20 bins gives AOD very close to the 100 bins data. Further decreasing of the number of bins, by using 10 and 6 bins, makes the uncertainty grow. The relative error of using 6 bins is shown in the second column of Table 4.2 for different aerosol types.

Table 4.2: Relative differences in AOD at 550 nm produced by of 0.1 kg of aerosol of different types situated in the cube of 10 m^3 with bins partitioned differently. Absolute AOD values, from which the relative differences are calculated, are presented in Table 4.1.

	Relative difference		
	6bin and 6bin_reg	6bin_reg and 100bin_reg	6bin and 100bin_reg
desert dust	4.2%	19%	15%
sea salt	21%	21%	0.01%
black carbon	8.2%	2.2%	11%
organic carbon	2.6%	26%	24%

The choice of bin ranges has also an important effect on uncertainties (the first column in Table 4.2). It can compensate (the case of sea salt) or increase (the case of black carbon for example) the error discussed above. We can estimate that these two factors, the number of bins and the choice of bin ranges, have a similar impact on the AOD uncertainty.

In the test, the AOD uncertainties vary among aerosol types. This could be explained by their different size distributions. The aerosol size distributions evolve during the time, which makes that these uncertainties also change during the time. It means that the values from Table 4.2 could not be used for any type of correction of AOD values in the model. We can note that the sectional approach and the choice of bin ranges in MOCAGE, in average, produce the AOD uncertainty of $10 - 15\%$.

Since the mass of aerosols is constant for all species in the test, Table 4.1 highlights the differences of aerosol extinction efficiency between species. The reasons are different size ranges of aerosol types and their different optical properties. There are also other uncertainties influencing AOD (Tsigaridis et al., 2008) that are not discussed here.

4.3.2 AOD sensitivity to hygroscopic growth

Hydrophilic aerosols are assumed to be in an equilibrium with the ambient relative humidity. The water uptake increases the size of aerosols and this is

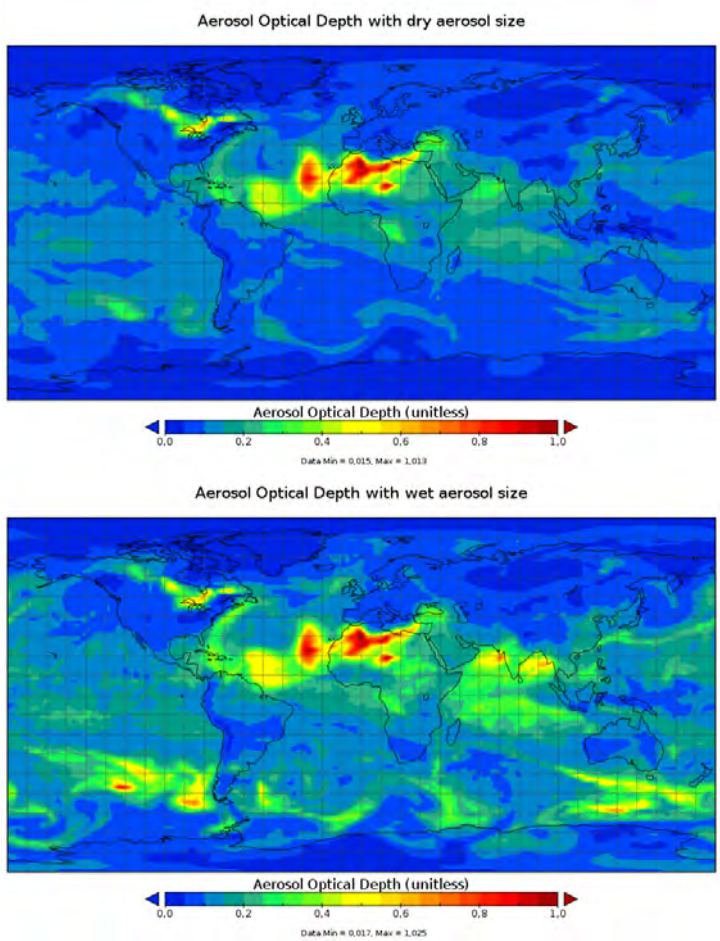

Figure 4.2: Global aerosol optical depth as simulated by MOCAGE for 30.06.2013. with: **(top)** dry sea salt aerosol size and **(bottom)** wet sea salt aerosol size by taking into account the water uptake.

Table 4.3: The change of diameters of sea salt aerosols for different relative humidities. Aerosol optical depth at 550 nm of 0.1 kg of dry mass of sea salt particles situated in the cube of 10 m^3 and calculated for different relative humidities.

	Relative humidity [%]						
	0	30	50	70	80	90	95
diameter	D_p	$\approx 1.14D_p$	$\approx 1.23D_p$	$\approx 1.37D_p$	$\approx 1.66D_p$	$\approx 2.04D_p$	$\approx 2.53D_p$
AOD	0.77	0.80	0.87	1.09	1.38	2.15	3.28

taken into account for sea salt aerosols in the AOD calculations. To calculate a wet particle size we use the hygroscopic growth formula from Gerber (1985):

$$r = \left(\frac{C_1 r_{\mathrm{d}}^{C_2}}{C_3 r_{\mathrm{d}}^{C_4} - \log \mathrm{RH}} + r_{\mathrm{d}}^3 \right)^{1/3}, \tag{4.3}$$

where $r_{\mathrm{d}}$is the dry particle radius [cm], RH is the relative humidity in percentage, r is the particle size at the RH relative humidity, and $C_1 = 0.7664$, $C_2 = 3.079$, $C_3 = 2.573 \times 10^{-11}$, $C_4 = -1.424$ are constants valid for sea salt particles. At very high values of relative humidity, this relation is not accurate enough (Fan and Toon, 2011). Thus, we limit relative humidity to 95 % to avoid unrealistic optical depths.

Besides the change in aerosol size, the water uptake changes particle refractive indices. The indices from GADS database were interpolated to the corresponding relative humidity. As the particle grows, the size parameter increases, and the refractive indices show weakening of scattering and absorbing capabilities of material (the real part of the indices increases, the complex part decreases). Table 4.3 shows the total effect of the hygroscopic growth on AOD at 550 nm for the same setup used for Tables 4.1 and 4.2. The growth of the sea salt aerosols increases the light extinction at 550 nm. The change is not linear, and becomes larger for high values of the relative humidity. Figure 4.2 illustrates this effect on the real modelled field. The AOD values over ocean are significantly larger due to the particle hygroscopic growth.

4.4 Development and evaluation of aerosol physical parameterizations

The improvements and further developments of aerosol physical parameterizations in the CTM MOCAGE done during this thesis, brought forth the publication submitted and published in the journal "Geoscientific Model Development":

B. Sič, L. El Amraoui, V. Marécal, B. Josse, J. Arteta, J. Guth, M. Joly, and P. Hamer. Modelling of primary aerosols in the chemical transport model mocage: development and evaluation of aerosol physical parameterizations. *Geoscientific Model Development*, 8(2): 381–408, 2015. doi: 10.5194/gmd-8-381-2015.

The article follows in its integrity, except for the bibliography which is common for the whole manuscript.

4.5 Article – Modelling of primary aerosols in the chemical transport model MOCAGE: development and evaluation of aerosol physical parameterizations

Authors:
B. Sič, L. El Amraoui, V. Marécal, B. Josse, J. Arteta, J. Guth, M. Joly, and P. Hamer, all from CNRM-GAME, Toulouse, France

Abstract

This paper deals with recent improvements to the global chemical transport model of Météo-France MOCAGE (Modèle de Chimie Atmosphérique à Grande Echelle) that consists of updates to different aerosol parameterizations. MOCAGE only contains primary aerosol species: desert dust, sea salt, black carbon, organic carbon, and also volcanic ash in the case of large volcanic eruptions. We introduced important changes to the aerosol parameterization concerning emissions, wet deposition and sedimentation. For the emissions, size distribution and wind calculations are modified for desert dust aerosols, and a surface sea temperature dependant source function is introduced for sea salt aerosols. Wet deposition is modified toward a more physically realistic representation by introducing re-evaporation of falling rain and snowfall scavenging

and by changing the in-cloud scavenging scheme along with calculations of precipitation cloud cover and rain properties. The sedimentation scheme update includes changes regarding the stability and viscosity calculations. Independent data from satellites (MODIS, SEVIRI), the ground (AERONET, EMEP), and a model inter-comparison project (AeroCom) are compared with MOCAGE simulations and show that the introduced changes brought a significant improvement on aerosol representation, properties and global distribution. Emitted quantities of desert dust and sea salt, as well their lifetimes, moved closer towards values of AeroCom estimates and the multi-model average. When comparing the model simulations with MODIS aerosol optical depth (AOD) observations over the oceans, the updated model configuration shows a decrease in the modified normalized mean bias (MNMB; from 0.42 to 0.10) and a better correlation (from 0.06 to 0.32) in terms of the geographical distribution and the temporal variability. The updates corrected a strong positive MNMB in the sea salt representation at high latitudes (from 0.65 to 0.16), and a negative MNMB in the desert dust representation in the African dust outflow region (from -1.01 *to* -0.22*). The updates in sedimentation produced a modest difference; the MNMB with MODIS data from 0.10 in the updated configuration went to 0.11 in the updated configuration only without the sedimentation updates. Yet, the updates in the emissions and the wet deposition made a stronger impact on the results; the MNMB was 0.27 and 0.21 in updated configurations only without emission, and only without wet deposition updates, respectively. Also, the lifetime, the extent, and the strength of the episodic aerosol events are better reproduced in the updated configuration. The wet deposition processes and the differences between the various configurations that were tested greatly influence the representation of the episodic events. However, wet deposition is not a continuous process; it has a local and episodic signature and its representation depends strongly on the precipitation regime in the model.*

4.5.1 Introduction

Atmospheric aerosols play a major role in a number of atmospheric processes and have an important global climate impact (IPCC, 2007). Increased effort has been made in the domain of aerosol modelling as knowledge of their importance has increased (Textor et al., 2006). The goal of the modelling has been to qualitatively and quantitatively represent aerosols in the correct way in order to better understand how aerosols affect atmospheric chemistry, air quality, climate, aviation, visibility, radiative budget and clouds. For this task, it is necessary to develop reliable parameterizations that describe how aerosols are emitted, transported and transformed, and, in the end, removed

from the atmosphere. Owing to this drive to improve model representation of aerosols, and due to the complexity of aerosol processes, a large diversity of parameterizations now exists. This variety produces a wide range of model results (Mahowald et al., 2003; Tegen, 2003; Textor et al., 2006). Therefore, the choice, development and validation of used parameterizations are crucial for the performance of the models (Lee et al., 2011).

Sources of aerosols are more difficult to define than those of gases (IPCC, 2007). In models, aerosol sources are characterized either by interactive emission parameterizations that depend on soil properties and/or wind intensity – which are, in the case of primary aerosols, generally used for desert dust and sea salt particles – or by existing emission inventories, mainly used for other primary aerosol types. Secondary aerosols are not directly emitted and they originate from gas-phase precursors or from reactions between dissolved or adsorbed gases and primary aerosols. The AEROCOM model inter-comparison run with and without harmonized emissions (Textor et al., 2007) showed that, although the uncertainties in emissions can be large, after the emission harmonization the inter-model diversity decreased slightly but remained large. The standard deviation of the total aerosol burden decreased from 18 Tg, for non-harmonized emissions, to 16 Tg, for harmonized emissions. Therefore, in addition to emissions, other model components, like the parameterizations of physical processes, contribute significantly to the model uncertainties.

Removal processes balance against the emission and production processes, and determine the lifetime of aerosols in the atmosphere. They are especially important for species that do not interact chemically (i.e. primary aerosols) because they represent their only available sinks. Mechanisms which remove aerosols are divided in two groups: "wet" deposition (scavenging) processes which take place in the interaction of aerosols with precipitation, and "dry" processes which include gravitational sedimentation (or gravitational settling) and dry deposition by interaction with the surface. The comparison of the models and their performance compared to dust measurements after long-range transport by Prospero et al. (2010) showed that the ratios of different deposition mechanisms varied greatly among the models and against the observed ratios. For example, the ratios of wet deposition to dry deposition ranged from about 1 : 1 to 30 : 1 in the models, in contrast from about 3 : 1 to 4 : 1 at the measurement stations. This and findings from the other studies demonstrate that aerosol deposition is complex and challenging to implement in an accurate way (Rasch et al., 2000; Sportisse, 2007; Prospero et al., 2010).

Wet deposition is the most efficient aerosol sink (Pruppacher et al., 1997), but it is regionally limited. Its uncertainty is augmented by the uncertainties in precipitation and aerosol properties, and wet deposition is identified as a key source of uncertainty in aerosol models (Vignati et al., 2010; IPCC, 2013).

Rasch et al. (2000) showed in an inter-comparison that model simulations differ most strongly in the upper troposphere for species undergoing wet scavenging processes. In all wet deposition processes, particles are indirectly transferred to the surface with the aid of precipitation. Inside clouds, in-cloud scavenging (rain-out) occurs when precipitation forms. Aerosols can act as condensation nuclei for the formation of water droplets and small cloud particles. When water vapour interacts with their surface, it can start to condense and allow the cloud droplets to grow. Additional aerosol particles can then be attracted and absorbed into them. When a droplet starts to precipitate, below-cloud scavenging (wash-out) takes place. While falling, a droplet can collide with aerosol particles and collect them from the air. Although less efficient than in-cloud scavenging, below-cloud scavenging is particularly important for coarse and very small particles (Andronache, 2003). Wet deposition is commonly parameterized by the scavenging coefficient Λ (s^{-1}) where $\frac{dc}{dt} = -\Lambda c$, c is the aerosol concentration. Many methods have been proposed in the literature to estimate the scavenging coefficient (e.g. Sportisse, 2007): more theoretical approaches, semi-empirical parameterizations with detailed modelling of various component processes that are responsible for aerosol deposition, or fully empirical approaches with a large number of different proposed formulations.

Aerosols undergo the influence of gravitational forces and tend to fall because their mass is not negligible. Near the surface, the dry deposition process acts together with gravitational sedimentation and it is especially efficient for coarse and very fine particles (Seinfeld and Pandis, 1998). Particles interact with the surface and objects in a thin layer of air next to the surface: they experience drag, change velocities and fall down. The velocity of dry deposition depends on properties of the surface, aerosols particles, and meteorological parameters (Seinfeld and Pandis, 1998).

Uncertainties in the models do not only come from the different formulations of deposition parameterization. Uncertainties in meteorological fields can also have a significant effect on model performance. Winds control the transport of species and can influence the interactive emission parameterizations. The humidity determines cloud coverage, rain localization and intensity – which are crucial for wet deposition processes – and hygroscopic particle growth, which is important for the particle settling and visibility.

In the present study we examine all of the previously mentioned processes in the chemical transport model (CTM) MOCAGE (Modèle de Chimie Atmosphérique à Grande Echelle). The CTM MOCAGE was developed at Météo-France and contributes to a wide range of scientific studies. Its applications cover both regional and global scales and extend to air-quality forecasts, climate–chemistry interactions (Teyssèdre et al., 2007; Lamarque

et al., 2013), desert aerosol studies (Martet et al., 2009), long-range transport pollution studies (Bousserez et al., 2007), "chemical weather" (Dufour et al., 2005), data assimilation of chemical species (e.g. El Amraoui et al., 2010), troposphere–stratosphere transport (Ricaud et al., 2009; Barré et al., 2012), etc. For its applications relating to aerosols, the CTM MOCAGE is implicated in a number of projects: MACC (`www.gmes-atmosphere.eu`), PREV'AIR (`www.prevair.org`), IMPACT2C (`www.hzg.de/mw/impact2c/`), and VAAC (Volcanic Ash Advisory Centre) predictions. The model outputs that are used in these projects are aerosol optical depth (AOD) and particulate matter concentrations ($PM_{2.5}$ and PM_{10} – particulate matter up to 2.5/10 µm in size).

Many aerosol processes are highly inter-connected; uncertainties and different formulations of processes lead to a large dispersion of model results as shown in comparative studies (Rasch et al., 2000; Textor et al., 2007; Prospero et al., 2010). This reveals the importance and complexity of aerosol physical parameterizations. In this paper, we present the recent developments on primary aerosol emissions and physical parameterizations in the CTM MOCAGE. Our main objective is to improve the aerosol representation in the model. To achieve this objective, we will, firstly, reexamine and modify primary aerosol emissions and parameterizations (wet scavenging and sedimentation) in MOCAGE; secondly, study sensitivities to different formulations of the mentioned processes in order to show how different treatments influence the aerosol representation in the model and to which extent their uncertainties affect the model performance; and thirdly, evaluate the new parameterizations for emissions, wet deposition, and sedimentation in MOCAGE by comparing the model outputs with different satellite and ground observations. We perform this evaluation for two physical quantities important for model applications: AOD and PM concentrations. The analysis and evaluation are based on the model output at the global scale for the year 2007.

The article is organized as follows. In Sect. 2 we present the general description of the model MOCAGE. The aerosol parameterizations in the model and their improvements are presented in detail in Sect. 3. Section 4 describes all observational data sets used for comparison with the model. In Sect. 5 we define the model experiments and explain the method used to assess model performance. Results and discussions are presented in Sects. 6 and 7 where we compare MOCAGE results with different independent observations and evaluate a new set of parameterizations in MOCAGE to estimate their impact on aerosol burden, lifetime, concentration, deposition and optical depth. Section 8 concludes this study.

4.5.2 General description of the model

MOCAGE is a global chemistry and transport model (CTM) developed at Météo-France. It is used as an operational air quality model simulating gases (Josse et al., 2004; Dufour et al., 2005) and primary aerosols (Martet et al., 2009). It transports atmospheric species by a semi-lagrangian advection scheme (Williamson and Rasch, 1989). Turbulent diffusion is implemented following Louis (1979), and convection following Bechtold et al. (2001). The dynamics within the CTM are forced by ARPEGE meteorological analysis fields (pressure, winds, temperature, specific humidity). ARPEGE is the operational global numerical weather prediction model of Météo-France. The precipitation field and liquid water content are calculated in MOCAGE in the same way as in ARPEGE. MOCAGE has 47 vertical hybrid sigma-pressure levels from the surface up to about 5 hPa. The vertical resolution is not uniform; levels are packed more densely near the surface, with a resolution of 40 m in the planetary boundary layer, about 400 m in the free troposphere and about 700–800 m in the upper troposphere and lower stratosphere. In the global configuration, simulations have a horizontal resolution of 2° latitude ×2° longitude.

Aerosols in MOCAGE are considered as an external mix of four primary aerosol species: desert dust, sea salt, black carbon (BC), organic carbon (OC) and volcanic ash. Volcanic ash aerosols are included only in the case of large volcanic eruptions and they are not considered in this study. The particle size distribution is divided across size bins, which are treated as passive tracers: aerosols are emitted, transported and removed from the atmosphere, and no transformations or chemical reactions between the different aerosol species or with gases are allowed. Each of the species has six size bins where we consider only the averaged mass and diameter of particles. The size ranges of bins for all considered aerosol species are shown in Table 4.4. The number of bins per species is limited to six in order to balance the operational cost and effectiveness. Two of the bins have their limits at 2.5 and 10 µm for practical air quality purposes in order to easily integrate the sum of $PM_{2.5}$ and PM_{10} particles. The other bin size ranges are distributed in a such manner as to have

Table 4.4: Bin ranges of individual primary aerosol species present in MOCAGE.

	bin1	bin2	bin3	bin4	bin5	bin6
desert dust (µm)	0.1–1	1–2.5	2.5–5	5–10	10–30	30–100
sea salt (µm)	0.003–0.13	0.13–0.3	0.3–1	1–2.5	2.5–10	10–20
black carbon (µm)	0.0001–0.001	0.001–0.003	0.003–0.2	0.2–1	1–2.5	2.5–10
organic carbon (µm)	0.0005–0.003	0.003–0.1	0.1–0.3	0.3–1	1–2.5	2.5–10

an optimal aerosol representation considering the initial size distribution and evolution of each aerosol species in the model.

Aerosol optical depth (AOD) in the model is calculated at 550 nm using Mie theory with refractive indices taken from the Global Aerosol Data Set (GADS, Köpke et al., 1997) and extinction efficiencies derived with Wiscombe's Mie scattering code for homogeneous spherical particles (Wiscombe, 1980).

4.5.3 Aerosol parameterizations in the model

In this section we describe the aerosol parameterizations in MOCAGE, as well as developments and updates that we have made to the parameterizations as part of this study. From now on, the present MOCAGE configuration will be referred to as SIM1, and the configuration with updated parameterizations as SIM2. For the complete description of the SIM1 and SIM2 configurations, the reader is referred to Section 4.5.5.

4.5.3.1 Dry deposition

Dry deposition of aerosol particles in the model is based on the Slinn and Slinn (1980) and Slinn (1982b) studies that describe the deposition process as a transport to the surface in terms of resistances in series aided by particle sedimentation. The complete scheme is described in detail in Nho-Kim et al. (2004). Briefly, the process of particulate dry deposition is composed of transport through the atmospheric surface layer governed by turbulent diffusion (aerodynamical resistance), the transport in the quasi-laminar layer influenced by diffusion, interception and impaction (quasi-laminar layer resistance), and adherence to the surface which is considered totally efficient. Each of these mechanisms contributes to the deposition velocity. The characteristics of the surface are defined as in the ARPEGE model which includes physical parameters of soils (roughness length, vegetation type) necessary for particle–surface interaction. The dry deposition velocity is defined as

$$V_{\mathrm{dd}} = \frac{1}{R_{\mathrm{a}} + R_{\mathrm{b}}} + V_{\mathrm{p}}, \tag{4.4}$$

where R_{a} is the aerodynamical resistance ($\mathrm{s\,m^{-1}}$), R_{b} is the quasi-laminar resistance ($\mathrm{s\,m^{-1}}$), and V_{p} is the settling velocity ($\mathrm{m\,s^{-1}}$). The aerosol dry deposition scheme is not a subject to the changes in this study.

4.5.3.2 Sedimentation

Gravitational settling of aerosol particles is implemented as described in Seinfeld and Pandis (1998). The settling velocity is based on Stokes law and is a function of particle diameter, particle density, and air viscosity:

$$V_{\mathrm{p}} = \frac{D_{\mathrm{p}}^2 \rho_{\mathrm{p}} g C_{\mathrm{c}}}{18 \mu_{\mathrm{a}}} \tag{4.5}$$

where D_{p} is the ambient aerosol diameter (m), ρ_{p} is the aerosol particle density ($\mathrm{kg\,m^{-3}}$), g is the gravitational constant ($\mathrm{m\,s^{-2}}$), μ_{a} is the dynamical viscosity of air (Pa s), and C_{c} is the slip correction factor which accounts for noncontinuum effects when the particle diameter and the air mean free path are of the same order of magnitude. C_{c} is defined as (Seinfeld and Pandis, 1998):

$$C_{\mathrm{c}} = 1 + \frac{2\lambda}{D_{\mathrm{p}}} \left[1.257 + 0.4 \exp\left(-\frac{1.1 D_{\mathrm{p}}}{2\lambda} \right) \right] \tag{4.6}$$

where λ is the mean free path of an air particle [m].

In the model configuration SIM1, we calculate the dynamical air viscosity using an assumed constant value of the kinematic viscosity. In the updated sedimentation calculations, in SIM2, we calculate it by Sutherland's law, an empirical relation connecting dynamical viscosity and temperature (White, 1991):

$$\mu_{\mathrm{a}} = \mu_0 \frac{T_0 + S}{T + S} \left(\frac{T}{T_0} \right)^{3/2} \tag{4.7}$$

where μ_0 is the reference dynamical viscosity of air at the reference temperature T_0 with values of $\mu_0 = 1.716 \times 10^{-5}$ Pa s and $T_0 = 273\,\mathrm{K}$, and $S = 111\,\mathrm{K}$ is the Sutherland's effective temperature (White, 1991).

Finally, in SIM2, to ensure the stability and the mass conservation of our explicit sedimentation scheme, sedimentation velocity is not allowed to exceed one gridbox height per model time step.

4.5.3.3 Wet deposition

The fraction of aerosols removed at each time step by interaction with precipitation (by both in-cloud and below-cloud scavenging) is calculated as

$$F = f_{\mathrm{prec}} (1 - e^{-\Lambda \Delta t}) \tag{4.8}$$

where F is the fraction of removed aerosols, f_{prec} is the fraction of precipitating cloud cover (the percentage of a cloud coverage in a gridbox where precipitation forms or falls); Λ is the scavenging coefficient (s^{-1}) which describes a rate of loss of particles due to scavenging; Δt is the model time step for scavenging (s). The scavenging coefficient, Λ, consists of the in-cloud scavenging coefficient, Λ_{ro}, and the below-cloud scavenging coefficient due to rainfall, Λ_{wo}, and due to snowfall, Λ_{so}. To calculate them, we use the respective in-cloud and below-cloud parameterized schemes described in the following.

Cloud cover of precipitation cloud cover In SIM1, we use a simple approach by considering that if precipitation forms in the gridbox it happens in all available cloud cover in the gridbox. To better represent the precipitating cloud cover in MOCAGE, we updated the model by adapting in SIM2 a scheme from Giorgi and Chameides (1986). To estimate the portion of the sky covered by precipitating clouds, this scheme considers typical conditions in stratiform and convective clouds during the formation of precipitation and compares them with the modelled gridbox mean precipitation formation rates. Precipitation formation rates are calculated by the diagnostic scheme that uses the cloudiness scheme from Xu and Randall (1996) and the precipitation scheme from Kessler (1969). For stratiform clouds, the fraction of precipitation forming clouds is (we also take all values of quoted parameters from Giorgi and Chameides (1986) if not stated differently)

$$f_{\mathrm{strat}} = \frac{Q}{(L_{\mathrm{st}} \cdot R_{\mathrm{st}} + Q)} \tag{4.9}$$

where Q is the gridbox mean rate of precipitation formation including both liquid and solid precipitation ($\mathrm{kg\,m^{-3}\,s^{-1}}$). L_{st} is the typical in-cloud liquid water content in precipitation forming stratiform clouds: $L_{\mathrm{st}} = 1.5 \times 10^{-3}\,\mathrm{kg\,m^{-3}}$, from Brost et al. (1991). It differs from the value originally proposed by Giorgi and Chameides (1986), $L_{\mathrm{st}} = 0.5 \times 10^{-3}\,\mathrm{kg\,m^{-3}}$, taken from Pruppacher and Klett (1978). The value from Giorgi and Chameides (1986) was corrected by Brost et al. (1991) and later adopted by Jacob et al. (2000) and Liu et al. (2001). R_{st} is the in-cloud rate constant of conversion of cloud water to precipitation for stratiform precipitation: $R_{\mathrm{st}} = 1 \times 10^{-4}\,\mathrm{s^{-1}}$.

For convective clouds, the fraction of precipitating cloud cover within a gridbox for any given time step is

$$f_{\mathrm{conv}} = \frac{F_0 Q \frac{\Delta t}{t_{\mathrm{c}}}}{Q \frac{\Delta t}{t_{\mathrm{c}}} + F_0 R_{\mathrm{cv}} L_{\mathrm{cv}}} \tag{4.10}$$

where F_0 is the maximum cumulus cloud cover assumed in the radiation calculations backed by observations, $F_0 = 0.3$; Δt is the model time step; t_{c} is

the typical duration of precipitation from a cumulonimbus cloud, $t_c = 30\,\text{min}$ (Liu et al., 2001); R_{cv} is the in-cloud rate constant of conversion of cloud water to precipitation in convective clouds, $R_{cv} = 1.5\times10^{-3}\,\text{s}^{-1}$; and L_{cv} is the typical in-cloud liquid water content in cumulonimbus clouds, $L_{cv} = 2\times10^{-3}\,\text{kg}\,\text{m}^{-3}$.

To estimate the scavenging coefficient, many parameterizations have been developed and Sportisse (2007) summarizes them adequately. In our model, the current parameterization for in-cloud scavenging is the Langner and Rodhe (1991) scheme and in this study it will be evaluated against the Giorgi and Chameides (1986) scheme. Besides, in this study we modified and re-evaluated the model's current below-cloud scavenging scheme based on Slinn (1977).

Implemented schemes To estimate the scavenging coefficient Λ and its components, many parameterizations have been developed and Sportisse (2007) summarizes them adequately. In our model, the current parameterization for in-cloud scavenging, used in SIM1, is the Langner and Rodhe (1991) scheme and in this study it will be evaluated against the Giorgi and Chameides (1986) scheme, which is implemented in the SIM2 configuration. Additionally, in this study we modified and re-evaluated the model's current below-cloud scavenging scheme based on Slinn (1977).

Rain below-cloud scavenging Below-cloud scavenging in the model acts in all gridboxes, and gridbox fractions, where precipitation falls. However, below-cloud scavenging cannot occur in the same gridboxes, or gridbox fractions, where precipitation forms. In order to calculate the fraction of a particular gridbox where below-scavenging acts we examine the overlying layers above that gridbox and find the layer with the maximum precipitation fraction. We then subtract from this maximum fraction, the fraction where in-cloud scavenging acts in the gridbox we are examining. The rain below-cloud scavenging coefficient is defined as in Seinfeld and Pandis (1998): (Seinfeld and Pandis, 1998):

$$\Lambda = \frac{3}{2}\frac{EP}{D_d} \tag{4.11}$$

where E_r is the collection efficiency of a raindrop to collect a particle during its fall, P is the precipitation rate in precipitating area ($\text{kg}\,\text{m}^{-2}\,\text{s}^{-1}$), and D_d is the raindrop diameter (m). To permit both, rain-out and wash-out, to take place in the same gridbox at the same time, we revised the condition for when and where wash-out occurs, and we now assume that it happens in all regions exactly below the rain-out area.

We calculate the collection efficiency using Slinn's below-cloud scavenging scheme (Slinn, 1977), described also in Seinfeld and Pandis (1998) and widely used in models (Wang et al., 2010). Slinn's scheme considers collisions between a falling raindrop and an aerosol particle, and accounts for Brownian diffusion, interception and impaction. The collision efficiency is a function of the sizes of raindrops and aerosols, and is expressed as (Slinn, 1977)

$$E = \frac{4}{Re\,Sc}(1 + 0.4Re^{1/2}Sc^{1/3} + 0.16Re^{1/2}Sc^{1/2}) + 4\phi[\omega^{-1} + (1 + 2Re^{1/2})\phi] + \left(\frac{Stk - Stk^*}{Stk - Stk^* + \frac{2}{3}}\right)^{3/2} \cdot \left(\frac{\rho_\mathrm{d}}{\rho_\mathrm{p}}\right)^{1/2} \tag{4.12}$$

where $Re = \frac{D_\mathrm{d} V_\mathrm{d} \rho_\mathrm{a}}{2\mu_\mathrm{a}}$ is the Reynolds number of the raindrops based on their radius, $V_\mathrm{d} = \frac{D_\mathrm{d}^2 \rho_\mathrm{d} g C_\mathrm{c}}{18\mu_\mathrm{a}}$ is the terminal raindrop velocity as used in SIM1 (expression based on Stokes law) ($\mathrm{m\,s^{-1}}$), ρ_a and ρ_d are the density of air and water ($\mathrm{kg\,m^{-3}}$), $Sc = \frac{\mu_\mathrm{a}}{\rho_\mathrm{a} D}$ is the Schmidt number of the collected aerosol particles, $D = \frac{k T_\mathrm{a} C_\mathrm{c}}{3\pi\mu_\mathrm{a} D_\mathrm{p}}$ is the aerosol diffusivity ($\mathrm{m^2\,s}$), k is the Boltzmann constant ($\mathrm{J\,K^{-1}}$), T_a is the air temperature (K), $Stk = \frac{2\tau(V_\mathrm{d} - V_\mathrm{p})}{D_\mathrm{d}}$ is the Stokes number of the collected particles, $\tau = V_\mathrm{p}/g$ is the characteristic relaxation time (s), $Stk^* = \frac{1.2 + \frac{1}{12}\ln(1+Re)}{1+\ln(1+Re)}$ is the critical Stokes number, $\phi = D_\mathrm{p}/D_\mathrm{d}$ is the ratio of diameters of the aerosol particle and the rain droplet, and ω is the viscosity ratio of air and water. Considering terminal raindrop velocity, the expression defined above, used in SIM1, covers only the Stokes flow regime. But, the majority of raindrops falls with velocities out of the Stokes flow regime where inertial forces must be regarded, that is true for $D_\mathrm{d} > 2 \times 10^{-5}\,\mathrm{m}$ (Seinfeld and Pandis, 1998). The expressions of the raindrop terminal velocity which cover the whole raindrop size range are based on experimental data. From Brown and Lawler (2003), in SIM2 we use

$$V_\mathrm{t} = \frac{V_\mathrm{d}}{1 + 0.17\sqrt{Re}} \tag{4.13}$$

where V_d is the Stokes flow velocity defined earlier, and Re is the corresponding Reynolds number at the Stokes velocity.

In SIM1, the raindrop diameter is presumed to be fixed with the value of 1 mm. To examine effects of this assumption we consider raindrops to be also distributed in size. In SIM2, we use the exponential raindrop distribution from Marshall and Palmer (1948).

The first term in the collision efficiency equation (Eq. (4.12)) describes Brownian diffusion and is the most important for the smallest particles ($D_\mathrm{p} <$

0.2 µm), while the second and the third terms describe interception and inertial impaction which dominate for bigger particles ($D_\mathrm{p} > 1\,\mu\mathrm{m}$) (Seinfeld and Pandis, 1998).

Phoretic and electric effects The scavenging calculated due to diffusion, interception and impaction showed possible underestimation of scavenged quantities when compared with field measurements (Davenport and Peters, 1978; Laakso et al., 2003). Some authors broaden scavenging by including more mechanisms – thermophoresis, diffusiophoresis, and electric effects (Davenport and Peters, 1978; Chate, 2005; Andronache et al., 2006). Thermophoresis makes particles move along a temperature gradient; diffusiophoresis makes particles move due to gas concentration gradients (e.g. motion toward the raindrop during condensation); and electric forces make charged particles interact with each other. We included these effects to (Eq. (4.12)) in the SIM2_BCPLUS configuration (Table 4.5) as (Davenport and Peters, 1978)

$$\text{Thermophoresis } E_\mathrm{th} = \frac{4\alpha\left(2 + 0.6Re^{\frac{1}{2}}Pr^{\frac{1}{3}}\right)(T_\mathrm{a} - T_\mathrm{s})}{V_\mathrm{t} D_\mathrm{d}} \tag{4.14}$$

$$\text{Diffusiophoresis } E_\mathrm{df} = \frac{4\beta\left(2 + 0.6Re^{\frac{1}{2}}Sc_\mathrm{w}^{\frac{1}{3}}\right)\left(\frac{P_\mathrm{s}^0}{T_\mathrm{s}} - \frac{P^0\mathrm{RH}}{T_\mathrm{a}}\right)}{V_\mathrm{t} D_\mathrm{d}} \tag{4.15}$$

$$\text{Electrostatic charge } E_\mathrm{ec} = \frac{16KC_\mathrm{c}a^2\gamma^2 D_\mathrm{p}}{3\pi\mu_\mathrm{a} V_\mathrm{t}} \tag{4.16}$$

where $\alpha = \frac{2C_\mathrm{c}\left(k_\mathrm{a} + \frac{5\lambda}{D_\mathrm{p}k_\mathrm{p}}\right)k_\mathrm{a}}{5P\left(1+\frac{6\lambda}{D_\mathrm{p}}\right)\left(2k_\mathrm{a}+k_\mathrm{p}+\frac{10\lambda}{D_\mathrm{p}k_\mathrm{p}}\right)}$, k_a and k_p are the thermal conductivity of air and aerosol particle ($\mathrm{J\,m^{-1}\,s^{-1}\,K^{-1}}$), P is the atmospheric pressure (Pa), $Pr = \frac{c_p\mu_\mathrm{a}}{k_\mathrm{a}}$ is the Prandtl number for air, c_p is the specific heat capacity of air ($\mathrm{m^2\,s^{-2}\,K^{-1}}$), T_s is the temperature at the surface of the raindrop and it is taken to be 1 K less that the air temperature (Slinn and Hales, 1971), $\beta = \frac{T_\mathrm{a}D_\mathrm{w}}{P}\left(\frac{M_\mathrm{w}}{M_\mathrm{a}}\right)$, $D_\mathrm{w} = 2.1\times10^{-5}\left(\frac{T_\mathrm{a}}{T_0}\right)^{1.94}\left(\frac{P}{P_0}\right)$ is the water vapour diffusivity (Pruppacher et al., 1997), M_w and M_a are the molecular weights of water and air, respectively, $Sc_\mathrm{w} = \frac{\mu_\mathrm{a}}{\rho_\mathrm{a}D_\mathrm{w}}$ is the Schmidt number for water vapour in air, P_s^0 and P^0 are the water vapour partial pressures (in Pa) at temperatures T_s and T_a, respectively, RH is the relative humidity, K is the Coulomb constant, a is a constant $a = 0.83\times10^{-6}$, and γ is the parameter of cloud electricity and it is taken as an averaged value $\gamma = 2$ (Pruppacher et al., 1997; Andronache, 2004).

Below-cloud scavenging due to snowfall We extended the scavenging module in SIM2 by adding snowfall scavenging. Often, precipitation in liquid state at

the surface originates from solid state precipitation at higher altitudes. Tests in MOCAGE show that snowfall wash-out occurs in a larger number of gridboxes than rainfall wash-out. Compared to rainfall scavenging, there are fewer studies of the scavenging due to snowfall and there is a wider set of necessary snowfall parameters (due to different types and shapes of snow particles), which lead to larger uncertainties in the aerosol scavenging due to snowfall in the models. Also, snow scavenging efficiencies measured by different authors have a wide range of values: some are similar to those of rainfall, but some are 1 order of magnitude larger or lower (Sportisse, 2007).

Within MOCAGE, we introduce the Slinn (1977, 1982a) snowfall scavenging formula, which is one of the most commonly used snowfall parameterizations (Gong et al., 1997; Croft et al., 2009; Zhang et al., 2013). All snow crystals in this study are assumed to be formed by riming. The snowfall below-cloud scavenging coefficient is given as (Slinn, 1982a):

$$\Lambda = \frac{\gamma E P}{D_{\mathrm{m}}} \tag{4.17}$$

where E_{s} is the collection efficiency of a snow crystal to collect a particle during its fall, γ is the dimensionless fractional constant (in our case 0.5), and D_{m} is the characteristic volume-to-area length scale (for the rimed crystals $D_{\mathrm{m}} = 2.7 \times 10^{-5}\,\mathrm{m}$, Slinn, 1982a).

The Slinn (1977, 1982a) formulation is aerosol size, aerosol-type and snow-crystal-type dependent. The collection efficiency of the snow crystals is

$$E = \left(\frac{1}{Sc}\right)^{\delta} + \left[1 - \exp\left[-\left(1 + \sqrt{Re_l}\right)\frac{D_{\mathrm{p}}^2}{l^2}\right]\right] + \left(\frac{Stk - Stk^*}{Stk - Stk^* + \frac{2}{3}}\right)^{3/2} \cdot \left(\frac{\rho_{\mathrm{s}}}{\rho_{\mathrm{p}}}\right)^{1/2} \tag{4.18}$$

where the exponent δ depends on the snow-crystal type, l is the characteristic length of collecting ice filaments, and Re_l is the corresponding Reynolds number; $\rho_{\mathrm{s}} = 100\,\mathrm{g\,m^{-3}}$ is the density of falling snow. For rimed snow crystals that we consider in the model: $l = 100\,\mu\mathrm{m}$, $Re_l = 10$ and $\delta = \frac{2}{3}$ (Slinn, 1977). Since we consider only rimed crystals of a fixed size, terminal settling velocity is considered constant: $V_{\mathrm{s}} = 0.9\,\mathrm{m\,s^{-1}}$ (Todd, 1964).

Re-evaporation We introduced precipitation re-evaporation in the below-scavenging module in SIM2. If the fraction f of precipitation evaporates at one level, then the corresponding $0.5f$ fraction of scavenged aerosols will be released back to the atmosphere. The factor of 0.5 (Liu et al., 2001) is due to the fact that water molecules are more efficiently released than aerosols. If precipitation evaporates completely, then all scavenged aerosols are released.

Sublimation of snowfall is not taken in account, and it is presumed that all solid precipitation would first melt and then evaporate.

4.5.3.4 Emissions

All considered species are emitted as particles, i.e. primary aerosols. For emissions of black carbon and organic carbon we use prepared emission inventories, while for desert dust and sea salt we use online parameterizations.

The anthropogenic carbonaceous aerosol emissions in the SIM1 configuration come from the monthly defined AeroCom emission inventory (Dentener et al., 2006). Dentener et al. (2006) is based on Bond et al. (2004), which used the reference year 1996. In the SIM2 configuration, the organic carbon and black carbon anthropogenic emissions come from the inventory of Lamarque et al. (2010). Lamarque et al. (2010) monthly defined emissions are based on Bond et al. (2007) and Junker and Liousse (2008), which are harmonized with the reference year 2000. Lamarque et al. (2010) updated these previous inventories using other studies regarding additional emission sources (coal burning, domestic biofuel, ship tracks). Biomass burning emissions for both organic carbon and black carbon come from the GFEDv3 project (van der Werf et al., 2010). In GFEDv3, the data from biogeochemical modelling and active fire satellite measurements (MODIS and GOES) are combined to a daily state-of-the-art biomass burning emission estimate (Mu et al., 2011). Biomass burning carbon emissions are injected more quickly to higher altitudes compared to other emissions, due to fire-induced convection. The maximal injection height depends on fire heat flux and environmental conditions, and varies significantly with latitude. In our model we have defined the maximal injection height in the tropical regions to be 1000 m, in midlatitudes 4000 m, and in the boreal regions 6000 m. Our choice is consistent with Williams et al. (2009).

The black carbon and organic carbon initial size-distribution is defined using a two-mode lognormal distribution with the number mode diameters of the two modes as $r_1 = 1.5 \times 10^{-8}$ m and $r_2 = 4 \times 10^{-8}$ m, the geometric standard deviation $\sigma_1 = \sigma_2 = 1.8$, and the mass distribution between modes $\mathrm{frac}_1 = 0.4$ and $\mathrm{frac}_2 = 0.6$ (Dentener et al., 2006).

Sea-salt source function Monahan and O'Muircheartaigh (1986) developed a formulation for the production of sea salt particles resulting from the bursting of wind-formed sea surface bubbles. Their semi-empirical formulation depends on the particle size and the intensity of surface winds. Gong (2003) addressed

the overestimation of small particles ($D < 0.2\,\mu m$) compared with observations and proposed an improved formulation. The rate of sea salt particle production (particles $m^{-2}\,s^{-1}\,\mu m^{-1}$) became (Gong, 2003)

$$\frac{dF}{dr} = 1.373 u_{10}^{3.41} r^{-A} (1 + 0.057 r^{3.45}) \cdot 10^{1.607 e^{-B^2}} \tag{4.19}$$

where r is the particle radius at relative humidity of 80 %, u_{10} is the wind speed at 10 m above the surface ($m\,s^{-1}$), and the parameters $A = 4.7(1 + 30r)^{-0.017 r^{-1.44}}$ and $B = (0.433 - \log r)/0.433$. Jaeglé et al. (2011) compared modelled data with AOD and sea salt measurements from coastal stations, satellites and ocean cruises, and found that the Gong (2003) function at high wind speeds ($> 6\,m\,s^{-1}$) overestimates sea salt concentrations over cold waters, and underestimates them over tropical waters. Their modified sea salt source function includes a sea surface water temperature dependence (Jaeglé et al., 2011):

$$\frac{dF}{dr} = (0.3 + 0.1T - 0.0076T^2 + 0.00021T^3) \cdot 1.373 u_{10}^{3.41} r^{-A} (1 + 0.057 r^{3.45}) \cdot 10^{1.607 e^{-B^2}} \tag{4.20}$$

where T is the sea surface temperature (SST; °C). The possible mechanisms of how sea surface temperature influences sea salt production are mentioned in Jaeglé et al. (2011): they are connected with kinetic viscosity of water and the gas exchange efficiency which leads to stronger whitecaps coverage in warmer waters (Lewis and Schwartz, 2004a; Anguelova and Webster, 2006). In MOCAGE, the sea salt source function proposed by Gong (2003) is used in SIM1, and the Jaeglé et al. (2011) modification is implemented in SIM2 and evaluated in this study. Both of these formulas use particle size at relative humidity of 80 %, and to calculate a dry particle sea salt source function we use the Gerber (1985) hygroscopic growth formula:

$$r = \left(\frac{C_1 r_d^{C_2}}{C_3 r_d^{C_4} - \log \mathrm{RH}} + r_d^3 \right)^{\frac{1}{3}} \tag{4.21}$$

where r_d is the dry particle radius (cm); RH is the relative humidity in percentage; r is the particle size at the RH relative humidity; and $C_1 = 0.7664$, $C_2 = 3.079$, $C_3 = 2.573 \times 10^{-11}$, and $C_4 = -1.424$ are constants valid for sea salt particles. The particle sizes are assumed to be in an equilibrium corresponding with the ambient relative humidity. The hygroscopic growth affects optical properties and deposition of sea salt aerosols, and Eq. (4.21)) is also used to calculate these effects. The Gerber (1985) relation is not accurate for high relative avoid unrealistic optical depths and deposition. In SIM2, the sea salt temperature used in Eq. (4.20)) is implemented from the Reynolds data set (Reynolds et al., 2002).

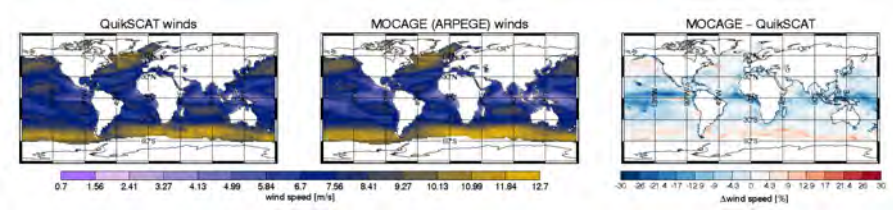

Figure 4.3: Mean annual surface winds for 2007: left – QuikSCAT measurements, middle – ARPEGE analysis, and right – their relative difference.

Due to the $u_{10}^{3.41}$ wind dependency (Eq. (4.20)), the sea salt source function is very sensitive to the quality of the wind field in the model. To assess winds used in the CTM MOCAGE we compared the surface wind speed of the ARPEGE analysis with satellite surface wind measurements from the SeaWinds scatterometer located on the QuikSCAT satellite. Spaceborne scatterometers are calibrated to measure the so-called equivalent neutral stability wind defined as the wind that would be observed under neutral stability conditions or atmospheric stratification. The equivalent neutral stability wind speed is very similar to actual wind speed, but they are not the same. The differences between the two can be as large as $0.5\,\mathrm{m\,s^{-1}}$ (Bourassa et al., 2003). We use the monthly level 3 (L3) QuikSCAT data set for 2007 with a resolution of $1° \times 1°$ (Bourassa et al., 2003), which is regridded to the MOCAGE $2° \times 2°$ resolution and averaged to get a mean annual wind field. The comparison of the mean 2007 wind fields from ARPEGE and QuikSCAT are presented in Fig. 4.3. The two fields have a very good agreement, with relative differences that are their strongest ($\sim 20\,\%$) in the regions dominated by low wind speeds. The differences are very similar to what Chelton and Freilich (2005) found by comparing ECMWF and QuikSCAT fields. A part of the disagreements can be explained by the differences between the equivalent neutral stability wind, which is observed by the scatterometer, and the actual wind, which is represented in the NWP (numerical weather prediction) analyses, and the fact that scatterometer retrievals typically overestimate buoy observations for relatively low wind speeds ($< 4\,\mathrm{m\,s^{-1}}$) (Bentamy et al., 1999; Chelton and Freilich, 2005). It should also be noted that Chelton (2005) remarked that NWP models do not represent well the influence of SST on low-speed winds over warm waters that could lead to a model underestimation in these regions.

Desert dust emission schemes The emission of mineral dust particles in arid zones depends on the surface characteristics and wind intensity. If the wind friction velocity is larger than the erosion threshold velocity for a given particle size and soil properties, particles can be emitted into the atmosphere (e.g. Zhao et al., 2006). A desert dust emission scheme takes into account all of the

main processes involved: achievement of the erosion threshold, saltation where particles start to move horizontally, and sandblasting where the fine particles are released from soil aggregates into the atmosphere due to impacts between the saltating particles and the surface.

In MOCAGE, two emission schemes have been implemented: the first one for African and Arabian deserts (Marticorena et al., 1997), and the second one for deserts in Asia (Laurent et al., 2006). The Marticorena et al. (1997) scheme covers Africa, Arabia and the Middle East (13–36° N, 17° W–77° E) with a resolution of $1° \times 1°$. The input soil properties and aerodynamical surface parameters are based on available pedological, topographical, geological and climatological data and analysis (Marticorena et al., 1997; Callot et al., 2000). The main sources were from the French National Geographic Institute (IGN) and Soviet topographic maps. Laurent et al. (2006) developed the emission scheme for north-eastern Asia that includes all arid areas in the region 35.5–47° N, 73–125° E. Typical soil characteristics are derived from soil samples (Mei et al., 2004) and statistically analysed and extrapolated to all known deserts in the domain. Aerodynamical surface parameters are determined from POLDER-1 surface bidirectional reflectance observations with a resolution of $0.25° \times 0.25°$.

Regarding the desert dust emission schemes in the different model configurations, in SIM2 compared to SIM1, we changed the wind fields interpolation method and the initial size distribution.

In SIM1, ARPEGE wind analysis is rebinned to the resolution of the emission schemes with the nearest-neighbour interpolation. In SIM2 we also take into account all adjacent gridboxes with the bilinear interpolation.

The initial emitted size-distribution is a three-mode lognormal distribution composed of fine, accumulation and coarse modes. The size distribution used in SIM1 has the number median diameters $r_1 = 1.7 \times 10^{-6}$ m, $r_2 = 6.7 \times 10^{-6}$ m, and $r_3 = 14.2 \times 10^{-6}$ m; geometric standard deviations $\sigma_1 = 1.7$, $\sigma_2 = 1.6$, and $\sigma_3 = 1.5$; and mass fractions $\mathrm{frac}_1 = 0.3$, $\mathrm{frac}_2 = 0.4$, and $\mathrm{frac}_3 = 0.3$. In this study we modified the size distribution following Alfaro et al. (1998) and Crumeyrolle et al. (2011), and in SIM2 our distribution is shifted towards smaller sizes with number median diameters $r_1 = 6.4 \times 10^{-7}$ m, $r_2 = 3.45 \times 10^{-6}$ m, and $r_3 = 8.67 \times 10^{-6}$ m; the standard deviations and the mass fractions are the same as above.

4.5.4 Observations

To evaluate the performance of the model we use large-scale satellite observations, ground-based photometer data and in-situ surface measurements. The MODIS (Moderate-resolution Imaging Spectroradiometer) instruments observe atmospheric aerosols aboard Terra (since 2000) and Aqua (since 2002) from complementary sun-synchronous orbits. We use MODIS aerosol optical depth Collection 5 retrievals at 550 nm from Terra and Aqua that have predicted uncertainties of $\Delta\tau = \pm(0.03 + 0.05\tau)$ over oceans and $\Delta\tau = \pm(0.05 + 0.15\tau)$ over land (Remer et al., 2005). We start with good-quality global level 3 (L3) daily MODIS data (QA-weighted products) and perform an additional quality control by rejecting all gridboxes with less than five level 2 (L2) observations per a L3 gridbox and more than a 50 % cloud fraction. To combine Terra and Aqua observations and to regrid from the original L3 $1^\circ \times 1^\circ$ grid to the MOCAGE $2^\circ \times 2^\circ$ grid we weight data by considering the number of L2 observations in each L3 gridbox. The data are processed in this manner to minimize the number of observations that are cloud contaminated and those with statistically low confidence, which often artificially increase AOD (Remer et al., 2008; Zhang et al., 2005; Koren et al., 2007).

AERONET (Aerosol Robotics Network) measures ground-based AOD from hundreds of automated stations with an accuracy of ±0.01 (Holben et al., 1998). We use L2 daily data from different stations and interpolate it in logarithmic space to 550 nm (to harmonize wavelengths between different stations and with the model) by using available neighbouring wavelengths: 440, 500, 675, and 870 nm.

Carrer et al. (2010) applied a multi-temporal approach to SEVIRI (Spinning Enhanced Visible and InfraRed Imager) geostationary observations to derive surface and aerosol properties simultaneously. They retrieved AOD over land using directional and temporal analysis of the signal, as opposed to spectral and spatial analysis done in MODIS retrievals (Ichoku et al., 2005). The data cover the SEVIRI field of view with a selected resolution of $1^\circ \times 1^\circ$, which is later regridded to the MOCAGE resolution. SEVIRI AOD observations are considered only if their relative uncertainty is estimated to be less than 75 %.

The EMEP (European Monitoring and Evaluation Programme – Cooperative programme on the long-range transmission of air pollutants in Europe) observation network consists of background stations and provides particulate matter measurements ($PM_{2.5}$ and PM_{10}) throughout Europe (Tørseth et al., 2012). We use measurements from the EMEP stations where primary

aerosols have a dominant effect. The considered stations have hourly or daily measurement frequencies.

4.5.5 Experiment design

We conduct our experiment to test the performance of the model in two main stages. First, we compare model outputs with observations. We define two main model configurations used as reference simulations and compare them with observations to evaluate the overall impact of the model updates. The reference simulations are called SIM1 and SIM2 and their configurations are presented in Table 4.5. SIM1 uses the configuration of MOCAGE with the current parameterizations while in SIM2 we use the updated parameterizations. Second, we evaluate the sensitivities of our results to the individual modules updates introduced in this study. To emphasize the separate effects of the parameterization updates, we have implemented different configurations based on the reference simulations. We separately analyse the impact of these updates on the emissions, sedimentation and wet scavenging (in simulations SIM2_EMI, SIM2_SED and SIM2_WDEP in Table 4.5), and we study the introduction of thermophoresis, diffusiophoresis and electric effects in the below-cloud scavenging (simulation SIM2_BCPLUS in Table 4.5). The simulations cover the globe for the year 2007 and use dynamics from 3-hourly meteorological fields from ARPEGE analyses downgraded to a resolution of the model ($2^\circ \times 2^\circ$). We have only primary aerosols in the model. Thus, to compare the model outputs with observations, we focus on the regions where primary aerosols dominate the aerosol optical depth field, and on strong, high-concentration aerosol events near the sources where we can presume that the contribution of other aerosols is minimal. Inspecting the averaged quantities (annual budget, burden, lifetime, emissions, depositions) allow us to evaluate the relative importance of different parameterizations and processes.

4.5.6 Results

In this section we evaluate MOCAGE SIM1 and SIM2 output and compare it to independent data. Figures 4.4 and 4.5 present the effects of the model updates, by showing horizontal geographical and vertical zonal distribution of aerosol species in MOCAGE for the SIM1 and SIM2 simulations. As shown in Fig. 4.4, the changes to the model in SIM2 compared to SIM1, resulted in

Table 4.5: Description of MOCAGE simulations used in this study.

Simulation	Description	
1. SIM1	The reference simulation using the current MOCAGE configuration	
	in-cloud scavenging:	the Langner and Rodhe (1991) scheme
	below-cloud scavenging:	the Slinn (1977) scheme with fixed raindrop size and Stoke's regime terminal raindrop velocity
	emissions	
	– sea salt:	the Gong (2003) source function
	– desert dust:	the Marticorena et al. (1997) and Laurent et al. (2006) schemes with the nearest-neighbour wind interpolation
	– carbonaceous aerosols:	AeroCom + GFED3 emissions
2. SIM2	The reference simulation using the updated model configuration	
	in-cloud scavenging:	the Giorgi and Chameides (1986) scheme and precipitation cloud cover
	below-cloud scavenging:	the Slinn (1977) rainfall scheme with the exponential raindrop size distribution, the parameterized terminal raindrop velocity and the precipitation re-evaporation; the Slinn (1977, 1982a) snowfall scheme
	emissions	
	– sea salt:	the Jaeglé et al. (2011) source function
	– desert dust:	the Marticorena et al. (1997) and Laurent et al. (2006) schemes with the bilinear wind interpolation and the Alfaro et al. (1998) desert dust initial distribution
	– carbonaceous aerosols:	Lamarque et al. (2010) + GFED3 emissions
	sedimentation:	introduction of Sutherland's law + stability check
3. SIM2_WDEP	As SIM2, but wet deposition module as in SIM1	
4. SIM2_SED	As SIM2, but sedimentation module as in SIM1	
5. SIM2_EMI	As SIM2, but emissions as in SIM1	
6. SIM2_BCPLUS	As SIM2 plus thermophoretic, diffusiophoretic and electric charge effects in the below-cloud scavenging scheme	

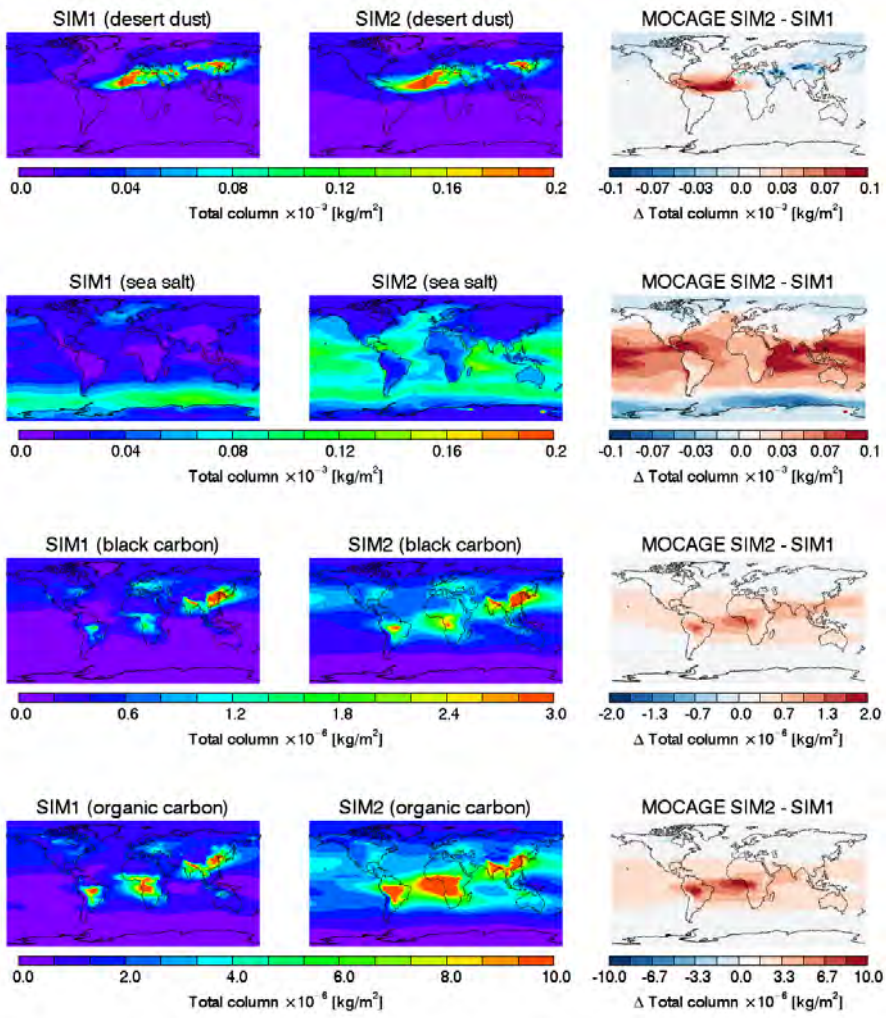

Figure 4.4: The geographic distribution of the mean annual burdens of all aerosol species in the CTM MOCAGE: for SIM1 on the left, for SIM2 in the middle, and their difference on the right.

less desert dust aerosols near sources in Asia and northern Africa, but more in the south-eastern part of the Sahara. Also, more aerosols are transported over the Atlantic, with the long-range transport eased by the shift in the initial size distribution towards smaller sizes in SIM2 (Sect. 4.5.3.4). Sea salt aerosols are more abundant globally in SIM2 compared to SIM1. Over cold waters, especially over southern oceans, we note a decrease and over warm waters an increase in the sea salt burden. This shift is mainly due to the introduction of the SST dependency in the sea salt emission scheme in SIM2. Having the

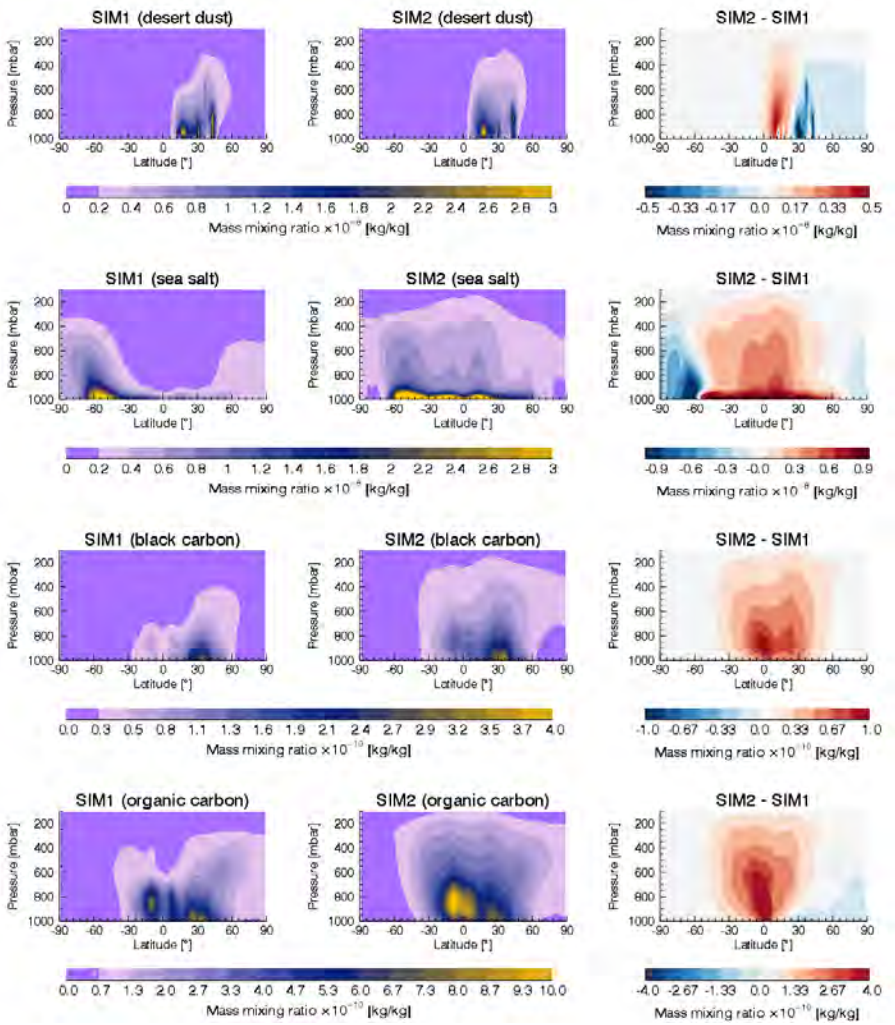

Figure 4.5: The annual and zonal mean vertical profiles of mass mixing ratios of all aerosol species in the CTM MOCAGE: for SIM1 on the left, for SIM2 in the middle, and their difference on the right.

black carbon and organic carbon emissions quite similar in SIM1 and SIM2, the differences between SIM1 and SIM2 reflect mainly the changes in the wet deposition scheme. The increase in their burden in SIM2 is the outcome of the weaker wet deposition in total in SIM2 than in SIM1. Figure 4.5 confirms these findings and, although a number of effects influence the mass mixing ratios, one can see that the updates generally produced more desert dust and sea salt aerosols toward higher altitudes. Regarding black and organic carbon aerosols, Fig. 4.5 shows their higher concentrations in the free troposphere in SIM2 than

in SIM1. This is the result of the weaker wet deposition in SIM2 than in SIM1 and of the shift in the wet deposition vertical distribution by having a weaker below-cloud scavenging and a stronger in-cloud scavenging in SIM2 compared to SIM1.

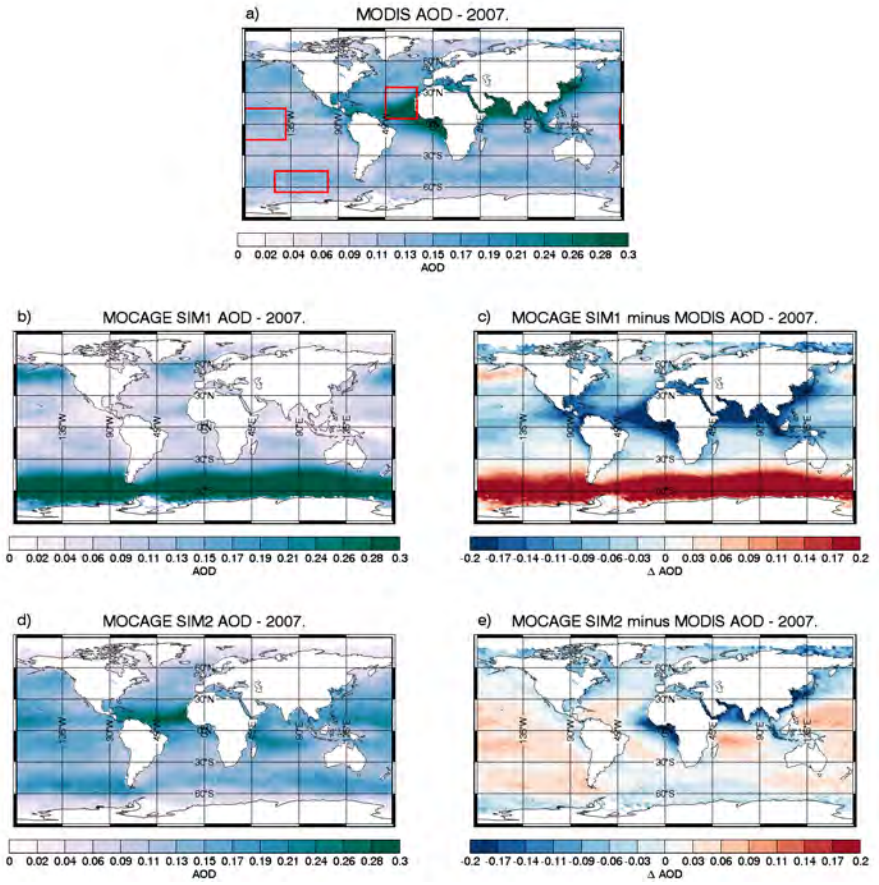

Figure 4.6: Global, mean aerosol optical depth at 550 nm for the year 2007 from MODIS (Aqua + Terra) **(a)**, SIM1 **(b)**, SIM2 **(d)**, and the difference between MODIS observations and model simulations **(c**, **e)**. The descriptions of the model simulations are in Table 4.5. The boxes in panel **(a)** correspond to the regions used in Fig. 4.8.

In Fig. 4.6, SIM1 and SIM2 AOD fields are compared with global yearly averaged MODIS AOD. Model AOD values are only sampled in the case of available MODIS observations on a particular day. Overall, SIM2 shows a significant improvement over SIM1 in terms of AOD. The modified normalized mean bias is decreased from 0.42 to 0.10 and the correlation is improved from 0.06 to 0.32 (Figs. 4.6 and 4.7; Table 4.6). The improvement is especially apparent in mid- to high-latitude Southern Hemisphere oceans (where the

modified normalized mean bias is lowered from 0.65 to 0.16) and the African dust outflow region (the modified normalized mean bias improved from −1.01 to −0.22). Near coasts, where the influence from the land is stronger, both model simulations underestimate AOD. This could be due to the absence of secondary aerosols in the model. The effect is more evident near south-eastern Asia, India, the Arabian Peninsula and in the Gulf of Guinea, and is less pronounced in SIM2 due to the changes in primary aerosol parameterizations. The cause of discrepancy over the Gulf of Guinea is not clear and a similar pattern is observed by Jaeglé et al. (2011) in the GEOS-Chem model. In MOCAGE, it could be due to the missing secondary aerosols, the insufficient biomass burning aerosol concentration or possibly the cloud contamination in the MODIS data. Another possibility that is less likely is the inaccurate sea salt emissions due to possible wind errors in the ARPEGE analysis; however, considering the low wind speeds in the region (Fig. 4.3) we do not expect a lot of sea salt particles. In the tropical oceans, compared to MODIS, model AOD shifted from a negative bias in SIM1 to a positive bias in SIM2. The results for SIM2 were significantly better, but the model still overestimates AOD with discrepancies that are larger than the MODIS expected error.

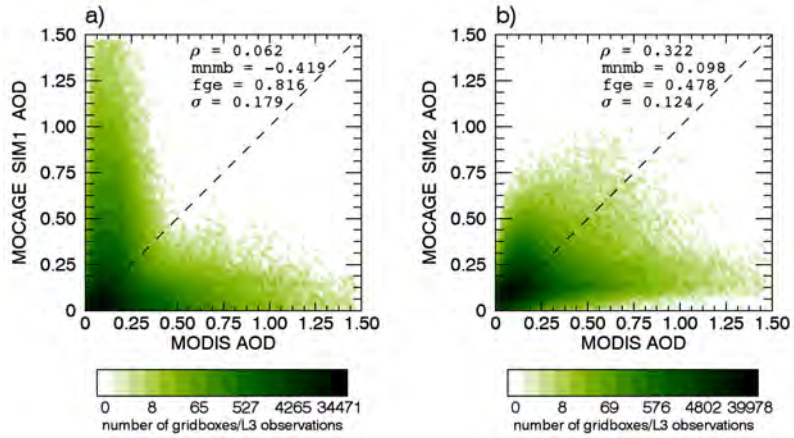

Figure 4.7: Scatterplots of aerosol optical depths from MODIS and the simulations: SIM1 **(a)**, SIM2 **(b)**. Scatterplots are contoured according to the number of the points in them. Each point in the scatterplot presents MODIS L3 observed AOD and the corresponding modelled AOD. In each panel, correlation (ρ), modified normalized mean bias (MNMB), fractional gross error (FGE) and standard deviation (σ) are noted. The descriptions of the model simulations are in Table 4.5.

The relationship between model simulations and observations are presented in Fig. 4.7. This figure confirms the improvement in the AOD field in SIM2 compared to SIM1, but with discrepancies with observations visible in both simulations. As we performed a strong quality control of the MODIS data,

we presume that these discrepancies are related to the model performance. Having in mind also Fig. 4.6, SIM1 (Fig. 4.7a) shows strong signatures of overestimated sea salt AOD, a lack of secondary aerosols and an underestimation of desert dust particles. SIM2 (Fig. 4.7b) has significantly better statistics, a better correlation and smaller standard deviation relative to the observations, but still displays the strong signature of the missing secondary aerosols.

Table 4.6: Number of observations, correlation (ρ), modified normalized mean bias (MNMB) and fractional gross error (FGE) between observations (MODIS and AERONET) and SIM1/SIM2. The number of MODIS observations includes the number of considered L3 gridboxes, and the corresponding number of L2 observations. EMEP observations are of hourly or daily frequency. MODIS regions correspond to Fig. 4.8a–c, and AERONET sites correspond to Fig. 4.9a–f.

	No. obs.		SIM1			SIM2		
			ρ	MNMB	FGE	ρ	MNMB	FGE
MODIS	L3	L2						
African dust outflow region	84 272	8.6×10^6	0.76	−1.009	1.009	0.797	−0.222	0.268
Tropical Pacific	91 322	9.8×10^6	0.647	−0.715	0.716	0.689	0.267	0.268
South Pacific*	23 687	3.0×10^6	0.334	0.652	0.676	0.363	0.158	0.278
AERONET		L2						
Tenerife Santa Cruz		5033	0.553	−0.527	0.663	0.687	0.192	0.447
Cabo Verde		5389	0.587	−1.019	1.034	0.632	−0.216	0.449
Nauru		3040	0.074	−1.508	1.519	0.217	0.513	0.564
Tahiti		1328	0.091	−0.697	0.989	0.277	0.805	0.813
Amsterdam Island		933	0.204	0.703	0.778	0.269	0.501	0.582
Crozet Island		361	0.076	1.161	1.168	0.181	0.644	0.723
EMEP								
Hyytiälä, FI (P2.5)		140	0.059	−1.236	1.24	0.545	−0.778	0.785
Lille Valby, DK (P2.5)		327	0.041	−1.02	1.041	0.042	−0.262	0.518
Ayia Marina, CY (P10)		302	0.266	−1.787	1.787	0.312	−0.374	0.602
Auchencorth Moss, GB (P10)		8428	0.064	−1.003	1.471	0.197	−0.706	1.106
Zingst, DE (P10)		333	−0.121	−0.904	0.939	−0.138	0.350	0.70

* Statistics calculated excluding the winter months because of very few observations

Figure 4.8 presents the temporal variability comparison of model simulations with MODIS observations over the selected regions, where primary aerosols dominate the AOD throughout the year and which are large enough to cover a statistically meaningful number of observations (usually thousands of observations per day). This figure confirms the positive effect due to the updates in the model parameterizations (statistics of Fig. 4.8 shown in Table 4.6). In the Saharan desert dust outflow region over the Atlantic (Fig. 4.8a), SIM2 agrees better with MODIS than SIM1, but with some underestimation of AOD in both simulations. We improved the intensities of the stronger dust events and overall correlation, and lowered the bias. Over the tropical waters of the central Pacific, SIM2 shows a slight statistical improvement (Fig. 4.8b): while SIM2 overestimates, SIM1 underestimates AOD. In the high-wind South Pacific region (Fig. 4.8c), SIM2 greatly improves the AOD values and reduces the bias. Correlations between the observations and the simulated AOD are smaller

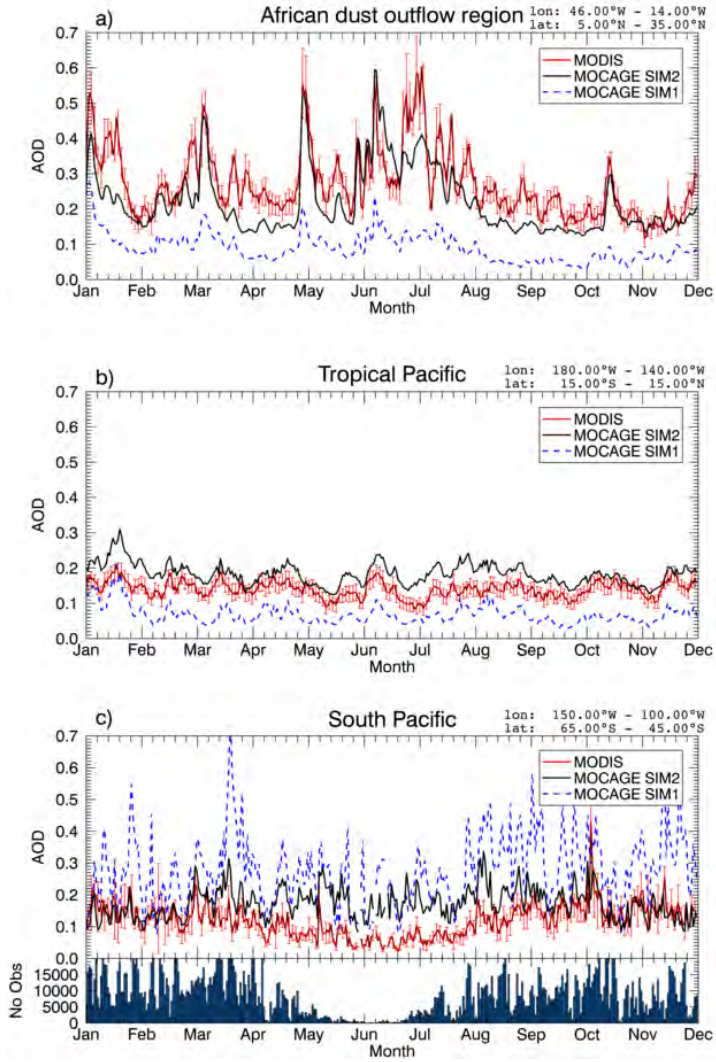

Figure 4.8: Time series of aerosol optical depth at 550 nm in 2007 of MODIS (Aqua + Terra) data, SIM1 and SIM2 over **(a)** the African desert dust outflow region (45–15° W, 5–35° N), **(b)** the tropical Pacific (180–140° W, 15° S–15° N), and **(c)** the South Pacific (150–100° W, 65–45° S). The regions are also marked in Fig. 4.6a. For the South Pacific region, the number of observations over the region is given for each day. Correlation, modified normalized mean bias and fractional gross error for both SIM1 and SIM2 as compared to MODIS data are given in Table 4.6. The descriptions of model simulations are in Table 4.5.

than in the other regions, which is possibly due to wind errors present in the ARPEGE analysis for this remote part of the world. However, by taking into

account the whole year data, SIM1 correlates better with MODIS than SIM2. The cause is a minimum in AOD in the Southern Hemisphere winter visible in the MODIS data, which is not present in the model. The noted minimum in the data is determined by only a small number of satellite observations (there are even days without observations over the whole region because of high cloudiness). Thus, statistical confidence in the observations over that period is low. In the model, winds (Fig. 4.3) and sea surface temperature in this region do not show important systematic errors and are therefore probably not responsible for the discrepancy. If we exclude the effect of the observed winter minimum from our analysis, correlations in SIM2 are superior to SIM1 (0.33 in SIM1, 0.36 in SIM2), which demonstrates the improvement in the representation of aerosols in this part of the globe.

We also compared the model AOD with the independent data set from AERONET for 2007 (Fig. 4.9). AERONET data are very accurate and often used for the validation of satellite data (Remer et al., 2005; Kahn et al., 2005; Schuster et al., 2012). However, the horizontal representativity of AERONET data is much smaller compared to that of satellite data. The data are less adapted to make comparisons with the model than satellite data – it is localized in a single spot for each station compared to the $2^\circ \times 2^\circ$ model data. It may be preferable to do multi-year analysis to improve statistics since some stations do not have the whole year record, and observations are especially scarce in the wintertime. For our study, we chose the stations with available observations where primary aerosols dominate AOD. The AERONET observations confirmed the findings from the comparison with MODIS (Fig. 4.9, with statistics shown in Table 4.6): SIM2 reduced the AOD underestimation in the African dust outflow region (stations on Tenerife and Cabo Verde), reduced sea salt overestimation in mid- and high-latitude regions (Amsterdam Island and Crozet Island), and had a minor impact on the absolute value of the bias – but changed its sign – over tropical regions (Nauru and Tahiti). We noted that AERONET stations on the oceanic islands show smaller AOD values than MODIS.

In Fig. 4.10 we compare the model simulation with the independent data from SEVIRI. We used the daily averaged only-land SEVIRI data (Carrer et al., 2010) to analyse an AOD field over Europe on a day (23 May 2007) when several strong primary aerosol events dominated the AOD field: several desert dust plumes visible over southern and central Europe, and sea salt aerosols to the north of the British Isles. In both model simulations we see the same AOD features, but they differ in intensity. The location and extent of the features in the model correspond well with the SEVIRI field, except that the desert dust plume over eastern Europe in the model is located more to the south. The AOD values in SIM2 are much closer to the SEVIRI data than in SIM1. Low background AOD values in the model reveal a systematic underestimation over continents. This could be due to an absence of secondary aerosols.

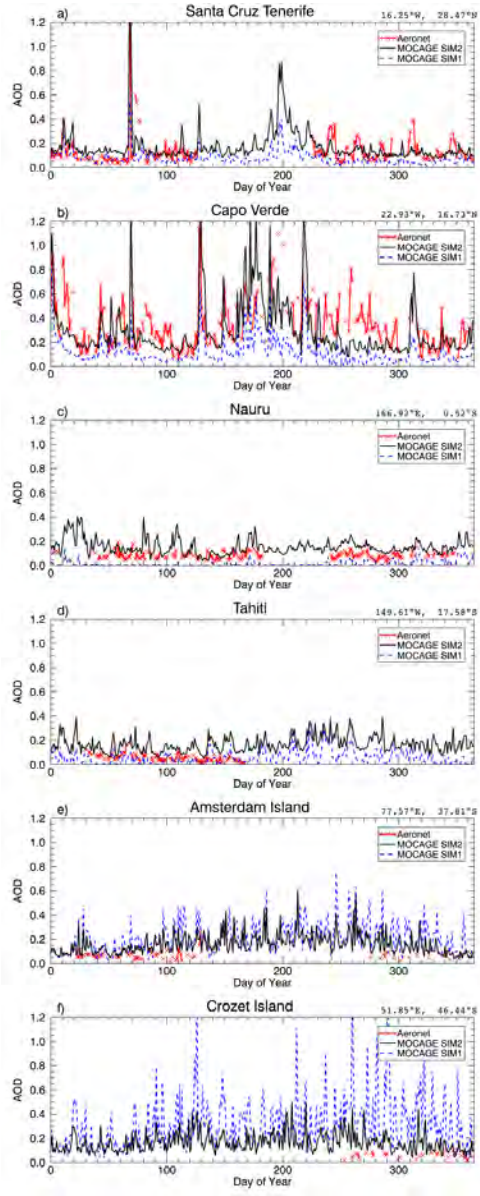

Figure 4.9: Time series of aerosol optical depth at 550 nm from the AERONET data, SIM1 and SIM2 for six AERONET stations: Tenerife Santa Cruz (16.25° W, 28.47° N), Cabo Verde (22.93° W, 16.73° N), Nauru (166.92° W, 0.52° S), Tahiti (149.61° W, 17.58° S), Amsterdam Island (77.57° E, 37.81° S) and Crozet Island (51.85° E, 46.44° S). Correlation, modified normalized mean bias and fractional gross error for both SIM1 and SIM2 compared to AERONET observations are given in Table 4.6. The descriptions of the model simulations are in Table 4.5.

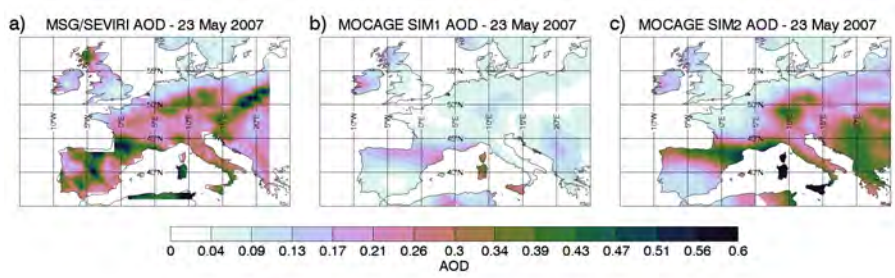

Figure 4.10: Aerosol optical depth fields over Europe for 23 May 2007 at 550 nm from SEVIRI **(a)**, SIM1 **(b)**, and SIM2 **(c)** simulations.

Besides AOD observations, we assess the MOCAGE performance with the particulate matter measurements from the EMEP surface network. When considering the EMEP network, the majority of stations are in or near urban zones where the signature of secondary aerosols is strong. Therefore, we use the measurements from selected stations which are chosen so that their locations are near coasts where usually sea salt aerosols dominate or in sites far from the urban zones. Figure 4.11 and Table 4.6 show how SIM1 and SIM2 compare against EMEP measurements from the selected stations. The comparison shows slight to significant differences due to the model updates, and confirms the overall improvement to the model performance.

Table 4.7 shows how the MOCAGE simulations compare to data from the AeroCom model inter-comparison (`http://aerocom.met.no/`, Textor et al., 2006, 2007). AeroCom data are not based on observations but is an independent data set which indicates how MOCAGE relates to performances of other models.

Table 4.7: Globally averaged annual burden, lifetime and emissions in SIM1 and SIM2 for individual aerosols species (DD – desert dust, SS – sea salt, BC – black carbon), compared to data from AeroCom project (Dentener et al., 2006; Textor et al., 2006). For a description of model simulations, see Table 4.5.

	SIM1			AeroCom		
	DD	SS	BC	DD	SS	BC
Burden (Tg)	9.66	9.70	0.24	$19.2 \pm 40\,\%$	$7.52 \pm 54\,\%$	$0.24 \pm 42\,\%$
Lifetime (days)	1.0	3.0	10.0	$4.1 \pm 43\,\%$	$0.5 \pm 58\,\%$	$7.1 \pm 33\,\%$
Emissions ($\mathrm{Tg\,yr^{-1}}$)	3476	1180	8.89	1678	7925	7.7
	SIM2					
	DD	SS	BC			
Burden (Tg)	11.2	34.1	0.34			
Lifetime (days)	2.9	1.5	14.2			
Emissions ($\mathrm{Tg\,yr^{-1}}$)	1395	8274	8.82			

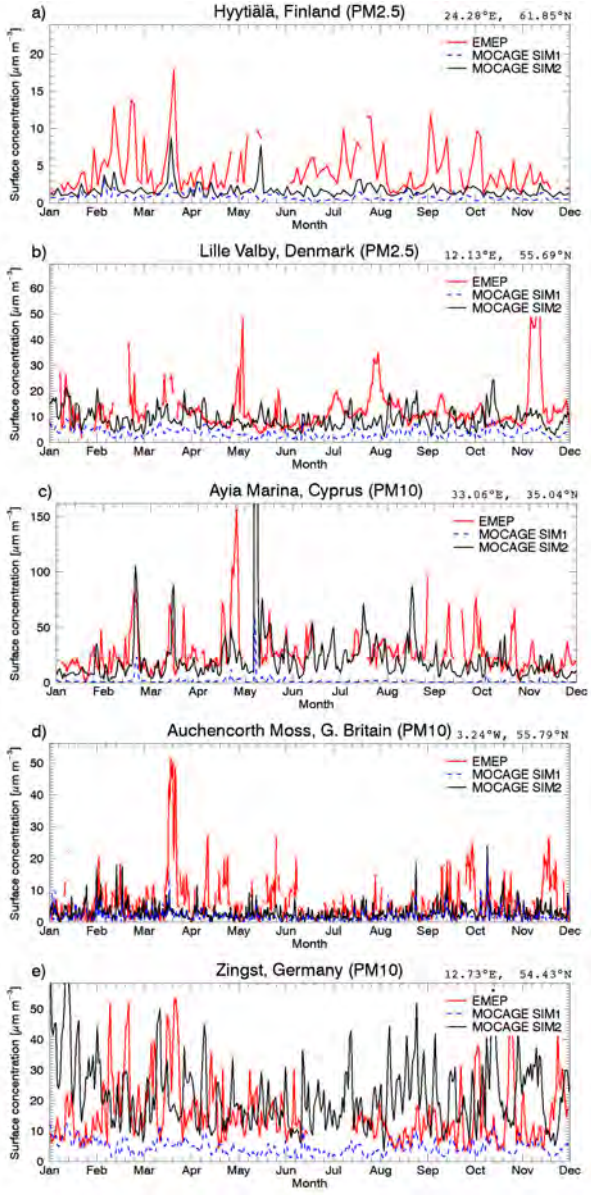

Figure 4.11: Time series of aerosol PM for 2007 from EMEP data, and for SIM1 and SIM2 for five EMEP stations: Hyytiälä, Finland ($PM_{2.5}$, 24.28° E, 61.85° N), Lille Valby, Denmark ($PM_{2.5}$, 12.13° E, 55.69° N), Ayia Marina, Cyprus (PM_{10}, 33.06° E, 35.04° N), Auchencorth Moss, Great Britain (PM_{10}, 3.24° W, 55.79° N) and Zingst, Germany (PM_{10}, 12.73° E, 54.43° N). Correlation, modified normalized mean bias and fractional gross error for both SIM1 and SIM2 compared to EMEP observations are given in Table 4.6. The descriptions of the model simulations are in Table 4.5.

Values from SIM2 compare better to AeroCom ranges, by improving several parameters over SIM1. Emitted quantities fit better in SIM2 and there is an improvement in desert dust and sea salt lifetimes as well. Black carbon emissions correspond well to the AeroCom model average. Both SIM1 and SIM2 black carbon burdens are within the AeroCom range, but the lifetime is by a factor of 2 larger in SIM2 than in AeroCom, which could indicate weak wet deposition in the regions of high black carbon concentrations in SIM2. The sea salt burden in SIM2 is larger than in SIM1, but the lifetime is improved in SIM2.

In summary, observations from MODIS, AERONET, SEVIRI and EMEP showed that changes in the aerosol parameterizations improved the model performance. SIM2 shows a significantly better agreement in AOD compared with different types of observations relative to SIM1, and this is confirmed by in situ observations.

4.5.6.1 Sensitivity to new parameterization components

The updates to the parameterizations, which are collectively compared to the observations in the section above, have different and separate effects on the model results. In this section we analyse separate impacts of the updates by dividing them into the three most important components: changes in emissions of sea salt and desert dust aerosols, in sedimentation of particles, and in wet deposition. In Fig. 4.12, simulations SIM2_SED, SIM2_EMI, and SIM2_WDEP are compared with the reference SIM2 run. This figure demonstrates that the improvements in the sedimentation make a modest overall change and that the changes to the emissions and wet deposition changes impact the results much more strongly. The total annual sedimentation in SIM2 decreased by 22 %, but this change influenced AOD only moderately: the results of the SIM2 and SIM2_WDEP simulations are very similar with the high correlation between them (0.92; Fig. 4.12a). In the atmospheric surface layer, sedimentation acts in conceit with dry deposition, and the impacts due to the changes to each process tend to compensate one another (Table 4.8).

Figure 4.12b presents the changes and major improvements in SIM2 that result from the modifications to the emissions compared to SIM2_EMI. The two distinct populations of points in the scatterplot represent overestimated sea salt particles and underestimated desert dust. In addition, both populations are likely affected by the missing secondary aerosols. In the SIM2 emissions, the desert dust aerosol distribution is shifted towards smaller diameters making the sedimentation process less important for aerosol removal, and consequently

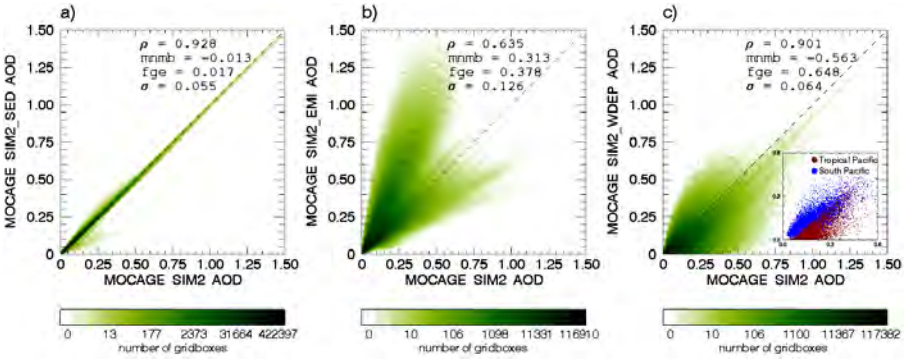

Figure 4.12: Scatterplots of aerosol optical depth from the model reference run SIM2 and the simulations: SIM2_SED **(a)**, SIM2_EMI **(b)**, and SIM2_WDEP **(c)**. These scatterplots show the impact of different model updates to the model performance and they are contoured according to the number of the points. Each point in the scatterplot presents modelled AOD in two corresponding simulations. In each panel, ρ, MNMB, FGE, and σ are noted. For the SIM2_WDEP simulation, a subgraph is presented showing the differences between the tropical Pacific and South Pacific regions (regions shown in Fig. 4.6a). The description of the model simulations is in Table 4.5.

their lifetimes are $\approx 50\,\%$ longer. The sea salt particle emissions in SIM2 are 7 times larger than in SIM2_EMI, which makes their burden larger in SIM2. Also, their global distribution changed – there are more particles in low and midlatitudes, which makes their lifetime shorter. Although emitted sea salt quantities hugely vary between different estimates (from 1000 to 30 000 $\mathrm{Tg\,yr^{-1}}$; Lewis and Schwartz, 2004a), emissions in MOCAGE are in agreement with the "best" estimate of Lewis and Schwartz (2004a) of 5000 $\mathrm{Tg\,yr^{-1}}$ (estimate uncertainty of the factor of 4) and with AeroCom data (Table 4.7). Desert dust aerosols are emitted by a factor of 2–3 less in SIM2 than in SIM2_EMI, with the decrease mostly in Asian deserts. The new value agrees better with the AeroCom estimate (Table 4.7). The change of wind interpolation in the desert dust emission schemes more strongly affected Asian desert dust because of the finer resolution of the scheme and the rougher topography present in this region. The differences between AeroCom and Lamarque et al. (2010) inventories for carbonaceous aerosols did not produce variation.

Figure 4.12c shows the impact of the wet deposition changes in the model between the SIM2 and SIM2_WDEP simulations. The two simulations are strongly correlated both temporally and spatially but they show important differences in AOD. Compared to SIM2, the below-cloud scavenging is overall stronger in SIM2_WDEP mainly due to the higher precipitating cloud fraction in SIM2_WDEP and missing precipitation re-evaporation (which is only introduced in SIM2). However, the AOD in SIM2 becomes both larger and

Table 4.8: Globally averaged annual burden, lifetime, emissions, and deposited mass due to wet deposition, dry surface deposition and sedimentation for different aerosol types (DD – desert dust, SS – sea salt, BC – black carbon, OC – organic carbon) in different model simulations to reveal the separate effects of different model updates. For a description of model simulations, see Table 4.5.

	SIM2				SIM2_BCPLUS			
	DD	SS	BC	OC	DD	SS	BC	OC
Burden (Tg)	11.2	34.1	0.34	1.74	11.1	33.6	0.34	1.72
Lifetime (days)	2.93	1.50	14.2	19.3	2.90	1.48	14.0	19.0
Emissions ($Tg\,yr^{-1}$)	1395	8274	8.82	33.0	1395	8274	8.82	33.0
Dry deposition ($Tg\,yr^{-1}$)	860	2689	3.23	9.71	859	2684	3.22	9.64
Sedimentation ($Tg\,yr^{-1}$)	317	3772	0.01	0.06	317	3766	0.01	0.06
Wet deposition ($Tg\,yr^{-1}$)	199	1759	5.53	23.2	200	1771	5.6	23.2
In-cloud scav. (%/100)	0.75	0.57	0.96	0.97	0.73	0.56	0.94	0.95
Below-cloud scav. (%/100)	0.25	0.43	0.04	0.03	0.27	0.44	0.06	0.05
	SIM2_EMI				SIM2_WETDEP			
	DD	SS	BC	OC	DD	SS	BC	OC
Burden (Tg)	14.4	15.5	0.45	1.53	8.9	28.0	0.24	1.21
Lifetime (days)	1.51	4.79	16.5	21.0	2.32	1.23	10.1	13.4
Emissions ($Tg\,yr^{-1}$)	3476	1180	9.89	18.0	1395	8274	8.82	33.0
Dry deposition ($Tg\,yr^{-1}$)	1824	344	3.29	4.2	867	2605	2.8	8.1
Sedimentation ($Tg\,yr^{-1}$)	1328	318	0.01	0.04	306	3715	0.01	0.05
Wet deposition ($Tg\,yr^{-1}$)	305	534	6.4	14.1	184	1908	6.1	25.3
	SIM2_SED							
	DD	SS	BC	OC				
Burden (Tg)	10.9	32.4	0.34	1.74				
Lifetime (days)	2.84	1.43	14.2	19.3				
Emissions ($Tg\,yr^{-1}$)	1395	8274	8.82	33.0				
Dry deposition ($Tg\,yr^{-1}$)	670	1912	3.23	9.71				
Sedimentation ($Tg\,yr^{-1}$)	521	4742	0.01	0.06				
Wet deposition ($Tg\,yr^{-1}$)	186	1576	5.53	23.2				

smaller in different situations; it decreased and increased depending on location with an overall tendency for weaker wet deposition in SIM2 (also shown in Table 4.8). In tropical regions, where convective systems are the cause of the majority of the scavenging and where re-evaporation has an important impact, aerosol particles are scavenged less in the SIM2 than in SIM2_WDEP (see the subgraph in Fig. 4.12c). Re-evaporation of precipitation effectively mitigates the wash-out of aerosols and in SIM2 it reintroduced into the atmosphere 9 % of aerosols scavenged by convective precipitation and 10 % of aerosols scavenged by stratiform precipitation. In the midlatitudes, the re-evaporation is less important and the cloud cover is a more important factor. In this region, the changes in the precipitating cloud fraction and other wet deposition updates made the wet scavenging a more powerful process in SIM2 than in SIM2_WDEP (subgraph in Fig. 4.12c). However, globally, the changes in the wet deposition

scheme resulted in 5 % less aerosols scavenged by wet deposition in SIM2 than in SIM2_WDEP. Modifications of the below-cloud scavenging scheme also included additional scavenging processes (thermophoretic, diffusiophoretic and electric charge effects) proposed in the literature (Andronache et al., 2006) and which are introduced in the SIM2_BCPLUS simulation. The additional processes moderately changed the efficiency of the below-cloud scavenging (Table 4.8). Scavenging increased by 5 %, but this only minimally influenced the resulting AOD field.

4.5.7 Discussion

The updated parameterizations improve the aerosol representation in the model and agree better with observations independent from one another. Compared to observations, the updated model still shows some overestimation over the sea salt dominated regions and an underestimation over the Atlantic region affected by the African desert dust outflow. The identified differences in AOD between the model and observations exceed prescribed observation errors and their degree is consistent with the results of other studies: Zhang et al. (2012a) with the ECHAM-HAM model compared to MODIS observations, Jaeglé et al. (2011) with the GEOS-CHEM model compared to both MODIS and AERONET observations, Su et al. (2013) using the GOCART model compared to the MODIS/MATCH AOD field. Zhang et al. (2012a) found that simulated AOD over sea salt regions was overestimated to a similar degree as with MOCAGE, while Saharan outflow desert dust AOD was overestimated with an absolute difference of greater than a factor of 2, as compared to MOCAGE. Jaeglé et al. (2011) found that AOD over sea salt regions of the global oceans was underestimated by less than 0.04 and over the African dust outflow region it was overestimated with the absolute difference greater by a factor of 2–3, as compared to MOCAGE. Su et al. (2013) compared GOCART with the assimilated MODIS/MATCH AOD that was "constrained to a large extent by MODIS" and found that AOD over the sea salt regions was overestimated slightly more than in MOCAGE, and that AOD over the African dust outflow region was underestimated a little less than in MOCAGE.

We noted in the previous paragraph that the present-day state-of-the-art models have similar performance compared to MOCAGE. Regarding this study, the biases could have different causes, and we should concentrate our further model developments to deal with these issues. Concerning desert dust aerosols, the peaks of the most intense desert dust events are well reproduced in MOCAGE, but in days with more moderate dust production we notice

weaker model AOD than in the observations. These weaker AOD values over the African dust outflow region were found both near and far from the sources, which hints that emissions of African desert dust may be too small. Wind uncertainties could be important in this region, which could lead to less fugitive sand and dust, or the soil characterization in the scheme might need a refinement (e.g. better resolution, satellite retrieved soil type/properties) (Laurent et al., 2008a,b; Bouet et al., 2012).

The sea salt discrepancy between MOCAGE and observations can possibly be caused by several factors: too high emissions, too weak below-cloud scavenging, and the missing sea salt chemical evolution in the model. First, we examine the possibility that the high sea salt burden results from emissions that are too large. Emitted sea salt quantities are in agreement with the AeroCom model average (Table 4.7), but the very large range in emissions in AeroCom indicates large uncertainties (Textor et al., 2007). Jaeglé et al. (2011) clearly showed the sea salt emission dependency on sea surface temperature, but their parameterization could be model dependent because they derived it by minimizing the bias of their model relative to in situ observations. Models could vary significantly and it might be necessary to separately fit the parameters of the Jaeglé et al. (2011) function to the individual model employed (which Jaeglé et al., 2011, also noted). This idea is supported by results from Spada et al. (2013), who implemented the sea salt function from Jaeglé et al. (2011) in the NMMB/BSC-CTM model and found the sea salt is overestimated in the tropical regions. Still, the parameterization depending on sea surface temperature undoubtedly improved the performance of MOCAGE.

The ratio of wet deposition to the total dry deposition (surface dry deposition + sedimentation) measured on cruise ships is 0.3 / 0.7 (Jaeglé et al., 2011), which corresponds well to the results from MOCAGE (Table 4.8). However, the longer mean atmospheric residence time of sea salt particles compared to the AeroCom model average could indicate that the wet deposition and, in particular, below-cloud scavenging, might be underestimated. The below-cloud scavenging is an efficient, episodic process, generally located near to sources, which can strongly influence the residence times of aerosols (Croft et al., 2009), and it is directly proportional to the precipitation intensity. The long lifetime of black carbon aerosols in the model can also indicate that wet deposition – by far the most important sink for black carbon particles (Textor et al., 2006) – could be too weak in MOCAGE. Compared with the data from the Global Precipitation Climatology Project, which is based on ground and satellite observations (Adler et al., 2003), the mean zonal distribution of precipitation in MOCAGE is correctly located, but its intensity is lower for $\approx 25\,\%$ (Fig. 4.13). This affects the simulated quantities that are scavenged and could lead to a longer residence time in MOCAGE than in the AeroCom model average.

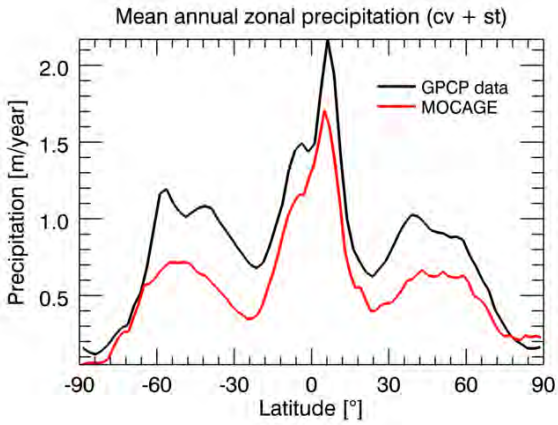

Figure 4.13: Mean annual zonal precipitation quantity (combined stratiform (st) and convective (cv) precipitation) from GPCP data and MOCAGE.

The chemical evolution of the sea salt aerosols could have an important impact on the sea salt burden (Lewis and Schwartz, 2004a). The tests of the secondary aerosol module performed in MOCAGE show that the dechlorination could be efficient in lowering the sea salt burden (and lifetimes) obtained in this study. Still, the whole impact of the reactions with sea salt aerosols will be possible to evaluate with the secondary inorganic aerosol module validated in the model.

Secondary aerosols can certainly account for the discrepancies between the model and the observations in the zones where anthropogenic aerosols have a major influence, as already discussed. However, the so-called unspecified primary anthropogenic aerosols can also play a role, but the secondary aerosols should have a stronger influence on AOD. The unspecified primary anthropogenic aerosols are not implemented in the configuration of the model used in this study, because they are not present in the emission inventories that we used, but they can be found in some models (e.g. Matthias, 2008).

Updates in the emissions created the largest improvement in our model. But in other studies, uncertainties in the other aerosol parameterizations are found to be bigger than in emissions (Textor et al., 2007). This is backed by the differences in the scavenged aerosols simulated by two different in-cloud scavenging schemes presented in SIM1 and SIM2 that are about 25 %. This implies that adding other refinements and aiming for more physically realistic parameterizations would likely further improve the model performance.

Inclusion of secondary aerosols will be the most crucial addition, it would make the aerosol family more complete and improve the model performance over regions where secondary aerosols and chemical reactions with aerosols play a major role.

4.5.8 Summary and conclusion

In this paper we introduced the improvements to the aerosol module in the chemical transport model MOCAGE and evaluated the impact on aerosol representation, properties, and global distribution. The ambition was to solve already known model biases and to have more physically realistic aerosol parameterizations. The updates include changes in emissions, wet deposition, and sedimentation. Regarding emissions, we added a SST dependence to the sea salt source function, and adjusted the size distribution (and the wind speed calculation) in the desert dust emission scheme. In the wet deposition scheme we used a new precipitation cloud cover calculation and in-cloud scavenging scheme. We also developed the below-cloud scavenging scheme by revising the calculation of raindrop size and terminal velocity, and by introducing re-evaporation and snowfall scavenging. The sedimentation module update strengthened the performance of the scheme: for example, the model demonstrated better mass conservation. The emission and wet deposition changes produced a stronger impact, while updates in sedimentation produced a less pronounced effect. Emission changes directly influenced known biases of sea salt and African desert dust aerosols, while the impact of wet deposition update is more complex and balanced – depending on the location, it decreased or increased AOD. The effects of the wet deposition updates vary widely, both temporally and spatially, mainly because the wet deposition depends on both the presence of aerosols and the occurrence of precipitation. Examples of the changes in the model field are the increase of AOD in tropical oceans due to introduced re-evaporation in SIM2 compared to SIM1, and the decrease in southern midlatitude oceans due to the changes in the precipitating cloud cover fraction and other updates in the wet deposition scheme.

We evaluated the impacts of these changes and compared them to AOD observations from satellite sensors (MODIS, SEVIRI), the AERONET and the EMEP stations, and the AeroCom model inter-comparison. Since in our model only primary aerosols are present, we focused the analysis on the regions where mainly primary aerosols dominate AOD. Compared to the model simulation with old parameterizations, we significantly improve agreement with the observations and the AeroCom data (Tables 4.6 and 4.7). The sea salt

and desert dust emitted quantities correspond better to both estimates from the literature and the model average from the AeroCom project (Table 4.7). The shift toward smaller particles in the desert dust size distribution and the modified geographical distribution of sea salt emissions had a positive impact on aerosol lifetimes. We examined the spatial and temporal variability of AOD and showed that the SST-dependent emissions solved the strong positive bias in sea salt aerosols in mid- to high latitudes that were previously seen in our model (Fig. 4.6). This lead to a lower AOD over these regions and stronger AOD values over the tropics, which agrees better with observations. In the Saharan desert dust aerosol outflow region, we reduced the bias and improved the correlation and intensity of the stronger events (Table 4.6). Overall, the updates had a positive effect on the correlation with observations. Quantitatively, as an example, in the comparison with MODIS observations on the global scale, the update of parameterizations improved correlation from 0.06 to 0.32. The comparison with particulate matter $PM_{2.5}/PM_{10}$ measurements from the EMEP network showed that in urban zones the model underestimates aerosols, but confirmed the findings obtained from the comparison with AOD measurements that the model updates have positively impacted the model performance.

The obtained results confirmed that large uncertainties in models can come from the use of parameterizations. Significant differences in parameterization formulations lead to big differences in model outputs, as also confirmed in the literature (Textor et al., 2007). Two different in-cloud scavenging schemes used in this study had efficiencies that differed by a factor of 2, and a few changes in different components in our semi-empirical below-cloud scavenging scheme produce very different results in the same scheme.

We found that the introduced updates enhanced the model performance, but some discrepancies with the observations remain: (a) underestimation in the regions where secondary aerosols could have an important impact, (b) some overestimation of sea salt aerosols, and (c) some underestimation of African desert dust aerosols. The future work will address these issues. The inclusion of secondary aerosols in MOCAGE, which is the most important deficiency, is already in progress. The African desert dust emission scheme with a better resolution and satellite-derived soil properties could bring better results over the region. Also, the addition of dust emissions in Australia, North and South America would fill the gap in the global dust emissions in the model.

As mentioned, aerosols have both direct and indirect effects on many atmospheric processes that have relevance to research themes in air quality and climate change. The current development is therefore a necessary stepping stone to being able to conduct studies on these important research topics. The mid-term aim, having added secondary aerosols, would be to carry out studies

of air quality studies and to determine the human exposure to aerosols. Another aim would be to calculate the aerosol radiative budget. Another possibility would be to improve the representation of aerosols by using data assimilation or data inversion in the cases where the source term is highly uncertain.

4.5.9 Appendix

This appendix defines the statistical metrics used in this paper. A more detailed review of these statistical terms is given by Huijnen and Eskes (2012), Seigneur et al. (2000) and Boylan and Russell (2006).

The bias is defined as the average difference between paired modelled predicted, p_i, and measured or reference, m_i, values:

$$\text{bias} = \frac{1}{N}\sum_{i=1}^{N}(p_i - m_i), \tag{4.22}$$

where N is the number of pairs (p_i, m_i). The bias is an estimation of the general overprediction or underprediction of the model with respect to the measurements.

The modified normalized mean bias, MNMB, is defined as

$$\text{MNMB} = \frac{2}{N}\sum_{i=1}^{N}\frac{p_i - m_i}{p_i + m_i}. \tag{4.23}$$

It is a measure of the model bias and ranges between -2 and 2.

The fractional gross error (FGE) is defined as

$$\text{FGE} = \frac{2}{N}\sum_{i=1}^{N}\left|\frac{p_i - m_i}{p_i + m_i}\right|. \tag{4.24}$$

It is a measure of model error and ranges between 0 and 2.

The MNMB and FGE weight equally overpredictions and underpredictions without overemphasizing outliers and do not consider measurements as the absolute truth. They are useful when prediction and measurement values are strictly positive.

The standard deviation, σ, indicated the spread from the average value and it is defined as

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - \overline{p})^2}, \tag{4.25}$$

where $\overline{p}$ is the mean of the predictions.

The correlation coefficient measures the extent to which patterns in the predictions match those in the measurements. It is defined as

$$\rho = \frac{\sum_{i=1}^{N}(p_i - \overline{p})(m_i - \overline{m})}{\sigma_{\mathrm{p}}\sigma_{\mathrm{m}}}, \tag{4.26}$$

where $\overline{m}$ is the mean of the measurements, and σ_{p} and σ_{m} are the standard deviations of the prediction and the measurements, respectively.

Code availability This paper is based on source code that is presently incorporated inside the MOCAGE model. The MOCAGE source code is the property of Météo-France and CERFACS, and it is based on libraries that belong to some other holders. The MOCAGE model is not open source and routines from MOCAGE cannot be freely distributed. Therefore, we cannot provide the code openly to the *GMD* website.

Acknowledgements This work is funded in France by Centre National de Recherches Météorologiques (CNRM-GAME) of Météo-France and Centre National de la Recherche Scientifique (CNRS). The authors would like to thank the AERONET PIs and their staff for establishing and maintaining the sites used in this investigation. We acknowledge the MODIS mission team and scientists for the production of the data used in this study. We also acknowledge NASA/Goddard Space Flight Center's Laboratory for Atmospheres for developing and computing GPCP combined precipitation data and NOAA/OAR/ESRL PSD, Boulder, USA, for providing it, AeroCom and Lamarque et al. (2010) for the emissions of carbonaceous aerosols and the Global Fire Emission Database project for the fire emissions that we used. We thank D. Carrer and his collaborators for developing and providing their SEVIRI-retrieved aerosol data. We also thank the two anonymous reviewers for their constructive comments.

Edited by: T. Butler

5

Assimilation of aerosols in CTM MOCAGE

This chapter describes the assimilation of the aerosol optical depth and lidar profiles in the CTM MOCAGE.

Contents

5.1 Motivation

The improvements of aerosol models lower uncertainties in the model results as seen in the previous chapter. Still, the model results can significantly differ compared to other models or observations. Remaining uncertainties can be further lowered by integrating the information obtained from observations by data assimilation. In order to improve the model performance, we developed a system to assimilate aerosol observations. Different types of aerosol observations are suitable to be assimilated. We focused our developments to data assimilation of the aerosol optical depth (AOD) and the lidar profiles, complementing the system by giving the information of the vertical aerosol profile. In this chapter we describe developments, tests and validation of our aerosol assimilation system.

The assimilation in the CTM MOCAGE, before the work on this thesis, was developed and used for atmospheric gases, predominately carbon monoxide (CO) and ozone (O_3) (e.g. Semane et al., 2007; El Amraoui et al., 2010; Claeyman et al., 2010; Rabier et al., 2010; Bencherif et al., 2011; Lahoz et al., 2012; Emili et al., 2014). Data assimilation in CTM MOCAGE is based on the assimilation algorithm called Valentina. The assimilation system MOCAGE+Valentina is guided by a dynamical coupler PALM (Projet d'Assimilation par Logiciel Multi-Méthodes) developed by CERFACS (Buis et al., 2006).

5.2 MOCAGE-PALM

The assimilation algorithm Valentina injects observations into the assimilation system, applies observation operators, makes an approximation of the matrices **B** and **R**, and performs minimization to find the analysis increments.

To couple the model and the assimilation algorithm, PALM arrange them separately into *branches*. Branches are composed of components (called *units*) which are arranged in a sequence. PALM guides the execution, scheduling and synchronization of the components and data exchanges (called *communications*) between them. The coupling is dynamical, i.e. the coupler is able to launch components either in parallel or successively, in loop or conditionally, all depending on the algorithm.

5.2.1 Assimilation algorithm

MOCAGE-PALM uses for assimilation various variational assimilation methods. Up to now, studies with MOCAGE-PALM used 3D-VAR, 3D-FGAT and 4D-VAR methods (Massart et al., 2010, 2012). The study of aerosol data assimilation presented in this thesis is realized using the 3D-FGAT method which is better performing compared to 3D-VAR. Also, besides an observational operator, 3D-FGAT does not need the model tangent linear and adjoint as in 4D-VAR. In the standard configuration of the MOCAGE, the model tangent linear and adjoint are developed only for the advection scheme, for the purpose of assimilation of gases in the case of strong dynamics in the stratosphere (Massart et al., 2012; Emili et al., 2014).

Preconditioning MOCAGE-PALM uses the incremental form of 3D-FGAT (Eq. (3.17)). In order to minimize the cost function more efficiently and to improve convergence, the increment δx is transformed into

$$v = \mathbf{B}^{-1/2}\delta x. \tag{5.1}$$

In this way the cost function becomes

$$\begin{aligned} J(x) =& \frac{1}{2}v^{\mathrm{T}}v + \\ & \frac{1}{2}\sum_{i=1}^{N}(d_i - \mathbf{H}_i\mathbf{B}^{1/2}v)^{\mathrm{T}}\mathbf{R}_i^{-1}(d_i - \mathbf{H}_i\mathbf{B}^{1/2}v), \end{aligned} \tag{5.2}$$

and its gradient

$$\nabla J(\delta x) = v + (\mathbf{B}^{1/2})^T \sum_{i=1}^{N} \mathbf{H}_i^T \mathbf{R}_i^{-1}(d_i - \mathbf{H}_i\mathbf{B}^{1/2}v). \tag{5.3}$$

In this formulation, there is no need for the explicit specification of the inverse matrix $\mathbf{B}^{-1}$. Other advantages of such approach are presented by Courtier et al. (1994). In their example, the minimization with the preconditioned form of the cost function gave the same results in 24 iterations, as the non-preconditioned form in 30 iterations. In our system, the cost function is minimized using the BFGS (Broyden-Fletcher-Goldfarb-Shanno) method.

The minimization of the cost function with the preconditioned form gives, as a result, an increment of the analysis in the space of variable v. After

the minimization, it is necessary to pass to the model space again:

$$\delta x = \mathbf{B}^{1/2} v. \tag{5.4}$$

More details on the assimilation algorithm are described by Pannekoucke and Massart (2008) and Massart et al. (2012).

Matrix B The background error covariance matrix $\mathbf{B}$ is a matrix of size $j \times j$, where j is the size of the model state vector. It can be represented as

$$\mathbf{B} = \mathbf{\Sigma C \Sigma}^T, \tag{5.5}$$

where $\mathbf{\Sigma}$ is the diagonal matrix of the square root of the variances, and $\mathbf{C}$ is the positive definite symmetric matrix of correlations. In the case of the preconditioned cost function, the matrix is formulated as

$$\mathbf{B}^{1/2} = \mathbf{\Sigma C}^{1/2}, \tag{5.6}$$

where $\mathbf{C}^{1/2}$ is the square root of the matrix $\mathbf{C}$. Usually we do not have enough information to explicitly estimate all correlation members, nor to store them in memory. For these reasons, the matrix $\mathbf{B}$ is modelled as an operator. To estimate the product of the matrix $\mathbf{B}$ and a vector, MOCAGE-PALM uses the technique of the integration of a generalized diffusion-type equation in a reduced space, proposed by Weaver and Courtier (2001).

The background error variances, which are located on the diagonal of $\mathbf{B}$ and which influence the weight of the model in the cost function, can be specified in MOCAGE-PALM either as the percentage of a control variable, or with the exact values in the units of the control variable. The background covariances, which influence the spread of the analysis to neighboring gridboxes, are specified either with a 3D correlation length field, or with constant correlation lengths in the horizontal and vertical. The constant correlation lengths are modelled using a Gaussian function (Pannekoucke and Massart, 2008) in terms of geographic degrees for the horizontal lengths, and in terms of pressure or number of model levels for the vertical lengths (Massart et al., 2009). The characteristics of the correlation lengths influence the method by which the generalized diffusion-type equation will be solved. In the case of constant correlation lengths in the global domain, the equation is solved by the spherical harmonics transforms, in the case of a limited-area domain by the finite-difference method.

Matrix R Errors of observations are considered to be non-correlated, which means that all non-diagonal members (covariances) in the matrix $\mathbf{R}$ are zero.

The matrix $\mathbf{R}$ is reduced to its diagonal with variances of measurements

$$\mathbf{R} = \mathbf{D}_y = \mathrm{diag}(\sigma^2_{obs}). \tag{5.7}$$

Variances can be specified either as the percentage of the measurements or as the exact values in the unit of the observations.

Slots The cost function is minimized every assimilation cycle (window). For a better model–observations comparison and memory optimisation, the cycle is divided into *time slots*. During each slot, observations are read, the observation operator is run, its output field is interpolated to locations and times of the observations and compared with the observations, and the misfit vector is calculated and stored.

5.2.2 The choice of control variable for aerosol assimilation

For aerosols, the modelled prognostic variable and observations are usually not the same physical quantity. In MOCAGE, the prognostic variable is the aerosol mass concentration of each bin, and the quantities that we want to assimilate are the aerosol optical depth and the lidar backscatter/extinction profiles. For assimilation, it is necessary to choose the control variable x (Eqs. (5.2) and (5.3)) in the way to be the best adapted to our system and its purpose. One of the first criteria is that the control variable should be the same for all types of observations intended to assimilate. The observation operator should be as simple as possible and easy to linearise.

In the literature we can find different choices for the control variable for aerosol assimilation:

- Liu et al. (2011) used directly the prognostic variable in their model as the control variable. In our case that would be the 3D concentration of 30 bins.

- Benedetti et al. (2009) used the 3D total concentration as the control variable. All bins have to be summed, in order to obtain a new hybrid summed aerosol species.

- Zhang et al. (2008) and Schroedter-Homscheidt et al. (2010) assimilated aerosol optical depth by using AOD directly as the control variable.

The different approaches have different impacts on the assimilation process. The first approach is rigorous and straightforward, where all unknowns are the control variable. This makes the control variables considerably bigger than in other approaches which can have a performance effect, in terms of memory optimisation and calculation time. Also, the matrix **B** has to include all bins separately with their variances and covariances which are difficult to define. But, the analysis is partitioned to all bins automatically by the system, which searches for the optimal solution for all unknowns in the minimization of the cost function.

The second approach makes the control variable smaller than in the first approach. The cost function searches a solution in terms of only one variable. The observational operator has first to sum all bins. In order to linearise it, its tangent linear and adjoint have to contain information on how to partition the increment of analysis to all modelled bins. In this way, we directly decide about the method of the partitioning, by hard-coding the repartition criteria in the model code.

The third approach is optimized for the assimilation of the AOD, where the observational operator would only do interpolation. The increment of analysis is also in quantities of AOD and after the minimization it is necessary to convert it to the 3D concentrations partitioned into the bins. The observation operator for the types of observations other than AOD would be more complicated, especially for observations which are not column-integrated. Thus, the control variable would have to be changed from total column AOD, to the profile of partial-column AOD in order to preserve the information in the vertical axis and avoid 2D/3D transformations and unnecessary hypothesis. Thus, this approach is not optimal for the assimilation of the observations that are not column-integrated.

Considering our needs and the characteristics of our system, we have chosen to use the 3D total aerosol concentration as the control variable as in Benedetti et al. (2009). With this choice, the problem of minimization of the cost function is better determined than in the first approach, where one observation would be used to constrain 30 unknowns (bins). Also, it is lighter in terms of memory usage and calculation performance. Still, in order to linearise the observation operator, it is necessary to make an assumption how to the analysis increment δx^a will influence each bin.

Linearisation choice In MOCAGE-PALM, we decided to keep the relative contribution of each bin constant in terms of their mass during the assimilation cycle. Bulk aerosol observations do not have any information of the contribution of different aerosol types. With this approach, the repartition of the analysis increment strictly depends on the background.

5.3 AOD assimilation

In MOCAGE-PALM, we developed the capacity of the system to assimilate the aerosol optical depth. In this section we describe the development of the observation operator, its implementation in the system, and the validation using linear tangent and adjoint tests, synthetic observations and different real observations.

5.3.1 Developments

The assimilation of aerosol optical depth in MOCAGE-PALM requires the development of an observation operator which would transform the control variable from the model space, which is in terms of total concentration, into the observation space in terms of aerosol optical depth. First, in the model it is necessary to sum all individual species in order to calculate the total concentration, and to calculate the relative mass contributions among bins which will stay constant during an assimilation cycle. The aerosol optical depth (τ) is calculated by taking into account concentrations (n_{bin}) and the optical properties of individual species (C_{ext}) calculated by the Mie code:

$$H = \sum_{bin} \sum_{lev} C_{ext}(D_p, \tilde{n}, \lambda) n_{bin} \Delta z_{lev} = \tau. \tag{5.8}$$

The observation operator, together with its tangent linear and adjoint operators, are developed to be able to take into account different wavelengths. Wavelengths which the observation operator recognize are those that are commonly used for AOD retrieval and lidar observations. The necessary particle optical properties are computed at the beginning of the simulation in a look-up table, and applied when needed by the observation operator.

Within an assimilation cycle, firstly, observations are loaded. Only observations corresponding to the time period of the current slot are selected, and the characteristics of the different instruments are collected. The observation operator is called for all different wavelengths that are found. These AOD fields are then interpolated at times and locations of observations and misfits are calculated. Then, the tangent linear of the observation operator is called, initiated with perturbations of the total concentration, to calculate the perturbation in AOD. In the first iteration, the perturbation of the total concentration is zero. Afterwards, the adjoint of the observation operator is

applied to obtain its back projection, which is used to calculate the gradient of the cost function. Due to the linearisation of the observation operator, in the adjoint operator: first is calculated the back projection with respect to individual bins, and second, the back projection with respect to the total concentration is obtained, by applying the relative contributions of bins which are kept constant. The calculated gradient, which is in the units of the variable control, is used as the perturbation for the next iteration. The minimization goes until the convergence criterion is met. The increment of analysis, which is in terms of total concentration, is repartitioned to all model bins using the relative mass contributions of each bin.

Tangent-linear and adjoint tests The developed tangent linear (TL) and adjoint operator were tested. The tangent linear operator in order to provide a linearised trajectory around the model state x has to satisfy the tangent linear test:

$$\lim_{\delta x \to 0} \frac{H(x+\delta x) - H(x)}{\mathbf{H}(\delta x)} = 1. \tag{5.9}$$

As long as the perturbation δx is small enough that it stays close to the model state x, the test has to give a value close to 1.

In our AOD assimilation system, we perturb only aerosol concentrations, and not any other variables. In TL test, we define the perturbation δx. In this way, the relative contributions of bins in the perturbations are known. Since there is a linear dependence between AOD and aerosol concentrations, our tangent linear trajectory strictly follows the model trajectory, even for the perturbed states not near the model state x. This means that the test will always return a value of 1, independent of the size of perturbation. We confirmed this by running the test varying the perturbation field δx.

During the assimilation, the relative contributions of bins in the analysis are not known – they are considered identical as in the background. Thus, the tangent linear test, in the best case, can only be close to the value of 1 for a small perturbation that is close to the model state.

The adjoint operator of the observation operator is a backward operator and has strictly to reverse the action of the tangent linear. We test it with a *dot product test* described by Eq. (3.22):

$$\langle \mathbf{H}x, y \rangle = \left\langle x, \mathbf{H}^{\mathrm{T}} y \right\rangle .$$

Independently of the size of perturbation, the test has to match up to the machine precision of the computer. We achieved this with our AOD adjoint operator.

Implementation The aerosol optical depth observation operator was the first observation operator in MOCAGE-PALM for which the controlled and observed variables were different physical quantities. In past experiences, the observation operators were doing only simple transformations, like integration over levels for total column observations, or multiplication in the case of profiles with averaging kernels, etc. For the AOD assimilation, it was necessary to adjust the assimilation flow in PALM and add necessary communications. This task was done in collaboration with the PALM team of CERFACS.

Some model fields are necessary to the tangent linear and the adjoint operator during the minimization of the cost function (performed in the PALM branch *observations*). The necessary variables include: total 3D concentration of all aerosol types as the control variable; constant relative mass contributions of bins (a link between the control variable and the model prognostic variables); and some meteorological fields like pressure, specific humidity, temperature, etc. They are not freely available in this PALM branch, and it is necessary to transfer them via PALM communications from the PALM branch *model*. The minimal update frequency of the variables in MOCAGE-PALM is one hour since the minimal *time slot* length is one hour. In the case when the assimilation cycle is composed of more time slots, because of the limited memory, we can transfer only the control variable at each time slot, but not other necessary variables. We have to approximate them by considering them to be constant during the whole cycle. We can make them available to this PALM branch differently: taking their values at the beginning of the cycle, or at the end of the cycle, or averaging them. We made sensitivity tests (explained in more details in Appendix B) to determine which choice is the best, and we concluded that the differences in the results using any of these choices were minor. A much more important effect on the results had our linearisation choice (to make the relative mass contributions of the bins constant) than the choice how to treat the necessary variables before passing them to the branch *observations*.

5.3.2 First validation with synthetic observations

First tests of the implemented AOD assimilation system were performed by using synthetic observations. The synthetic observations are defined to be at the center of a gridbox, and observed at the beginning of an assimilation cycle. In this way, the observation operator does not perform any type of interpolation.

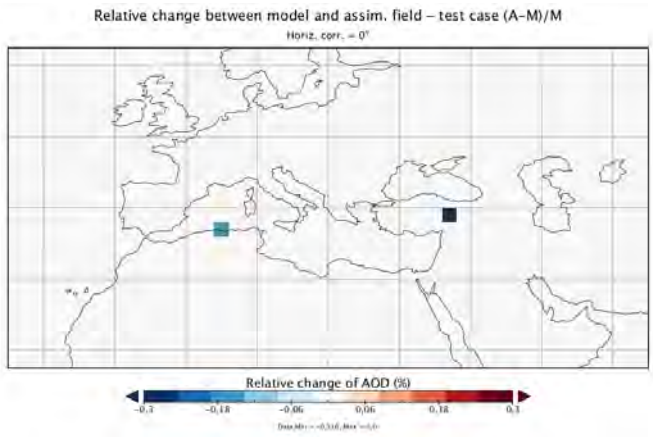

Figure 5.1: Assimilation of two synthetic observations with equal matrices **B** and **R**. Here is shown the relative change of analysis compared to the background. The values of backgrounds, observations and analyses are presented in Table 5.1.

Tests with the equal matrices B and R In this test both matrices **B** and **R** are defined as diagonal, with correlation members equal to zero. Their diagonals are identical, with equal values of errors for each observation and for the corresponding model gridbox. To have the matrices of the same size, we do not consider the whole aerosol profile, but only one level. In this case, we avoid the vertical integration in the AOD calculation. If the vertical integration was present in the observation operator, this test would not be valid. As the final result in this test, the analysis should not spread onto the neighboring gridboxes (no interpolation nor correlation), and its value should be exactly at the half between the values of AOD which are calculated in the model and which are observed. This test checks the assimilation system in total because it involves the functioning and calculation of the covariance matrices, observation operator, the cost function and its gradient, minimiser, etc.

Figure 5.1 shows the position of two synthetic observations used in our test, and Table 5.1 shows the results with the AOD values of the background,

Table 5.1: AOD values of the background, observations, and analyses for the assimilation of two synthetic observations in the test corresponding to Fig. 5.1 where matrices **B** and **R** are identical.

	background	observation	analysis
Obs. 1	0.864	0.50	0.682
Obs. 2	2.006	1.50	1.753

synthetic observations and analyses for those two points. The system successfully passed the test since the analyses are exactly between the values of the background and the observations.

Figure 5.2 shows an example of analysis for the same synthetic observations when covariance members of the matrix **B** are not zero. The horizontal correlation is modelled by the Gaussian function in terms of the geographic coordinates expressed in latitude and longitude. The horizontal correlation $\rho_h(\delta)$ between two points separated by the distance δ is

$$\rho_h(\delta) = \exp\left(-\frac{\delta^2}{2(L_x^2 + L_y^2)}\right), \tag{5.10}$$

where L_x and L_y are correlation length scales in S-N and W-E direction in [km]. In Fig. 5.2 both length-scales, L_x and L_y, are set to 2°. In this case, analyses spread symmetrically onto the neighboring gridboxes as expected.

Influence onto the vertical The aerosol optical depth is an integrated quantity and does not carry any information on the vertical profile. Bearing that in mind, the analysis is expected to have the same vertical profile shape as the background: the vertical profile of the analysis should be just the background profile multiplied by a factor. Still, if aerosol species are not uniformly distributed vertically, we should expect some changes in the vertical profile. The reason is that the back projection calculated by the adjoint carries the information of different optical properties of different bins and aerosol types. Consequently, different bins have a different specific weight to the adjoint back

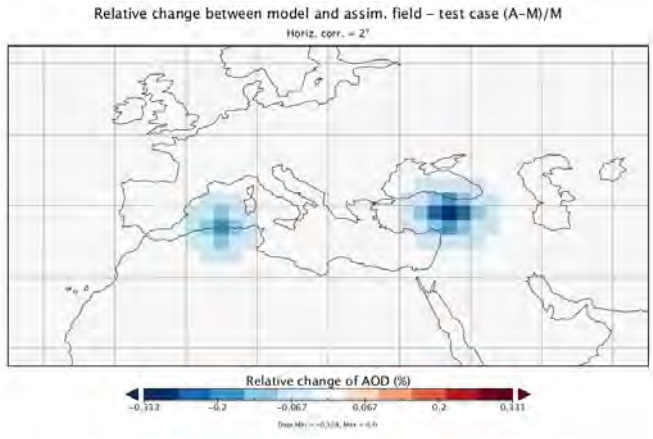

Figure 5.2: Assimilation of the same field as in Fig. 5.1, except the horizontal correlation is set to 2°.

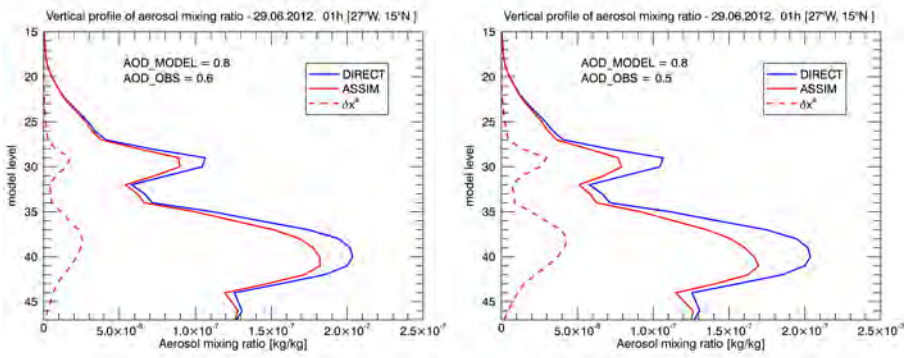

Figure 5.3: Aerosol mass mixing ratio profiles for the direct model and assimilated synthetic observations with different values of AOD: **(left)** AOD misfit of 0.2, and **(right)** AOD misfit of 0.3. Matrices **B** and **R** are constructed so that the observations have two times smaller errors, and that covariance members are zero. In the model, the ground corresponds to Level 47.

projection, and this can have an effect on the analysis if the aerosols types are not vertically uniformly distributed. This is demonstrated in Fig. 5.3 where in the two assimilation runs one synthetic observation is assimilated by varying its AOD value. The direct model profile over a point in the Atlantic ocean is characterized by two layers of dust (centered around model Level 31 and Level 40) and mixed aerosols located in the first few levels. By changing the AOD value of the synthetic observations, it is clear that in the analysis the particles in the centers of dust plumes are relatively slightly more influenced than the aerosols at the surface or in other levels. Still, the described effect is small and in the first approximation we can consider that the profile shape does not change during a single assimilation cycle.

5.3.3 Validation during TRAQA and ChArMEx

After the validation of AOD assimilation using synthetic observations, we make an extensive evaluation of the AOD assimilation system using real observations. Aerosol data assimilation shows its full capabilities if the available observations are assimilated continuously for an extended period of time. Here, we make such an experiment by assimilating MODIS AOD data for the time periods of both the TRAQA and ChArMEx field campaigns, in the summer of 2012 and 2013, respectively. Since in the model we do not have secondary aerosols, we choose these campaigns that are carried out in the summer, when primary

aerosols dominated the region. Different desert dust and biomass burning events, together with sea sat aerosols in the basin marked these periods.

After assimilation, a comparison with independent observations is used to determine the exact usefulness of the MODIS data assimilation. We compare our assimilated aerosol field against satellite measurements from SEVIRI, and ground-based measurements from AERONET. The more detailed evaluation of the AOD assimilation capabilities is possible by the comparison with aerosol measurements other than AOD. For this, we use data from the TRAQA and ChArMEx intensive measurement campaigns.

5.3.3.1 Assimilated observations

MODIS The MODIS (Moderate-resolution Imaging Spectroradiometer) instruments observe atmospheric aerosols aboard Terra (since 2000) and Aqua (since 2002) from complementary sun-synchronous orbits. The Terra overpass time is around 10:30 local solar time at the equator in its descending mode, and the Aqua overpass time is around 13:30 local solar time at the equator in ascending (daytime) mode. We use MODIS Aerosol Optical Depth Collection 5 retrievals at 550 nm from Terra and Aqua that have predicted uncertainties of $\Delta\tau = \pm(0.03 + 0.05\tau)$ over oceans and $\Delta\tau = \pm(0.05 + 0.15\tau)$ over land (Remer et al., 2005). Over bright desert areas, we use the "Deep blue" MODIS product (Hsu et al., 2006). For the assimilation, we only considered the best quality data, with the highest possible quality flag in used products: the ocean product retrieved with the "best solution", the reflectance corrected land product and the Deep Blue product over bright arid surfaces.

MODIS L2 resolution of 10×10km is superior than the model resolution of $0.2° \times 0.2°$ over the control domain in which the assimilation is performed. Also, the large number of observations over the control domain during one slot can be a limiting factor. For that, we use the technique of upscaling of observations (making so-called *super-observations*, Daley, 1993) in which we average all observations measured over the same model gridbox within the time period of 5 min. In this way, the model and upscaled observations have the same resolution, and the location of the observations corresponds to the gridbox centers. The maximal number of observations per one hour slot is reduced in this way to $\approx$15 000, which is manageable with our system. To combine Terra and Aqua observations, we weight data by considering the number of level 2 (L2) observations in each super-observation gridbox.

5.3.3.2 Independent observations

SEVIRI SEVIRI geostationary observations over ocean are retrieved at 550 nm by Thieuleux et al. (2005). ICARE operationally implemented this algorithm and makes AOD data available on its website (www.icare.univ-lille1.fr). This product was evaluated against other satellite products and AERONET measurements by Thieuleux et al. (2005) and Breon et al. (2011). The instrument makes an image of the whole Earth disk every 15 minutes. With such a high temporal resolution we expect to have a big amount of available SEVIRI data. Therefore, we use only the SEVIRI product over waters because the retrieval over dark surfaces is usually more accurate. The Earth disk in the SEVIRI field is divided into 3712×3712 pixels. A nadir horizontal resolution is 3 km, while over Europe it is $\approx$5 km. In our study, the SEVIRI data – used for the comparison with model fields before and after assimilation – was regridded to the model resolution.

AERONET AERONET (Aerosol Robotics Network) measures ground-based AOD from hundreds of automated stations in the world with an accuracy of ± 0.01 (Holben et al., 1998). As high accuracy observations, AERONET observations are often used for the validation of satellite products (including MODIS and SEVIRI; Remer et al. (2005); Thieuleux et al. (2005)). We use all available L2 data from different stations and interpolate it in logarithmic space to 550 nm (to harmonize wavelengths between different stations and with the model) by using available neighboring wavelengths: 440 nm, 500 nm, 675 nm, 870 nm.

TRAQA observations TRAQA (*fr.* TRAnsport à longue distance et Qualité de l'Air dans le bassin méditerranéen) was a scientific project including a measurement campaign intended as a pre-ChArMEx (CHemistry and AeRosol MEditerranean EXperiment) experiment (http://charmex.lsce.ipsl.fr). It took place over the north-western Mediterranean basin during the summer 2012. The main scientific objectives were: the characterization of the dynamic processes of export of polluted air masses from remote continental regions around the Mediterranean Sea; the quantification of exchanges between the boundary layer and the free troposphere above the Mediterranean basin; the study of ageing and mixing of pollution plumes in the lower troposphere (gases and aerosols) with Lagrangian tracking; the analysis of representative case studies against long-term measurements to estimate the impact of plumes of pollution on air quality. From 26 June to 11 July seven intensive observation periods were realized

with the ATR-42 aircraft of Météo-France, atmospheric balloons (sounding and drifting) and ground instruments measuring trace gases and aerosols. During the campaign, an event of desert dust outbreak from Africa arrived to the basin. It was well observed around 29 June with several instruments.

PCASP In our study, we use the data measured by the passive cavity aerosol spectrometer probe (PCASP) which was on board of the ATR-42 aircraft. The PCASP measures the aerosol concentration and the aerosol size distribution with its 30 channels (Strapp et al., 1992). Particles size range that PCASP can measure cover particles from 0.1 µm to 3 µm, with channel ranges and calibration methods reported by Cai et al. (2013). The high temporal resolution of the instrument is lowered for a comparison with the model from 1 s to 1 min. With its variable crusing speed, the horizontal resolution with this data thinning can be estimated on average to about 5 km.

LOAC Also, we use data from LOAC instruments (Light Optical Particle Counter) collected during TRAQA. LOAC is a light aerosol optical counter measuring aerosols in 19 size classes within a diameter range from 2 µm to 100 µm in the version of the instrument used in the TRAQA campaign. LOAC uses the technique of measuring aerosol at different scattering angles. When using a small scattering angle (below 20°), measurements are not very sensitive to the particle nature (Lurton et al., 2014). When using a larger scattering angle, the scattering depends on the nature of aerosols. When LOAC is mounted on the meteorological sounding balloon as during TRAQA, its vertical resolution depends on its measurement frequency. For the processed data that we use in our analysis, the vertical resolution in the troposphere is about 0.3 km to 0.4 km which is similar to the model resolution.

ChArMEx observations ChArMEx is a part of MISTRALS (Mediterranean Integrated STudies at Regional And Local Scales) programme, and is intended as a mid-term project. Its scientific objectives are: assessing the present state of the Mediterranean atmospheric environment; quantifying the impacts of aerosols and reactive gases; predicting future evolution of the budgets of aerosol and gases and its impacts. To achieve these goals, it is intended to carry out several field campaigns. One of the extensive ChArMEx field campaigns took place in summer 2013 in the western Mediterranean basin, with the first intensive observation period from 13 June to 5 July. During this period, different instruments were deployed on aircrafts, balloons and the ground. Similarly to the TRAQA data, we will use the particle concentration PCASP measurements, and LOAC data from meteorological sounding balloons. The

diameter range that LOAC measured in the ChArMEx campaign was 0.2 µm to 50 µm (Renard et al., 2013).

5.3.3.3 Assimilation experiment setup

To run the assimilation system it is necessary to specify the control domain and the error covariance matrices. For this study, the control domain, where we assimilate MODIS data, is defined with a resolution of $0.2° \times 0.2°$. It spatially covers the Mediterranean basin and the Sahara desert. The precise boundaries are $[20°W - 40°E, 16°N - 52°N]$ (Fig. 5.4). The lateral boundary conditions are provided by the global domain which is run with the resolution of $2° \times 2°$.

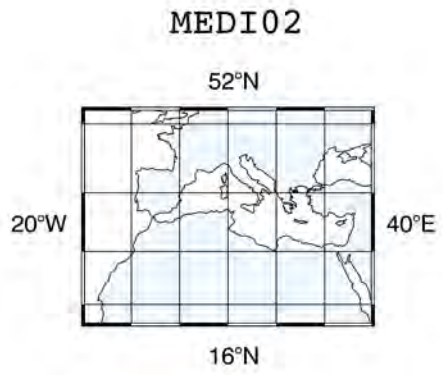

Figure 5.4: Control domain used in the assimilation study.

Error covariance matrices The good quality MODIS data still can contain the cloud contamination and outliers (Zhang et al., 2005; Sič et al., 2015), which contribute to uncertainties of observations. Different MOCAGE uncertainties are discussed in Section 2.2.2 and Chapter 4. Because of the errors, uncertainties and biases of datasets, the model and observations covariance matrices must be carefully characterized for successful assimilation. In this study, the error variances in matrices **B** and **R** are defined as the percentage of the modelled control variable for MOCAGE and the percentage of measured AOD for the observations. We used the χ^2 diagnostics to try to estimate optimal values for errors of these datasets (Ménard et al., 2000; Talagrand, 2003). The χ^2 test is a posteriori diagnostic which defines properly specified errors if

$$E\left(\frac{2J_{min}}{p}\right) \sim 1, \tag{5.11}$$

where E is the expectation (statistical average), J_{min} is the value of the cost function at the minimum and p is the number of observations. For this test, it is necessary to run the assimilation system for a prolonged period of time and, in the case of sufficient number of observations, the matrix $\mathbf{B}$ will not depend any more on its initial value. Because it is quite computationally expensive, a very rigorous optimization of the errors is difficult to do. We carried out several test runs to determine optimal parameters for the matrices a posteriori. We experienced significant variations in the χ^2 diagnostics. Roughly, we estimated as the optimal parameters the percentage for the errors for the model (24%) twice larger than for the observations (12%).

We also considered a possible underestimation of AOD in the model, because of missing secondary aerosols in MOCAGE. Thus, the smaller percentage for the error of observations could be compensated by possible smaller AOD values of the model. And, it is also considered that in the case of doubt, it is better to define a $\mathbf{B}$ matrix which is overestimated, than vice-versa (Talagrand, 2003).

The assimilation cycles in the experiment have a length of one hour. The cost function is minimized until the convergence criterion is reached, with the maximum number of iterations equal to 200. Horizontal and vertical correlations are homogeneous, with a 0.4° horizontal correlation, and a vertical correlation of one model level (it has no strong effect, but it reinforces the smoothness of the analysis increment). The analysis increment is added to the beginning of each assimilation cycle.

5.3.3.4 Model emissions

Emissions of aerosols in this study are the same as in Section 4.5.3.4, except for biomass burning aerosols. Biomass burning aerosol emissions are introduced from the Global Fire Assimilation System (GFAS) version 1.1 (Kaiser et al., 2012). The GFAS assimilates the fire radiative power observed by MODIS, corrects the cloud cover gaps, filters anthropogenic and volcanic activities, and finally calculates biomass burning aerosol emissions. The temporal resolution of the dataset which we use is one day, and the spatial 0.1°.

5.3.3.5 Experiment

Our goal in this experiment is to assimilate the MODIS AOD observations in MOCAGE during the TRAQA and ChArMEx field campaigns, and to evaluate

the impact of such assimilation on the aerosol field in MOCAGE. For that, we run two MOCAGE configurations, one with and one without assimilation. The simulation without assimilation we refer as the *direct model run* and the simulation with assimilation as the *assimilation model run*. The model fields, both direct and assimilated, we compare with different observations: independent AOD satellite and ground-based measurements, and the campaign data of the aerosol concentration and size distribution. All used independent datasets in this study are not considered as a "reality", because they have there own uncertainties and errors, but they represent independent platforms for comparison with different measurement methods and characteristics.

The concerned periods are from 25.06.2012 until 13.07.2012 for TRAQA, and from 12.06.2013 until 05.07.2013 for ChArMEx. For both periods, the model is run with spin-up periods of at least 45 days. In the assimilation runs the spin-up is split in the direct model spin-up period of one month, and the assimilation model spin-up period of ten days. In this way, at the beginning of the studied periods the assimilation model run already achieved important AOD correction over a major part of the field.

5.3.3.6 TRAQA results

The assimilation has an important impact on the modelled field. This is illustrated in Fig. 5.5 where we compare AOD fields of the direct model run and the assimilation model run with the independent SEVIRI AOD field. During the campaign, over the western Mediterranean basin were elevated concentrations of desert dust aerosols coming from the Sahara desert. This dust event occurred during 4 to 5 days over the region, giving the possibility to study a strong dust event with multiple instruments. As seen from the SEVIRI observations, the event was widespread over the whole western Mediterranean, bringing AOD values above 0.5 over the large area. The direct model run underestimated the intensity of the event, although simulated correctly its spatial extent. The assimilation strongly influenced the modelled field and corrected the AOD values in the model. By comparing with independent SEVIRI observations, the assimilated field match well the spatial extent, with AOD values being similar to what is observed by the satellite.

OmF and OmA analysis We evaluated in Fig. 5.6 the impact of the data assimilation to the modelled fields (forecast and analysis) compared with the assimilated observations. The histograms of OmF (Observation minus Forecast) and OmA (Observation minus Analysis) present the difference between the

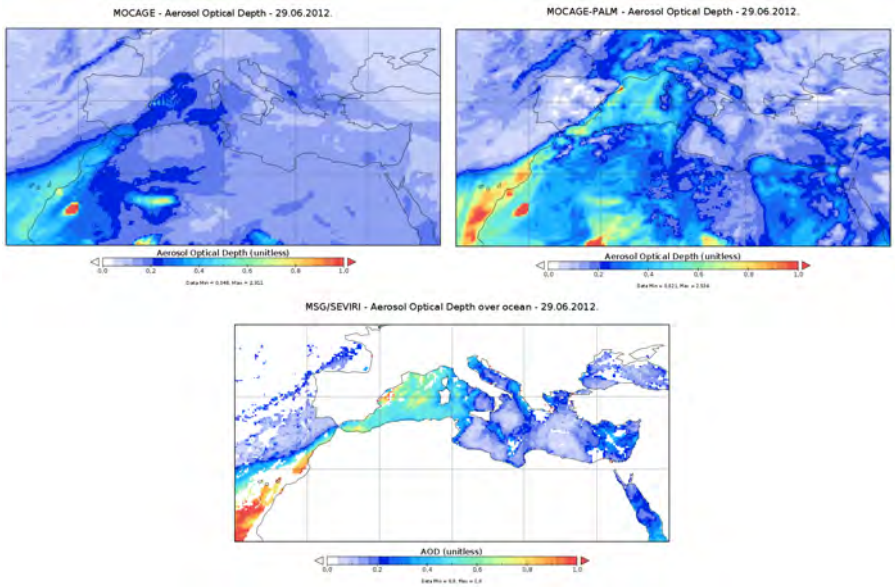

Figure 5.5: The aerosol optical depth over Europe on 29.06.2012 at 12h UT, (**top left**) simulated in MOCAGE by the model direct run, **(top right)** simulated in MOCAGE by the MODIS assimilation model run, and **(bottom)** observed by SEVIRI as the independent dataset.

forecast and analysis with MODIS observations. The forecast is considered at the very beginning of each assimilation cycle and takes into account effects

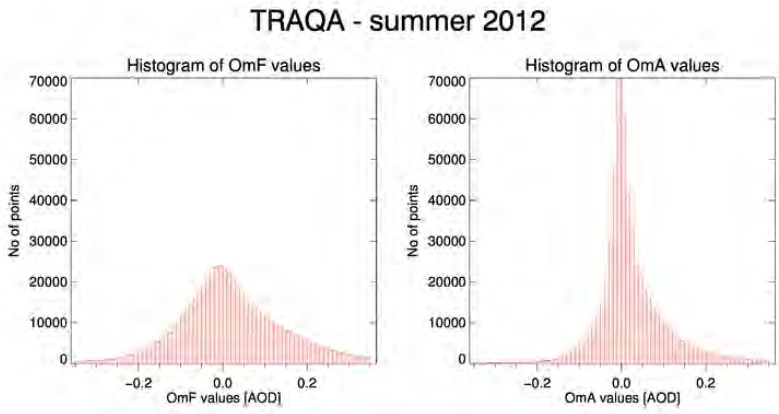

Figure 5.6: Histograms of (**left**) OmF (Observation minus Forecast), and (**rigth**) OmA (Observation minus Analysis) for the period of the TRAQA campaign from 25.06.2012 until 13.07.2012.

of observations already assimilated, while the analysis is considered after the assimilation process in the cycle. The forecast in our case is a one hour forecast. The OmF differences are well centered and their spread shows the degree of underestimation and overestimation of the forecast. This degree is considerably smaller in the analysis showed by the narrower OmA histogram compared to OmF. This reduced difference between assimilated observations and the model, considered as the main goal of the data assimilation, confirms its positive impact on the model performance.

Figure 5.7 shows the distribution and gives the statistics of the forecast and analysis with the assimilated MODIS observations. It clearly shows the benefits of assimilation. The assimilation improves the correlation, root mean square error and standard deviation of the model compared to the observations, which confirms the efficiency of AOD assimilation to make model closer to observations.

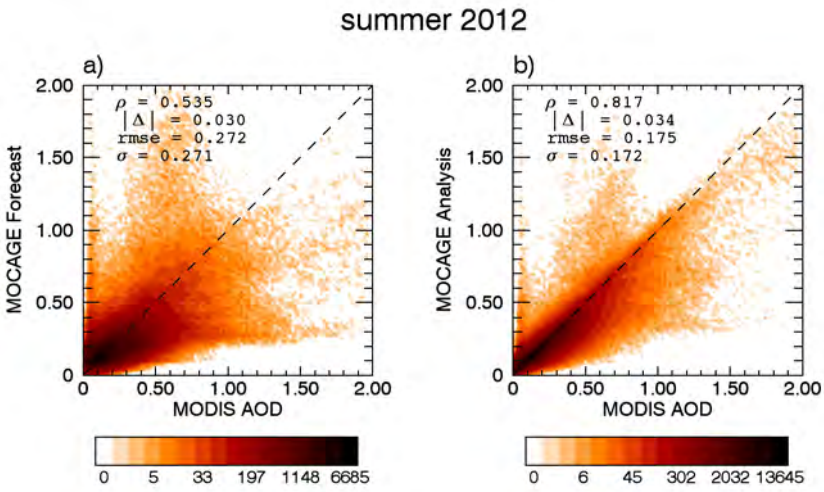

Figure 5.7: Scatterplots of aerosol optical depths of assimilated MODIS observations and: **(a)** the forecast (of one hour), and **(b)** analysis. In each panel, correlation (ρ), absolute bias (Δ), root mean square error (RMSE) and standard deviation (σ) are noted. The assimilated data correspond to the period of the TRAQA campaign from 25.06.2012 until 13.07.2012., and covers the MEDI02 domain (Fig. 5.4).

Comparison with SEVIRI A similar conclusion can be underlined when comparing the direct model and the assimilation model field with independent observations. In Fig. 5.8, we use available SEVIRI observations (over water surfaces). It should keep in mind that, as MODIS and SEVIRI are different

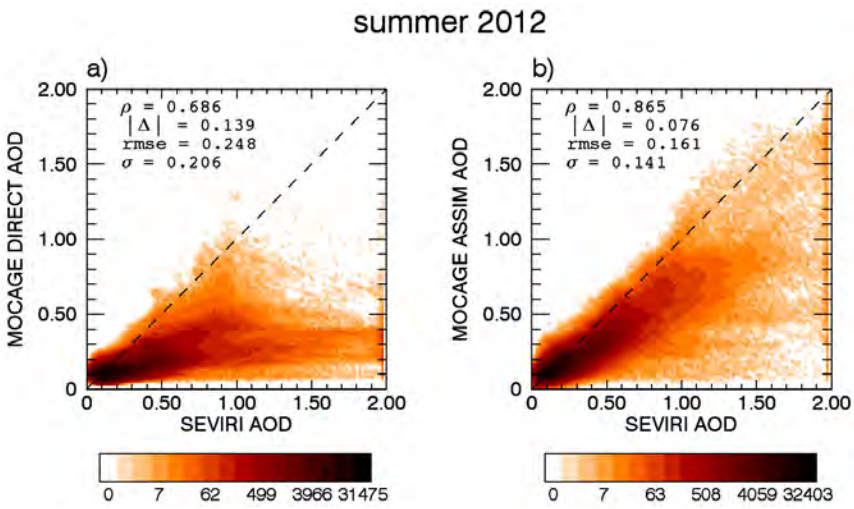

Figure 5.8: Scatterplots of aerosol optical depths from the independent observation dataset (SEVIRI) and the simulations: the direct model run **(a)** and the assimilation model run **(b)**. In each panel, correlation (ρ), absolute bias (Δ), root mean square error (RMSE) and standard deviation (σ) are noted. The included data correspond to the period of the TRAQA campaign from 25.06.2012 until 13.07.2012., and covers the whole MEDI02 domain (Fig. 5.4).

products, they show the differences for the same observations (which make them independent). The majority of observed points correspond to weak AOD values where the direct model and observations agree well. For larger observed values, the direct model run underestimates the AOD. These points mostly correspond to the strong desert dust outbreak, whose intensity the direct model underestimates. The assimilation reduced this disagreement, with the analyses situated between the direct model and the observations.

In Fig. 5.9 and Table 5.2 of SEVIRI timeseries, we consider only the western Mediterranean basin where the TRAQA campaign took place. We

Table 5.2: Correlation (ρ), absolute bias (Δ) and root mean square error (RMSE) between SEVIRI observations and MOCAGE direct/assimilation model run for the western Mediterranean during the TRAQA campaign between 25.06.2012 and 13.07.2012. Mean number of SEVIRI observations per hour is also given. The data correspond to Fig. 5.9 with the observations localized in the region 0–10° E, 35–45° N.

		MOCAGE direct			MOCAGE assimilation		
	$\overline{N_{obs}}$ [h^{-1}]	ρ	Δ	RMSE	ρ	Δ	RMSE
SEVIRI - TRAQA	20875	0.83	0.14	0.17	0.96	0.08	0.09

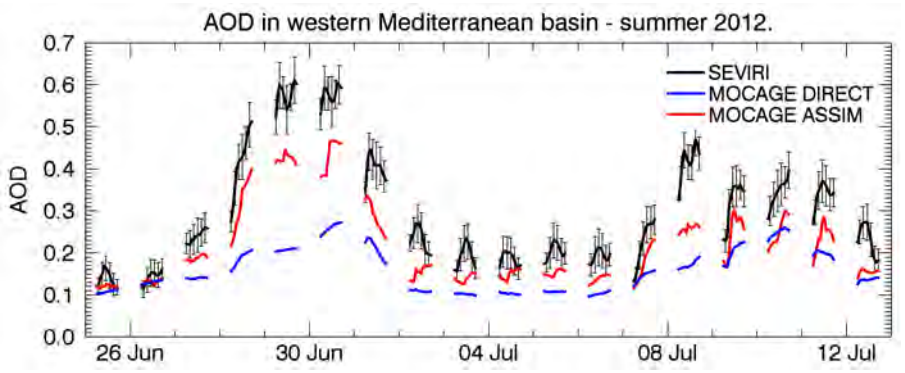

Figure 5.9: Hourly timeseries of aerosol optical depth at 550 nm of SEVIRI data, the direct model and the assimilation model run over the western Mediterranean (0–10° E, 35–45° N) for the period of the TRAQA campaign from 25.06.2012 until 13.07.2012. The considered region is also marked in Fig. 5.10 by the grey box. Correlation, bias and root mean square error for both the direct model and the assimilation model run as compared to SEVIRI data are given in Table 5.2.

confirm that the TRAQA period over the western Mediterranean was marked by a strong desert dust outbreak coming from the Sahara desert. It elevated AOD values during 4 to 5 days over the region. This figure also confirms that the intensity of this dust event is underestimated in the direct model run. The second desert dust event occurred at the end of the TRAQA period, but it was weaker and it was localised only in the part of the western Mediterranean close to the coast of Africa. The AOD values of the second dust event, and also of the period between the two events, were underestimated in the direct model run, although not as strongly as during the first event. Data assimilation reduced the difference between the model and the observations, improving all considered statistical parameters (Table 5.2). The MODIS overpasses the region twice during the daytime, and this provided sufficient information to pick up well the AOD time variations in the assimilated field (Fig. 5.9).

Comparison with AERONET We compared the model direct run and assimilation model run with the AOD data from AERONET stations. All considered stations are in or around the Mediterranean basin, and their locations are presented in Fig. 5.10. Timeseries plots for ten stations are presented in Fig. 5.12, and the statistics for all stations in Table 5.3. In total, for the TRAQA period, we considered measurements from 35 AERONET stations. The timeseries of the stations in the western part of the Mediterranean basin and in Spain are marked by the strong desert dust event, which was already discussed earlier. Stations in Spain registered the event before the stations in France, where it arrived a couple of days later. Nevertheless, the event covered the whole region

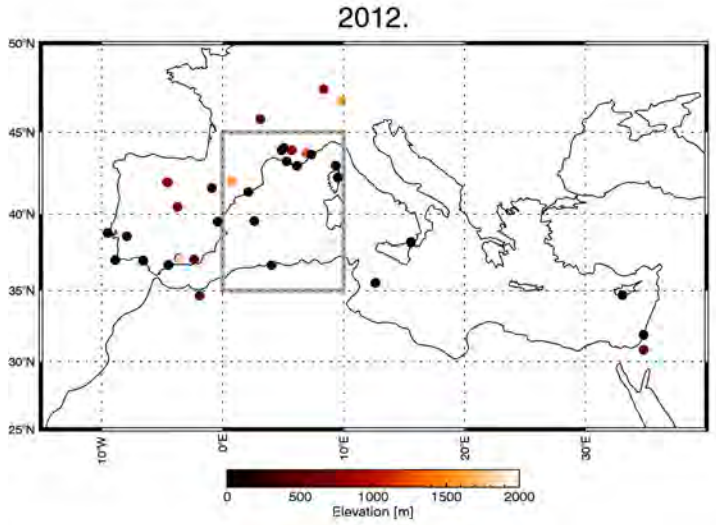

Figure 5.10: Positions of AERONET stations used in this study for the period of the TRAQA campaign in summer 2012. The grey box mark the region from which we considered SEVIRI data used in Fig. 5.9, Table 5.2.

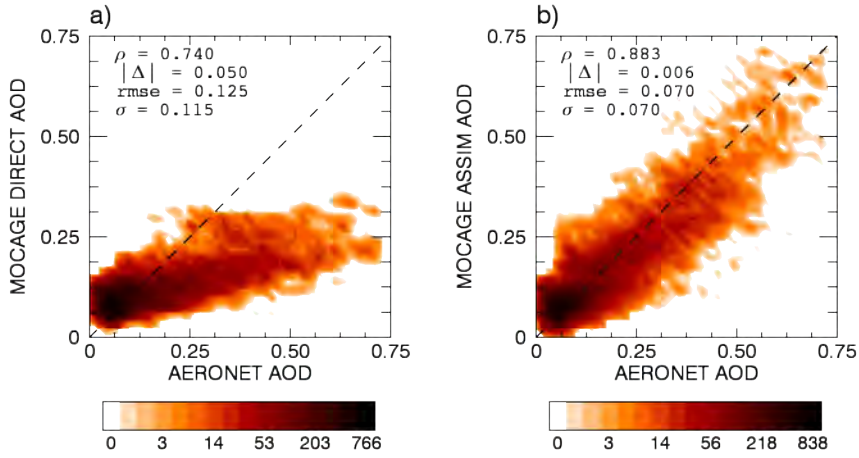

Figure 5.11: Scatterplots of aerosol optical depths from the independent observation dataset (AERONET) and the simulations: the direct model run **(a)** and the assimilation model run **(b)**. In each panel, correlation (ρ), absolute bias (Δ), root mean square error (RMSE) and standard deviation (σ) are noted. The included data correspond to the period of the TRAQA campaign from 25.06.2012 until 13.07.2012., and covers all stations presented in (Fig. 5.10).

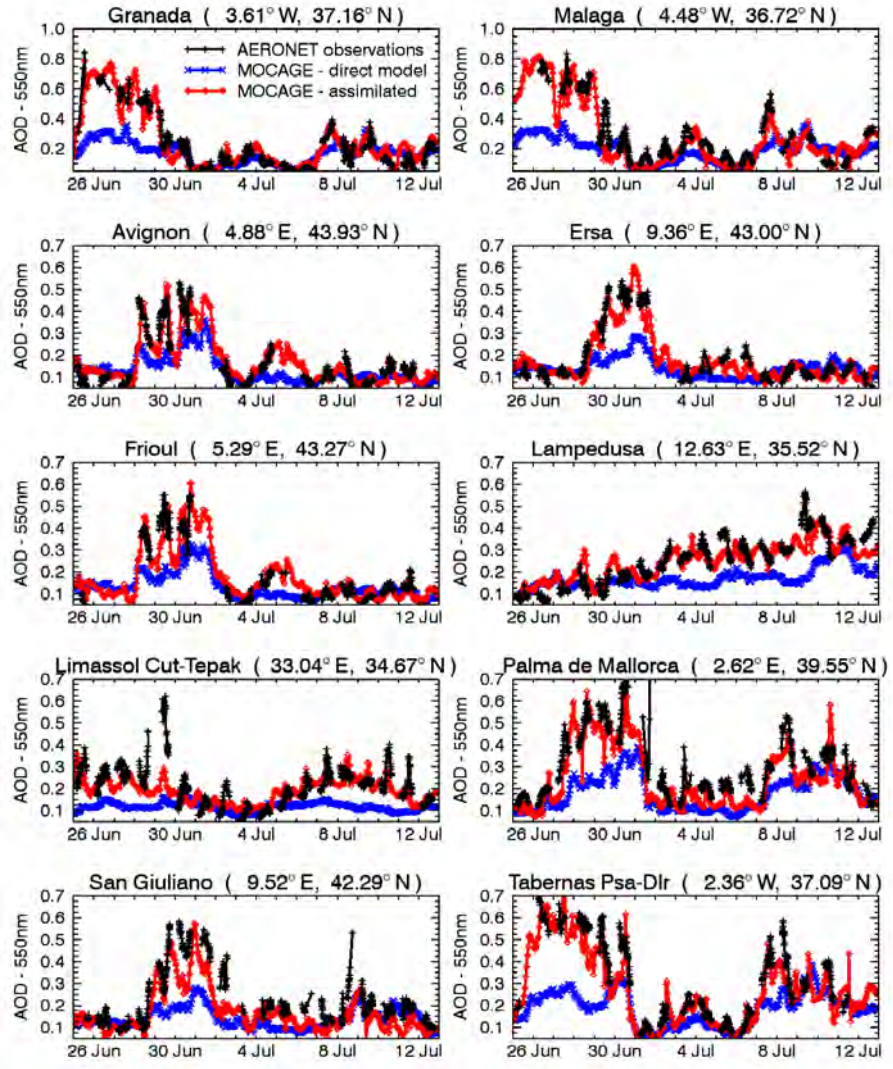

Figure 5.12: Time series of aerosol optical depth at 550 nm of AERONET data, the direct model and the assimilation model run for the period of the TRAQA campaign from 25.06.2012 until 13.07.2012. The presented AERONET data are from ten stations: Granada (ESP), Malaga (ESP), Avignon (FRA), Ersa (FRA), Frioul (FRA), Lampedusa (ITA), Limassol (CYP), Palma de Mallorca (ESP), San Giuliano (FRA), Tabernas (ESP). Correlation, bias and root mean square error for both the direct model and the assimilation model run as compared to AERONET data are given in Table 5.3.

Table 5.3: Correlation (ρ), absolute bias (Δ) and root mean square error (RMSE) between AERONET observations and MOCAGE direct/assimilation run for the period of the TRAQA campaign between 25.06.2012 and 13.07.2012. AERONET site locations are presented in Fig. 5.10.

		MOCAGE direct			MOCAGE assimilation		
	N_{obs}	ρ	Δ	RMSE	ρ	Δ	RMSE
Aubiere	225	0.506	0.041	0.110	0.859	0.024	0.074
Autilla	685	0.769	0.012	0.085	0.882	0.008	0.051
Avignon	846	0.851	0.024	0.087	0.896	0.005	0.055
Barcelona	378	0.802	0.110	0.169	0.900	0.038	0.084
Burjassot	488	0.681	0.132	0.191	0.815	0.055	0.119
Cabo da Roca	77	0.965	0.130	0.263	0.939	0.032	0.120
Calern OCA	509	0.784	0.013	0.093	0.905	0.009	0.052
Carpentras	738	0.774	0.023	0.085	0.876	0.005	0.055
Cerro Poyos	193	0.632	0.034	0.061	0.670	0.055	0.074
Davos	210	0.518	0.064	0.091	0.677	0.027	0.063
Ersa	675	0.760	0.043	0.112	0.946	0.011	0.045
Evora	886	0.826	0.010	0.128	0.932	0.019	0.061
Frioul	658	0.871	0.037	0.096	0.952	0.014	0.044
Granada	883	0.677	0.041	0.129	0.930	0.003	0.057
Huelva	1002	0.793	0.010	0.153	0.936	0.034	0.083
Laegeren	208	0.586	0.077	0.128	0.630	0.037	0.103
Lampedusa	1058	0.573	0.084	0.124	0.845	0.006	0.061
Limassol	978	0.338	0.080	0.115	0.640	0.003	0.067
Madrid	904	0.731	0.011	0.097	0.878	0.004	0.058
Malaga	786	0.702	0.101	0.173	0.910	0.047	0.088
Messina	573	0.519	0.068	0.111	0.835	0.020	0.060
Montsec	528	0.662	0.016	0.078	0.892	0.009	0.044
Nes Ziona	593	0.266	0.053	0.111	0.788	0.014	0.063
OHP Observatoire	657	0.742	0.019	0.089	0.886	0.000	0.053
Oujda	330	0.459	0.202	0.221	0.756	0.090	0.116
Palencia	649	0.859	0.030	0.104	0.919	0.002	0.051
Palma de Mallorca	797	0.754	0.129	0.163	0.888	0.048	0.084
Porquerolles	637	0.805	0.005	0.071	0.923	0.020	0.044
Sagres	405	0.901	0.017	0.197	0.958	0.023	0.088
Sede Boker	950	0.240	0.009	0.080	0.552	0.065	0.095
San Giuliano	768	0.675	0.084	0.137	0.908	0.066	0.089
Tabernas	740	0.754	0.129	0.184	0.927	0.038	0.078
Tizi Ouzou	241	0.686	0.195	0.203	0.764	0.079	0.095
Villefranche	480	0.707	0.064	0.113	0.873	0.033	0.067
Zaragoza	916	0.722	0.053	0.100	0.816	0.040	0.073
All sites	29840	0.740	0.050	0.115	0.883	0.006	0.070

of the western Mediterranean. The duration of the event was well simulated by both the direct model run and assimilation model run in all stations, but the intensity was underestimated in the direct model run. However, the assimilation model run matched well the intensity. The second, smaller desert dust event at the end of the TRAQA period was observed only at southern stations. Similarly, the assimilation model run corrected its intensity, underestimated by the direct model run. The stations in the east, like stations in Lampedusa and Cyprus, were not influenced by these dust events. They are mostly influenced by sea salt aerosols, and the data assimilation also here had a very positive impact. The assimilation model run, with only two MODIS overpasses per day, showed also improved hourly variations of AOD in these stations. These variations are not clearly visible in the model direct run, but they are present in AERONET data with the similar amplitudes as in the assimilation model run. The statistics of all AERONET stations confirm the overall positive effect of assimilating MODIS data (Table 5.3).

The AERONET findings confirmed those obtained by the comparison with SEVIRI observations. The scatterplot of all AERONET observations reinforce the same conclusion (Fig. 5.11). Very similarly as in Fig. 5.8, the points with AOD in the direct model are well correlated with the small AOD AERONET observations, while the higher AERONET AOD values are underestimated in the model. This is mostly attributed to the underestimated influence of desert dust events in the direct model run. The assimilation model run corrects AODs and significantly improves the statistical parameters.

In-situ aircraft concentration measurements The comparison with AOD independent observations showed an undoubtedly positive effect of the AOD assimilation on the model performance. The important question is what is the impact of AOD assimilation on other aerosol properties: we assessed impact on the aerosol concentration, the vertical profile and the size distribution. We first compared the modelled aerosol number concentrations with the aerosol concentrations measured by the PCASP instrument. During the TRAQA period, flights with the ATR-42 were effectuated during 9 different days, carrying the PCASP onboard. They passed over the whole western Mediterranean basin using Toulouse, Marseille and Corsica airports, and they covered different meteorological and air quality conditions. Fig. 5.13 presents three representative timeseries from such flights: flight A of 26.06.2012 from Corsica to Toulouse (Fig. 5.13a), flight B of 27.06.2012 from Marseille to Corsica and back to Toulouse (Fig. 5.13b), and flight C of 29.06.2012 from Corsica to Toulouse (Fig. 5.13c).

During Flight A (Fig. 5.13a), aerosol concentrations were rather low, except for the pollution around the Toulouse agglomeration measured at the

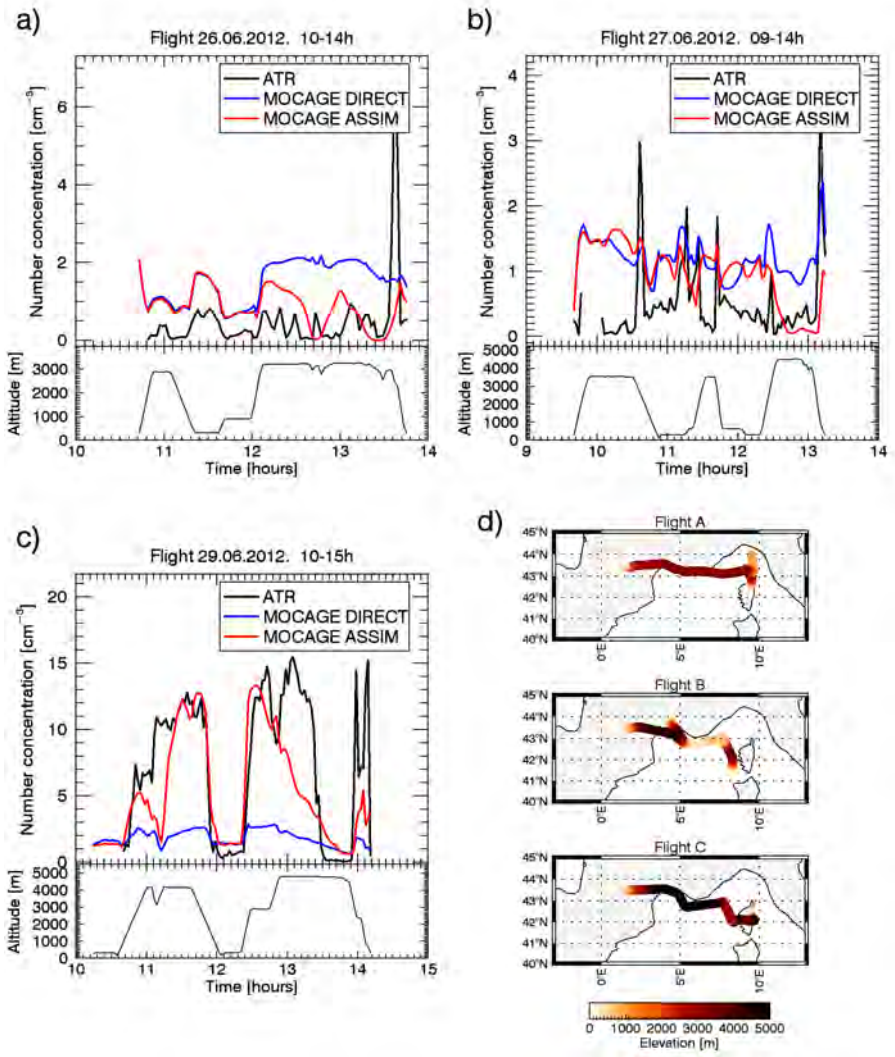

Figure 5.13: Aerosol number concentration [cm^{-3}] from the PCASP instrument onboard of the ATR aircraft for three different flights: Flight A on 26.06.2012 **(a)**, Flight B on 27.06.2012 **(b)**, and Flight C on 29.06.2012 **(c)**. The altitude of the aircraft is also given for all three flights. Also, the maps of flight tracks are presented **(d)**.

flight end. The aircraft first visited the area of the Gulf of Genoa where, because of no AOD observations, the direct model run and assimilation model run show the same aerosol concentrations. The variability in that part of the flight is well simulated with slightly higher modelled aerosol concentrations at these heights than what is measured. Later, on the path to Toulouse, with some

available satellite observations the assimilation model run somehow lowered AOD and approached the measured concentration values to certain limits. With no observations or with sparse ones, data assimilation is not able to have a major effect.

Following the path of Flight B (Fig. 5.13b) we see again rather clear air conditions. The assimilation had a certain effect on the form of the timeseries curve, but it is fairly chaotic and it does not improve noticeably the field compared to the measurements. This could happen if the simulated shape of the aerosol profile somehow differ from the measured one, or if new satellite AOD observations are not available but the declining effect of previous assimilation cycles is still visible.

During Flight C (Fig. 5.13c), the aircraft flew directly through the desert dust plume. The concentrations were elevated in a wide range of heights. The assimilation model run significantly improved the number concentrations, by having them closely resembling the measured ones for most of the plume. If satellite measurements are accurate, after assimilation concentrations at one height can closely correspond to measured ones only if a shape of the vertical profile is well simulated in direct model run. To check this, the modelled and measured vertical profiles should be compared.

In-situ balloon concentration measurements During TRAQA, LOAC flew on three balloons, all launched from Martigues, near Marseille (FRA). Two flights on 29.06.2012, and one on 06.07.2012 are presented in Fig. 5.14. First two flights flew through the desert dust plume. The path of the second flight was near the path of the aircraft Flight C. With the similar time of measurements of both instruments we can directly compare the two types of measurements. The total horizontal motion of LOAC is quite small. Therefore, we will presume that LOAC measurements represent the aerosol vertical profile above the launch place.

The first two flights (Figs. 5.14a and 5.14b) are launched at two different times of the same day, through the same dust plume. In both cases, the assimilation model run matched very closely the measurements. It well simulated both, the shape of the profile and the aerosol number concentration. This was possible by having that the direct model run underestimated the aerosol number concentration, but simulated well the shape of the vertical profile. In the assimilated field, the profile concentrations are increased by assimilation. LOAC measurements acquired during the second flight matched well with the assimilation model run profile, confirming that assimilated aerosol concentration at one height (the case of aircraft measurements, Fig. 5.13c) can be correct only if the profile shape is well simulated by the direct model.

The third LAOC flight (Fig. 5.14c) measured moderate aerosol concentrations coinciding with an air pollution episode. The assimilation model run matched well with measured concentrations. The direct model run underestimated the concentrations only in lower levels. But, when compared with the assimilation model run, the assimilation changed the aerosol vertical profile significantly: the concentrations were increased much more in the lower levels, while in higher levels the change was less important. The different shape of the profile in the direct model run and the assimilation model run is a result of the continuous assimilation of AOD over many assimilation cycles, and the mixing of the aerosols coming from different levels and regions where they were already assimilated (or not) in previous assimilation cycles. This demonstrates that the continuous assimilation of good quality AOD observations can correct a shape of the aerosol vertical profile, although a single AOD assimilation cycle

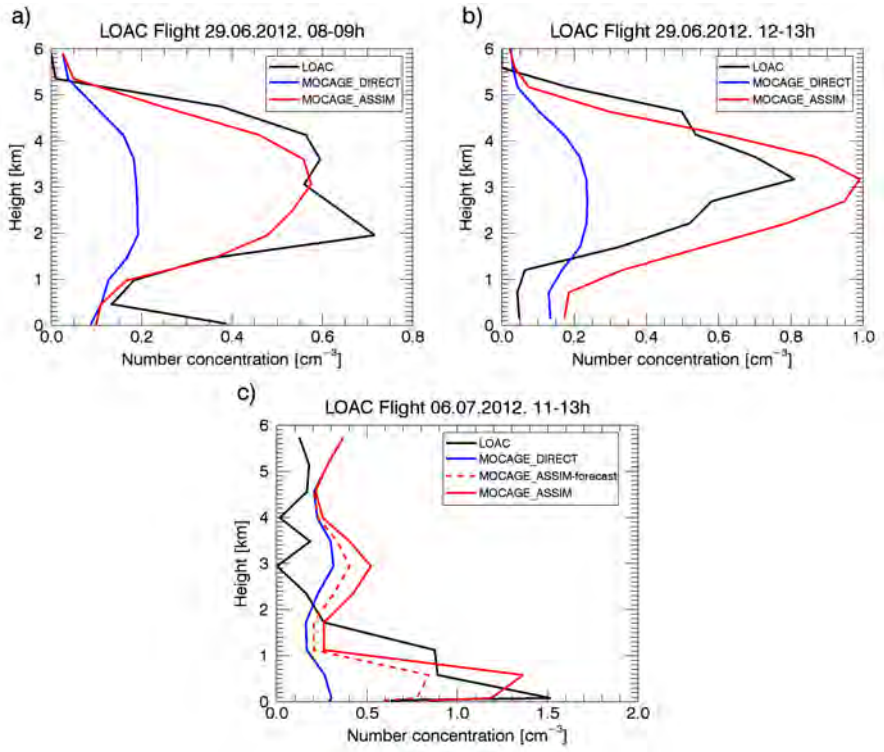

Figure 5.14: Aerosol number concentration (cm^{-3}) from the LOAC instrument onboard on meteorological sounding balloons. The presented flights are performed: in the morning of 29.06.2012 **(a)**, at noon of 29.06.2012 **(b)**, and on 06.07.2012 **(c)**. LOAC measurements are compared with the direct model run and the assimilation model run. For the third flight, we also present the forecast of the assimilation model run (dashed line).

can only expand or shrink the profile shape. The first two flights (Figs. 5.14a and 5.14b) did not change the profile shape, having aerosols coming from the same direction influenced by the strong low over the British isles.

Size distribution from PCASP The Flight C (Fig. 5.13c) flew through the desert dust plume. We assumed that during this part of the flight the PCASP instrument in its size range detected only desert dust particles, and, with this assumption, we compare the measured size distribution with the size distribution of desert dust aerosols from MOCAGE. The shapes of the size distributions are presented in Fig. 5.15. This figure shows that the size distribution in the model stayed the same after assimilation as expected. As a result, the AOD assimilation only added the aerosol mass, and it did not change the characteristics of the aerosol size distribution. In the measured distribution we see the desert dust distribution only up to 2.5 µm, which is the largest PCASP size class. It appears that the PCASP smaller size classes (below 0.5 µm) contain also other aerosol types and not only the desert dust. The overlapped sizes between the measured and modelled dust distribution correspond only to two MOCAGE bins, from 0.1 µm to 1.0 µm and from 1.0 µm to 2.5 µm. The regridded distribution was presented with dashed line (Fig. 5.15). The first bin of regridded distribution has a higher quantity of aerosols than in the modelled

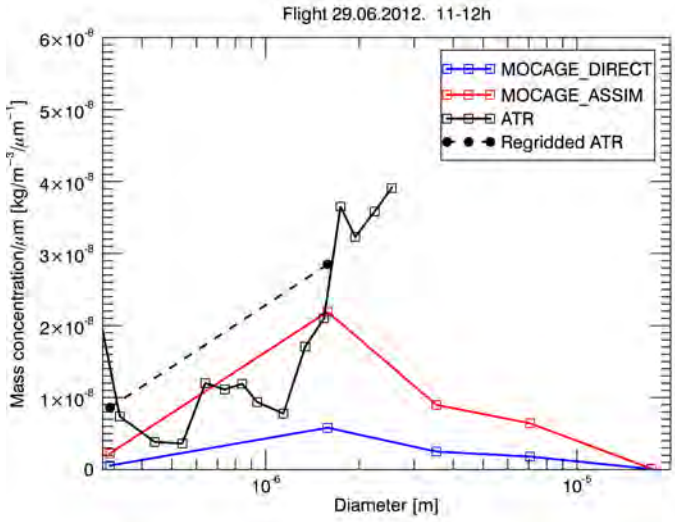

Figure 5.15: Aerosol size distribution measured by the PCASP instrument during the part of Flight C (Fig. 5.13c) compared with the desert dust size distribution from the direct model run and the assimilation model run. The measured size distribution is also regridded to correspond to the MOCAGE bins (dashed line).

distributions, because of possible mixture of different types of aerosols in the aircraft data. The second regridded bin contains more desert dust aerosols than what show the modelled distributions. The assimilation model run has a more similar distribution amplitude. Still, it is difficult to make a more detailed comparison because of the small number of overlapping bins.

5.3.3.7 ChArMEx results

Similar analyses, as performed for the TRAQA period, we repeated for the ChArMEx period. The results are presented in details in Appendix C. Briefly, data assimilation significantly improved the model performance. The ChArMEx period was characterized with more, spatially smaller desert dust and sea salt events. This made statistical parameters in the direct model run slightly worse than for TRAQA. Over the region of western Mediterranean, one desert dust event and one trans-atlantic biomass burning event stand out. The comparison with independent SEVIRI and AERONET data showed that he desert dust event was overestimated in the direct model run. The biomass burning event was present in the model run thanks to GFAS emissions. The stronger part of the event was slightly displaced in the direct model run, which assimilation managed to correct. AOD and number concentration measurements confirmed the positive effects of the AOD assimilation seen for the TRAQA period.

5.3.3.8 Discussion

Results showed that the AOD data assimilation in MOCAGE-PALM is a very efficient technique to bring the model closer to observations. Assimilated fields had better statistical performances than the direct model run in comparison with the assimilated observations, and also with independent AOD observations.

The data assimilation can give its best results if the assimilation system is optimally defined. OmF and OmA analysis (Figs. 5.6 and C.2) showed that the assimilation made a significant improvement of the model fields. When looking closer on OmF and OmA histograms, we note that they are not symmetrical. The model and observations are assumed non-biased, and the histograms are expected to be symmetrical, and to follow a Gaussian distribution. The histograms in Figs. 5.6 and C.2 are skewed toward observations for both the TRAQA and ChArMEx periods, having the forecast and the analysis with a tendency to more frequently underestimate than overestimate observations. In

our system, there are more factors that can have an effect on such behavior. First, our model lacks secondary aerosols. Their inclusion would increase the total aerosol mass, and consequently AOD in the system. Although the forecasted field is already strongly influenced by assimilation, by correcting AOD of many gridboxes, the lack of secondary aerosols is supposed to have an influence on OmF statistics. A secondary aerosol module in MOCAGE is being developed in parallel with the developments made during this thesis (the inclusion of the secondary aerosol module is one of objectives of PhD thesis of Johnatan Guth at CNRM), and we expect this will improve OmF statistics.

The skewness of OmA statistics is influenced by another element. Our assimilation system is more efficient in lowering overestimated AOD values in the model than increasing underestimated values. This can be seen in Figs. 5.7 and C.1, and it is also noticeable in the comparison with independent observations Fig. C.3. Its cause is directly the way how matrices **B** and **R** are defined in this experiment. By defining variances as the percentage of modelled and observed quantities, and making this factor two times smaller for observations, we penalized the high AOD values in the model. In the assimilation process, the overestimated model value can be a lot more easily moved toward the observed value, than the underestimated model value. This directly affects OmA statistics, and reflects in OmF statistics.

One of possible solutions would be to change the matrix **R**. Probably one of the simplest (ad-hoc) approaches would be to limit the observation error up to a fixed value, which would give more weight to observations in the case of high observed AOD. This would have a partial effect, influencing only observations above certain AOD. Another approach would be to try to define the matrix **B** differently. Studies show that a rigorously defined matrix **B** can slightly improve the analysis quality (Massart et al., 2012). In MOCAGE-PALM, in the framework of MACC project, the influence of differently defined matrices **B** was assessed for the case of ozone assimilation (Jaumouillé et al., 2012). One of the approaches was the percentage method used in our experiment. The second approach was the monthly posteriori diagnostics (Desroziers et al., 2005) computed from the data of a month before. It is a method used operationally in MACC. It is better adapted for operational purposes since the data from the past is readily available. The final approach is calculating diagnostics from an ensemble of runs with perturbed emissions with homogeneous or calculated correlation length scales. The main conclusion is that all methods significantly improve the modelled field with small differences in the statistics obtained by different methods. The percentage method that we use is the least costly to calculate the matrix **B**.

For systems that are biased, because of assumptions in both observations and models, a method of bias reduction can be implemented (e.g. Dee and

Uppala, 2008). First, it is necessary to estimate the model bias, for example a regression model where coefficients can be estimated by first guess departures. Then, for assimilation, the computed bias should be subtracted from observations. These modified observations would be assimilated in a bias-free system.

The impact of the AOD assimilation to the model found in our study is coherent with findings of other studies (Zhang et al., 2008; Liu et al., 2011; Schutgens et al., 2010). Our approach is similar to the approach used by Benedetti et al. (2009) since we used the same control variable. They derived the matrix **B** using the NMC method (Parrish and Derber, 1992). Satellite AOD errors are defined for retrievals over water by using a multiregression formula, and for retrievals over land by using the percentage approach with defining a minimal possible error. Their 4D-Var analysis results showed qualitatively a very similar impact of assimilation as in this study.

We assimilated MODIS data that have two overpasses per day during daytime. Satellite data with higher temporal resolution exist. SEVIRI data with a temporal resolution of 15 minutes was used as independent data to evaluate the results. Assimilating such data would further improve the agreement between observations and the assimilation model run, Both considered periods in our experiment are in summer with a good probability of having a cloud-free field, and two overpasses per day were able to cover a significant part of the control domain each day. Possibly, a higher temporal resolution data for assimilation could have a stronger effect, especially during the winter-time.

We would like to point out the importance of the resolution of the model for the desert dust dynamical emission scheme. In our simulations, we have tried two different limited-area domains, MEDI02 that covers the region of desert dust emission scheme over Sahara and another domain that does not cover it. The MEDI02 has an superior resolution than the scheme, and all input data is regridded to the resolution of the scheme. In the second case the scheme used input data from the global domain of $2^\circ \times 2^\circ$, which limited the scheme to this coarser resolution. The emission with the global domain were noticeably lower in comparison with the emissions with MEDI02 domain. The MEDI02 run was better correlated with observations over Mediterranean basin. We suspect that such behavior of the coarser resolution arrives from the smoothing of wind field. These findings are in agreement with the analysis of Bouet et al. (2012).

5.3.3.9 Conclusion

In this study we presented the developments and the validation of the aerosol optical depth assimilation system in MOCAGE-PALM. The system assimilates AOD with the 3D-FGAT method and uses total 3D aerosol concentrations as the control variable. We defined the background and observations errors as the percentage of their respective values. Assimilated observations are super-observations with the same resolution as the model, derived from MODIS L2 data. The data is assimilated with a one hour assimilation cycle over the region covering the North Africa, Mediterranean basin, and South Europe for the periods of both TRAQA and ChArMEx campaigns in summer of 2012 and 2013. Comparing the direct model and assimilated run with assimilated observations, the analysis values came closer to observations, improving all statistical parameters. The comparison with SEVIRI AOD observations, as an independent dataset, confirmed the significant positive effect of the AOD assimilation to the model.

The TRAQA and ChArMEx campaign data, with their collection of in-situ measurements, were used to evaluate the impact of the AOD assimilation on aerosol concentrations, vertical profile and size distribution. The comparison with aircraft data showed that the assimilation can have from a strongly positive to neutral effect on aerosol concentrations at one measured height. Concentrations can be improved by the assimilation when the shape of the vertical profile is correctly simulated by the direct model. The shape of the aerosol vertical profiles cannot be changed during one assimilation cycle because AOD observations do not contain any information on the vertical particle distribution. The shape in the analysis can be considered as the shape of the forecast changed by a certain multiplicative factor. This is also seen in the comparison with the balloon data. But, comparisons also showed that the shape of the vertical profile can be changed and improved by the AOD assimilation in the case of a continuous assimilation over many assimilation cycles. Aerosol that are coming from different locations and directions, if assimilated and corrected in previous cycles, can improve the shape of the vertical profile. The AOD assimilation can have a slightly different effect on different types of aerosols and bins because of different optical parameters, but the characteristics of the size distribution did not change during the assimilation, and only the aerosol quantity was adjusted.

The AOD assimilation proved to be a very efficient technique to improve the model performance. The quality of the analysis could be improved by tuning the error covariance matrices, and/or using different methods to calculate the

matrix **B**. Also, the anticipated inclusion of the secondary aerosol module is expected to improve the assimilation results.

5.3.4 Assimilation of the volcanic plume

The volcano Eyjafjöll in Iceland in April and May 2010 provoked major disruptions of air traffic over Europe. The volcano was active from 14 April to 23 May, with two periods of strong explosive activity from 14 April to 18 April, and 5 May to 20 May, during which high quantities of fine volcanic ash were ejected. The small size of ash promotes its long-range transport. During the eruption, models and remote sensing measurements were used by VAACs (Volcanic Ash Advisory Center) for making decisions of flight bans. The change from zero tolerance to new ash threshold values as the flight-ban criterion, raised during the Eyjafjöll eruption in 2010, required an additional effort in modelling to make ash plume quantitative predictions.

The main uncertainty source in volcanic ash prediction is the emission term. For an accurate prediction it is necessary to know the ejected mass, the size distribution of ash particles and the plume height and vertical distribution as a function of time. Eruptions vary a lot (Woods, 1995), and often these are hard to estimate (Zehner, 2012). The experience from the Eyjafjöll eruption showed that the predictions of plume extent and ash concentrations are difficult to make and it is necessary to achieve progress in combining models with available Earth observation data (Zehner, 2012). Possibilities include inverse source estimation and data assimilation. The inverse source modelling is recently used for SO_2 eruptions (Eckhardt et al., 2008; Kristiansen et al., 2010), and for ash plumes (Kristiansen et al., 2012), while data assimilation is only now emerging (Lange and Elbern, 2014). These techniques improve the usage of information from observations, having in mind that in 2010 VAACs often relied on qualitative comparison with satellite imagery using brightness temperature differences (Millington et al., 2012).

In this experiment we assimilate satellite AOD measurements and try to evaluate the impact of data assimilation on a forecasted volcanic ash plume during the Eyjafjöll eruption.

Table 5.4: The mean diameter of bins for volcanic ash particles and their mass fractions in the time of eruption used in CTM MOCAGE.

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6
diameter [µm]	0.2	0.65	2	6.5	20	65
mass fraction	0.001	0.005	0.052	0.207	0.69	0.045

5.3.4.1 Model setup

To define the emission term in MOCAGE, we calculated the ejected mass indirectly from the column height. We used the relationship from Mastin et al. (2009) (Eq. (4.1)). The column heights, that we use, were reported by VAAC during the eruption. They were compiled from observations and subjectively adjusted by VAAC if necessary (Flemming and Inness, 2013). The vertical distribution within the eruption column was considered uniform. The size distribution of the ash is presented in Table 5.4. The values are based on the work of Hobbs et al. (1991). The mass density of volcanic ash is considered to be $2300\,\mathrm{kg\,m^{-3}}$. The Eyjafjöll eruption produced mostly silica-rich trachyandesite ash (Langmann et al., 2012), and for the optical properties of volcanic ash we used andesite measurements by Pollack et al. (1973).

The limited-area domain used in this experiment covers Europe and Iceland with a resolution $0.2° \times 0.2°$. The precise boundaries are $[26°W - 46°E, 28°N - 72°N]$ (Fig. 5.16). The boundary conditions are provided by the global domain which is run with the resolution of $2° \times 2°$. ARPEGE analysis meteorological fields are used to dynamically force the model. Model emissions are equivalent to those used in Section 5.3.3.4, except that biomass burning emissions come from the inventory of Lamarque et al. (2010), which are monthly climatological data.

5.3.4.2 Assimilation system setup

To define the error covariance matrices $\mathbf{B}$ and $\mathbf{R}$, we specify variances as percentage of the values of the background and observations. When compared with the previous study (Section 5.3.3), uncertainties in the model are apparently larger because of the larger uncertainties of the source term of the

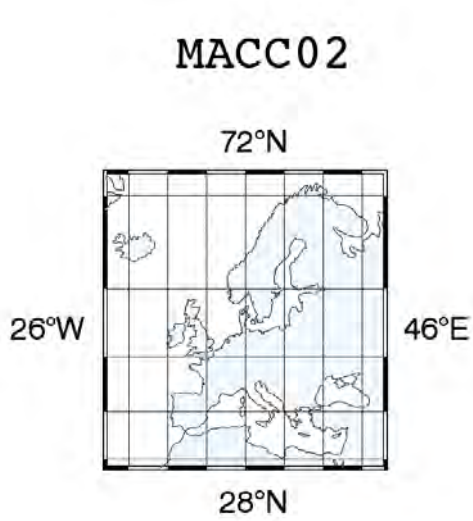

Figure 5.16: The limited-area domain used in this experiment. This is also the control domain for assimilation.

eruption. AOD observations in situations of volcanic plumes can also have larger uncertainties. In plumes, volcanic ash can be mixed with water or ice and this was especially pronounced during the first part of the Eyjafjöll eruption from 14-18. April (Langmann et al., 2012; Zehner, 2012). Having this in mind we decided to define the error covariance matrices as in Section 5.3.3.3.

The assimilation cycle in the experiment has a length of one hour. The cost function is minimized until the convergence criterion is reached, with the maximum number of iterations equal to 200. Horizontal and vertical correlations are homogeneous, with a 0.4° horizontal correlation, and a vertical correlation of one model level. The analysis increment, when found, is added to the end of each assimilation cycle.

5.3.4.3 Observations

The assimilated data from MODIS, and independent data from SEVIRI are described in Section 5.3.3.1 and Section 5.3.3.2. The only difference is that in MODIS data the "Deep Blue" product is not considered due to the fact that bright surfaces cover only a small part the control domain.

5.3.4.4 Experiment

The goal in this experiment is to assimilate the AOD observations during the Eyjafjöll eruption from 14.04.2010 until 21.05.2010, and to evaluate the impact of assimilation on the prediction of the volcanic plume. To do so, we run two MOCAGE configurations, one with and one without assimilation. The simulation without assimilation we refer as the *direct model run* and the simulation with assimilation as the *assimilation model run*.

The model is run with a spin-up period of two weeks, which in the assimilation run is split in the direct model spin-up period of nine days, and the assimilation model spin-up period of five days. In this way, at the beginning of the concerned period the assimilation model run already achieved important AOD correction.

5.3.4.5 Results

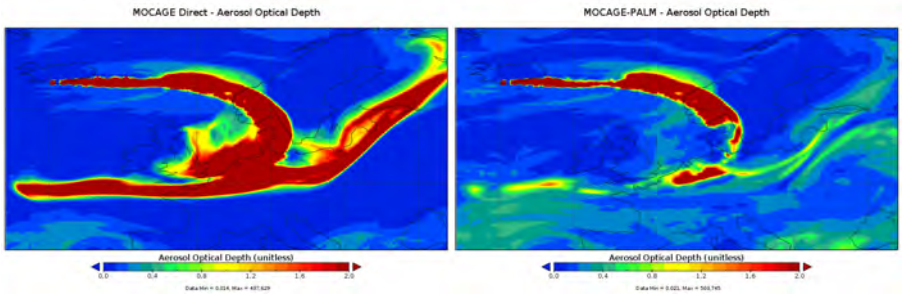

Figure 5.17: The aerosol optical depth over Europe on 16.04.2010 at 14h UT, (**left**) simulated in MOCAGE by the model direct run, and **(right)** simulated in MOCAGE in the MODIS assimilation model run.

The AOD assimilation has a large impact on the prediction of the volcanic plume. The Fig. 5.17 illustrates an example of the simulated difference between the direct model run and the assimilation model run. Although, in a period with conditions of important cloud coverage over the Atlantic and Europe (http://landweb.nascom.nasa.gov/browse), data assimilation significantly influenced the volcanic plume, its intensity and extent. The general conclusion is that the direct model run often overestimated the plume from the Eyjafjöll volcano, in terms of AOD and its spatial extent. Another example

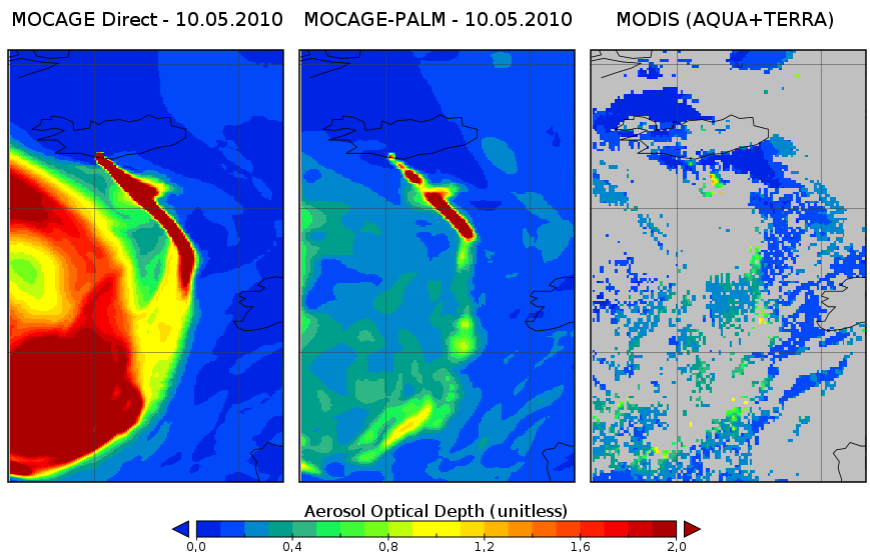

Figure 5.18: The aerosol optical depth over Iceland and the Atlantic ocean south of Iceland on 10.05.2010, (**left**) simulated in MOCAGE by the model direct run, **(center)** simulated in MOCAGE by the MODIS assimilation model run, and **(right)** assimilated MODIS data.

(Fig. 5.18), where we compare the model fields with assimilated observations, shows that the extent of the plume was well simulated in the model direct run, but the intensity was too high. After assimilation, the intensity was more similar to what was observed. This represents a very important correction for determining the concentrations above or below threshold values. Also, it should be noted that the difference between the direct model run and the assimilation model run comes also from previous observations, not only observations present in Fig. 5.18. The OmF (one hour forecast) and OmA statistics confirm the positive assimilation impact also for this period (Fig. 5.19). The bias between observation and model is lowered by the assimilation.

A comparison with independent AOD observations from SEVIRI shows that the assimilated field have better statistics than the direct model run (Fig. 5.20). Also, it shows that the direct model run largely overestimated AOD during this period. We have to be careful with this conclusion. First, satellite products (both MODIS nad SEVIRI) during the volcanic ash retrieval use the models that do not have a sophisticated description of the ash, which leads to difficulties or errors in the retrieval. Secondly, when looking at the distribution of available SEVIRI data, is is noticeable from Fig. 5.20 that we do

not have a lot of SEVIRI observations with high AOD. Compared to MODIS data in the same region, MODIS data provide a larger number of higher AOD than SEVIRI data. The reason is that there are very few SEVIRI observations

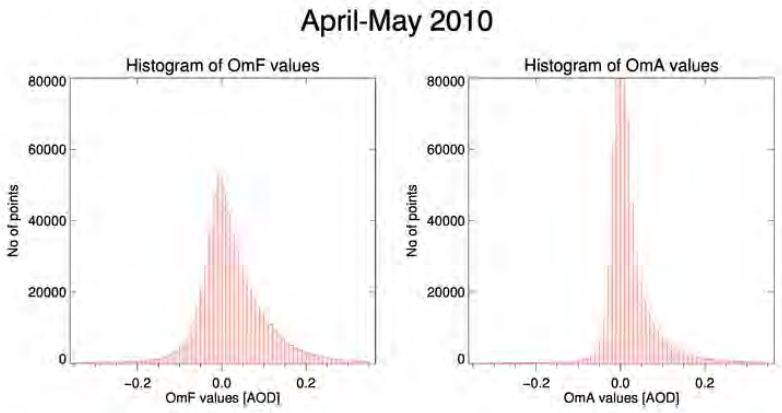

Figure 5.19: Histograms of (**left**) OmF (Observation minus Forecast), and (**rigth**) OmA (Observation minus Analysis) for the period from 14.04.2010 until 21.05.2010. over the MACC02 domain Fig. 5.16.

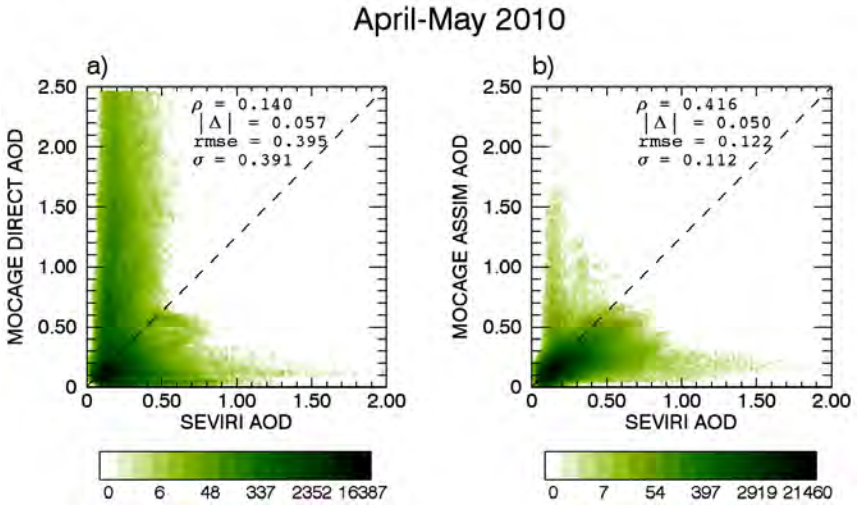

Figure 5.20: Scatterplots of aerosol optical depths from the independent observation dataset (SEVIRI) and the simulations: **(a)** the direct model run, and **(b)** the assimilation model run. In each panel, correlation (ρ), absolute bias (Δ), root mean square error (RMSE) and standard deviation (σ) are noted. The assimilated data correspond to the period from 14.04.2010 until 21.05.2010, and covers the domain with boundaries $[26°W - 20°E, 45°N - 72°N]$ (Fig. 5.16).

of the plume. In MODIS data, the plume is often hidden by clouds, but the lack of plume observations is a lot more pronounced in SEVIRI data, and this provides generally low AOD of SEVIRI. Often, AOD of the plume are not retrieved even in the cloudless conditions in SEVIRI. Factors that possibly have an influence on this are: large zenith angles because of high latitudes of the plume where the retrieval is more difficult, internal limitations in the retrieval, like e.g. cloud mask, that could include also plume pixels. Although we do not see plume pixels in the SEVIRI data, there are observations in the vicinity of the plume. Often in these regions, the MODIS assimilation constrained the plume if it was overestimated in the direct model run. This is consistent with SEVIRI data in Fig. 5.20.

5.3.4.6 Discussion

Data assimilation had a substantial importance for the prediction of the volcanic plume. With observations it is possible to affect the intensity and extent of the plume. As the input data we have used the *a priori* column heights inferred during the eruption. After the event, corrected column heights were also assembled (Arason et al., 2011; Webster et al., 2012), but are not used in this experiment. This makes our model run comparable to the real-time prediction, and we can use this run to evaluate its quality. The only difference in the model between the direct model run and the real time prediction of our system is using ARPEGE analysis to dynamically force the model instead of ARPEGE forecasted fields. The MODIS data that we assimilated is the MODIS standard processed data. MODIS has also a near real-time (NRT) products generated within 3 hours of observation time, with a similar quality as the standard product (http://lance modis.eosdis.nasa.gov). This means that the assimilation model run that we performed gives us also an idea of the prediction potential of assimilation for the fast-response output (in the case when NRT products would be used just upon their publishing).

The results confirm that the biggest uncertainty lies in the source term. Emission rate is indirectly calculated, and estimates can be highly uncertain (Tupper et al., 2009). Also, it should be noted that the emission rate of volcanic aerosols is quite sensitive to the column height: in our parameterization the 7 km column height gives 62% more emitted aerosols than a 6 km high column. The uncertainty of the column height in radar observations is in the range of 1 km (Arason et al., 2011), and it is even bigger for *a priori* column height estimates (Webster et al., 2012).

The direct model run more often overestimated than underestimated the extent of the plume in our experiment. In our assimilation system, overestimation in the volcanic plume extent in the model is a better option than its underestimation. Assimilation changes aerosol quantities according to the background information. If the extent is underestimated compared to observations, the observed AOD would correspond to very few modelled volcanic aerosols in that location. Then, the analysis increment would be distributed among other aerosol types available at that location. In the case of the Eyjafjöll eruption, these are mostly sea salt and other aerosols in the lower troposphere that would be added. Volcanic aerosols can be significantly added to the model only if they are already significantly present at that location compared to other aerosols. Adding lower tropospheric, instead of volcanic aerosols would not contribute to the realistic transport and prediction of the plume.

Assimilation can reduce uncertainties in quantity and extent of the plume, but not in the size distribution. The size distribution is important for processes such as wet deposition or sedimentation, and it can vary in the course of an eruption. There is a possibility to implement size distribution correction by comparing modelled and observed deposed particles during the eruption, but observations are usually not readily available during eruptions.

A possibility to further improve the prediction of plumes is to combine data assimilation and inverse modelling. In assimilation, we use available information from observations to improve modelled fields and their subsequent evolution. When the plume is observed by a satellite for the first time, data assimilation is the only tool necessary to extract available information from observations. But, inverse modelling could further improve model performance for the period between emission and the time of first available observations of the plume. Also, inverse modelling could also give some information of the vertical distribution of aerosols in an eruption column.

One of the reasons why the plume extent is often overestimated in the direct model run is the fact that the vertical distribution in the column is assumed uniform. In reality, aerosols are more and less abundant in different parts of the column. The uniform distribution limits the places where ash is not predicted but it is present in reality. Also, it makes ash present in places and levels where there is no ash. To get the non-uniform vertical distribution in the column in the first place, one should use ash column models. They simulate processes in the column in details, account different eruption styles, uncertainties in magma discharge, etc. To run them, it is necessary to combine volcanological a priori data, local and satellite observations.

Part of the forecasts of VAAC centers rely on brightness temperature differences observed in thermal infrared, which reveals volcanic ash and distinguish it from water/ice clouds (Watkin, 2003; Millington et al., 2012). Similar approach is used to retrieved ash mass loading [$\mathrm{kg\,m^{-2}}$] and different volcanic ash properties (Prata, 1989). This technique works well if volcanic plume is not too thick, too dispersed, in view with ice clouds or mixed with them (Prata, 2009). There are ash mass loading retrievals from different satellites. An integrated quantity as the mass loading could be assimilated in the assimilation system without significant difficulties. This could supplement AOD observations in the situations where it is more difficult to retrieve AOD over plume pixels and during the night-time when the AOD retrieval in visible wavelengths is not possible.

5.3.4.7 Conclusion

This experiment showed that for successful volcanic ash prediction it is necessary to rely, besides on models, also on observations. The source term in the volcanic aerosol emission has high uncertainties in emitted mass, column height, its vertical distribution, and particle size distribution. Using observations it is possible to constrain some of these uncertainties. Data assimilation proved to be a very efficient tool for improving volcanic plume prediction, even in periods and regions with high cloud cover. Observations affect modelled intensity and extent of the plume. We can conclude that with present-day's satellite observation products that are rapidly released after the time of acquisition, data assimilation can become an important, complementary technique for improving model predictions.

5.4 Lidar assimilation

The information on the aerosol vertical profile can be obtained from lidar observations. Incorporating this information in MOCAGE-PALM is an important improvement in the model. For data assimilation of lidar profiles, it is necessary to develop an observation operator which links the total concentration in the model space with observed lidar quantities in the observation space. By using 3D total concentration as the control variable, it was possible to develop the system which is able to efficiently assimilate both aerosol optical depth and lidar profiles (in the same time or separately).

In this section we present the developments which enabled the assimilation of lidar profiles in MOCAGE-PALM, and also the first tests which show the good functioning of the system.

5.4.1 Observation operator

We developed the lidar profile observation operator in MOCAGE-PALM. First it sums all individual species to calculate the total concentration. Also, it determines the relative mass contributions among bins which, during an assimilation cycle, are considered constant. Then, the lidar observation operator transforms a control variable into lidar observed quantities. The observation operator simulates measurements of an elastic backscatter lidar. Quantities that are considered and that can be assimilated are:

- backscatter signal
- aerosol extinction coefficient: α_{aer}
- aerosol backscatter coefficient: β_{aer}.

We implemented three wavelengths for which the observation operator can calculate these quantities. One wavelength is in the UV part of the spectrum at 355 nm, one is in the visible at 532 nm, and one is in the infrared part at 1064 nm.

The backscatter signal is described by the lidar equation (Eq. (1.38)). In our case, the observation operator considers that the signal is and range-corrected by its initial signal intensity (P_0):

$$P_l = K\beta \exp\left(-2\int_0^z \alpha\, dz\right). \tag{5.12}$$

The aerosol extinction coefficient (Section 1.8.2.2) and the aerosol backscatter coefficient (Section 1.10.2) are contained in the total (molecular + aerosol) backscatter coefficient β and the total extinction coefficient α in Eq. (5.12). To calculate the aerosol backscatter coefficient from the aerosol extinction coefficient, it is necessary to make an assumption on the lidar ratio (Eq. (1.39)) for each species in the model. The lidar ratio in our observation operator is predefined with the values derived by Burton et al. (2013).

To calculate the molecular part in the backscatter signal (for both extinction and backscattering), it is necessary to take into account different processes. This includes the scattering of the atmosphere, and the influence of certain gases at certain wavelengths.

For all used wavelengths, the part of the signal is attenuated and backscattered by the Rayleigh scattering of the atmosphere. The Rayleigh extinction coefficient is calculated as (Collis and Russell, 1976; Flamant et al., 2013):

$$\alpha_{mol_{Ray}} = 1.16 \cdot 10^{-5} \left(\frac{\lambda_0}{\lambda}\right)^{4.09} \frac{p}{p_0}\frac{T_0}{T}, \tag{5.13}$$

where $\lambda_0 = 550\,\text{nm}$ is the reference wavelength, $p_0 = 101325\,\text{Pa}$ is the reference pressure, and $T_0 = 288\,\text{K}$ is the reference temperature.

Besides the Rayleigh scattering from the atmosphere, atmospheric gases can absorb and reemit radiation. At the implemented wavelengths, ozone (O_3) and nitrogen dioxide (NO_2) absorb light in the UV and visible part of the spectrum, sulfur dioxide (SO_2) in the UV, and oxygen (O_2) and carbon dioxide (CO_2) in the infrared part.

Absorption in the UV and visible part of the spectrum is caused by electronic energy level transitions. To calculate the extinction coefficients of these gases, we used the absorption cross sections reported by Bogumil et al. (2003). The absorption cross section is a function of wavelength and temperature. It is related to the absorption coefficient by:

$$\alpha_{mol_{gas}}(\lambda, T) = n_{gas}\sigma_{gas}(\lambda, T), \tag{5.14}$$

where n_{gas} is the gas number concentration [m^{-3}], and σ_{gas} is the absorption cross section [m^2].

Absorption of CO_2 and O_2 in the infrared is caused by vibrational and rotational transitions, which have a lot finer and a more complex structure than electronic states. To calculate their absorption it is necessary to use a band model. Different types of band models exist. We considered statistical narrow band models, the same approach as in the 6S radiative transfer model

(Kotchenova et al., 2006; Vermote et al., 2006) which we used as a reference. Statistical narrow band models divide bands into segments, consider lines within a segment as randomly distributed and line intensities to follow a certain probabilistic function. We implemented two statistical narrow band models. The Goody model (Goody, 1952) which considers the probability distribution of line intensities as:

$$p(S) = \overline{S}^{-1} \exp\left(-\frac{S}{\overline{S}}\right), \tag{5.15}$$

where S is the line intensity, and $\overline{S}$ is the mean line intensity in the band segment; and the Malkmus model (Malkmus, 1967) which assumes the probability distribution as:

$$p(S) = S^{-1} \exp\left(-\frac{S}{\overline{S}}\right). \tag{5.16}$$

This Malkmus model probability function gives a larger probability to weaker lines than the Goody model. Considering a Lorentz line profile, the transmittance of a band segment for the Goody model is (Ellingson and Gille, 1978):

$$T_r = \exp\left(-\frac{\overline{S}m}{\delta_l}\left(1 + \frac{\overline{S}m}{\pi\alpha_L p}\right)^{-1/2}\right), \tag{5.17}$$

and for the Malkmus model (Rodgers, 1968):

$$T_r = \exp\left(-\frac{\pi\alpha_L p}{2\delta_l}\left(\left(1 + \frac{4\overline{S}m}{\pi\alpha_L p}\right)^{1/2} - 1\right)\right), \tag{5.18}$$

where m is the amount of absorber, δ_l is the mean spacing between lines in the band segment, α_L is the average half width of the line, and p is the pressure [atm]. These relations are valid for a homogeneous atmospheric path, where temperature and pressure do not vary. To account for an inhomogeneous path with varying pressure and temperature, we use the Curtis-Godson approximation. This approach reduces the inhomogeneous path to a homogeneous path by scaling variables with averaged constant pressure and temperature. The scaled variables are (Ellingson and Gille, 1978):

$$\overline{m} = \int \Phi(T)\rho_a \, dz \tag{5.19}$$

$$\overline{p} = \frac{1}{\overline{m}} \int \Psi(T) p \rho_a \, dz \quad \Rightarrow \quad \overline{mp} = \int \Psi(T) p \rho_a \, dz, \tag{5.20}$$

where $\overline{m}$ is the temperature-scaled amount of absorber, $\overline{p}$ is the averaged pressure [atm], Φ and Ψ are functions of temperature, ρ_a is the density of absorber [$\mathrm{kg\,m^{-3}}$], and p is the pressure [atm]. Φ and Ψ functions can be

approximated by (Vermote et al., 2006):

$$\Phi(T) = \exp\left(a_\Phi(T - T_0) + b_\Phi(T - T_0)^2\right) \tag{5.21}$$

$$\Psi(T) = \exp\left(a_\Psi(T - T_0) + b_\Psi(T - T_0)^2\right), \tag{5.22}$$

where a_Φ, b_Φ, a_Ψ and b_Ψ are constants calculated by Vermote et al. (2006) using the data from the AFGL 1991 compilation of line parameters, and $T_0 = 250\text{K}$ is the reference temperature. With the Curtis-Godson approximation the transmittance for the Goody model becomes:

$$T_r = \exp\left(-\frac{\overline{S}\overline{m}}{\delta_l}\left(1 + \frac{\overline{S}\overline{m}^2}{\pi\alpha_L\overline{mp}}\right)^{-1/2}\right), \tag{5.23}$$

and for the Malkmus model:

$$T_r = \exp\left(-\frac{\pi\alpha_L\overline{mp}}{2\delta_l\overline{m}}\left(\left(1 + \frac{4\overline{S}\overline{m}^2}{\pi\alpha_L\overline{mp}}\right)^{1/2} - 1\right)\right). \tag{5.24}$$

Coefficients $\frac{\overline{S}}{\delta_l}$ and $\frac{\pi\alpha_L}{\delta_l}$ are calculated from the experimental data, from the HITRAN database, with an interval of the band segment of $10\,\text{cm}^{-1}$ (Vermote et al., 2006). The Goody and Malkmus models are similar, but observations showed the Malkmus model gives slightly better results for weaker bands (Soufiani and Taine, 1997). This, we decided to use the Malkmus model. From the transmittance we can get the optical depth:

$$\tau = -\ln T_r, \tag{5.25}$$

and from there we calculate the extinction coefficient.

For all molecular components, the backscattering coefficient is obtained from the extinction coefficient by applying the molecular lidar ratio (Ansmann and Müller, 2005):

$$L_{mol} = \frac{\alpha_{mol}}{\beta_{mol}} = \frac{3}{8\pi}\text{sr}. \tag{5.26}$$

When considering the signal at wavelengths implemented in the observation operator, in the UV (355 nm) it is the Rayleigh scattering that has the strongest signature in the signal (Fig. 5.21). In the visible (532 nm), the influence of the Rayleigh scattering decreases, because it depends on the wavelength as $f(\lambda) \sim \lambda^{-4}$. The aerosol influence does not typically vary significantly with wavelength in the considered part of the spectrum. Because of it, in the visible, and particularly in the infrared, aerosols in average have stronger influence on the signal than the Rayleigh scattering. When considering the gas absorption,

the most important influence has ozone in the visible. Typically, it contributes to the total column optical depth of ≈0.02, which is still a lot lower than the contribution of the Rayleigh scattering or aerosols. Other gases influence even less, considering all wavelengths.

5.4.2 Implementation

To assimilate a lidar profile in our system, similarly to the implementation of the AOD observation operator, all necessary variables, besides the control variable, are averaged during the cycle to be used for the minimization.

Tangent-linear and adjoint tests The results of the tangent linear and adjoint tests (Eqs. (3.22) and (5.9)) confirmed the good functioning of the observation operator, its tangent linear and adjoint operator, for all the three implemented quantities: the backscatter lidar signal, extinction coefficient and backscatter coefficient. In the case of the extinction and backscatter coefficients the tangent linear test is passed whatever the size of the perturbation is. This is true, similarly as for the AOD operator, because of the linear relationship between the concentration and these quantities. For the backscatter lidar signal, because of its more complex dependency on the aerosol and gas concentrations, the result of the tangent linear test was close to the value of 1 (within a couple of percent) for a perturbation up to ≈ 15% of the model state.

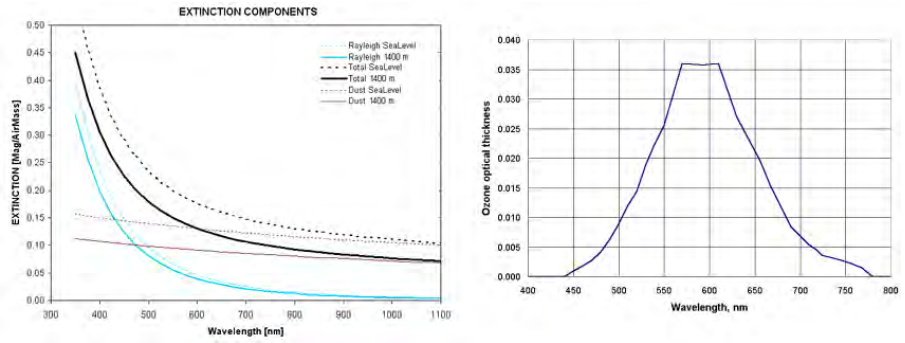

Figure 5.21: Extinction magnitudes of the Rayleigh scattering and of aerosols as a function of wavelengths **(left)** (from Gary (2007)). Ozone optical depth due to extinction in the Chappuis band **(right)** (according to Brooks (2006) and Goody (1952)).

5.4.3 Validation

The observation operator recognizes three different physical quantities that can be obtained from the same lidar profile, and the assimilation can be performed with each of them. The main goal of this validation experiment is to assimilate the same lidar profile with all three quantities separately, and verify that the analyses from all three approaches are coherent with each other.

5.4.3.1 Observations

The lidar profile data that we assimilate is obtained by the R-MAN510 dual polarization lidar located at Météo-France in Toulouse. This lidar system has three channels: parallel and perpendicular polarization channels at 355 nm, and Raman channel at 387 nm (Royer et al., 2013). For assimilation of the backscatter signal, we use measurements of the parallel polarization channel.

The backscatter signal is range-corrected and sky background corrected, and measured in [V m^2]. The system constant K (Eq. (1.38)) is unknown, and in order to eliminate it we have to assume that it does not depend on altitude. Because the constant K contains information on the geometry of the receiver, this is true only for heights above a couple hundred of meters from the lidar (Measures, 1992). Then, to obtain a signal that can be compared to the model, it is necessary to normalize it. The normalization is done with the measured signal value at the altitude z_{ref} where only the molecular backscattering ($\beta_{aer} = 0$) is present:

$$P_n(z) = \frac{P(z)}{P_{ref}} = \frac{P_0 K \beta \exp\left(-2\int_0^z \alpha\, dz\right)}{P_0 K \beta_{ref} \exp\left(-2\int_0^{z_{ref}} \alpha\, dz\right)}. \tag{5.27}$$

In the model, the normalization is done with the modelled signal at the same z_{ref} altitude.

We made the assimilation tests of the system with different profiles obtained by the R-MAN510 lidar system. Here, in this section, we show only one representative lidar profile, which displays an interesting case with a layer of desert dust in the free troposphere, and elevated aerosol concentrations in the planetary boundary layer. The profile is observed on 02.04.2014. It also shows mid- to high clouds in the signal. The reference signal for the normalization was chosen at the altitude below the cloud layers, at 4500 m. The high vertical resolution of the signal of 15 m is degraded to ≈150 m to correspond better to the model vertical resolution.

The extinction and backscatter coefficients are retrieved internally by the system software from information from all three channels using the Klett inversion method (Klett, 1981). The inversion by the system is performed only for altitudes above ≈400 m and below ≈4000 m.

5.4.3.2 Model and assimilation experiment setup

The control domain, is defined with a resolution of $0.2° \times 0.2°$. It spatially covers the Mediterranean basin and the Sahara desert. The precise boundaries are $[20°W - 40°E, 16°N - 52°N]$ (Fig. 5.4). The lateral boundary conditions are provided by the global domain which is run with the resolution of $2° \times 2°$.

ARPEGE analysis meteorological fields are used to dynamically force the model. Model emissions are equivalent to those used in Section 5.3.3.4, except for biomass burning emissions which come from the inventory of Lamarque et al. (2010), which are monthly climatological data.

In order to do an analogous assimilation with all three physical quantities and have their analyses directly comparable to check their coherence, the observations should have the same weight relative to the background in all three cases. The backscatter signal contains the aerosol and molecular part. In assimilation, we consider only the error of the aerosol part of the signal, which we have to convert to the error of the total signal. In order to have equivalent errors in all three assimilation runs, we estimated that, for 355 nm, to the 10% error of the extinction and the backscattering coefficient corresponds to the ≈1 % error of the backscatter signal. These values are used in our tests, together with 25% for the error of the background. The correlation between observations is not considered in any of the three cases. The vertical correlation is set to two model levels, with no horizontal correlation.

5.4.3.3 Results

We performed three different assimilation runs, assimilating separately the extinction coefficient, the backscatter coefficient and the backscatter signal obtained from the same observed lidar profile. Modeled, observed and assimilated profile for each of the three runs are shown in Fig. 5.22 in the units of the observations. The direct model run shows a desert dust plume above ≈1500 m. Heights of the plume are well simulated, but the quantities of aerosols in the

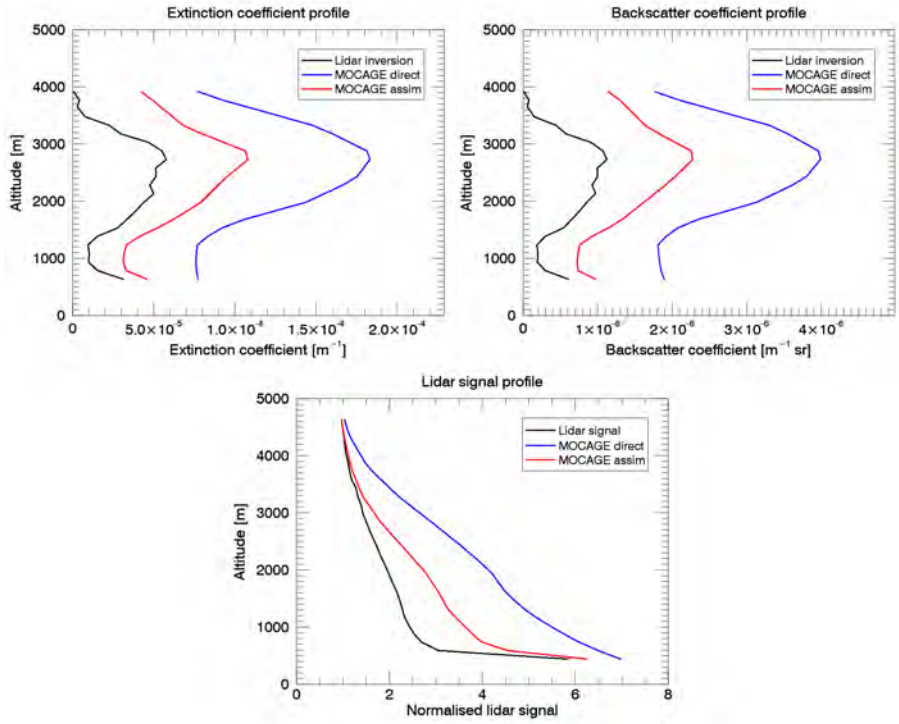

Figure 5.22: Forecast, observations and analysis in observed units for there different assimilation runs assimilating: **(top left)** the extinction coefficient [m^{-1}], **(top right)**, the backscatter coefficient [m^{-1}], and **(bottom)** the normalized backscatter lidar signal (Eq. (5.27)) over Toulouse [$1.5°E, 43.6°N$] on 02.04.2014 at 20:00 UTC.

plume are overestimated compared with the lidar observations. The analysis is between the model and the observations in all three cases.

We can directly compare the results of all assimilation runs by looking at their influence on the aerosol concentration profile (Fig. 5.23). The assimilation of the backscatter signal added a lot of aerosols in the planetary boundary layer which are not present in the direct model run. The backscatter coefficient and extinction coefficient assimilation runs could not add them because the lidar data was not inverted for these altitudes. Above the boundary layer, the desert dust plume quantities are lowered by the assimilation, and the results from all three runs are coherent between each other. This shows the good performance of all variants of the observation operator and the assimilation system.

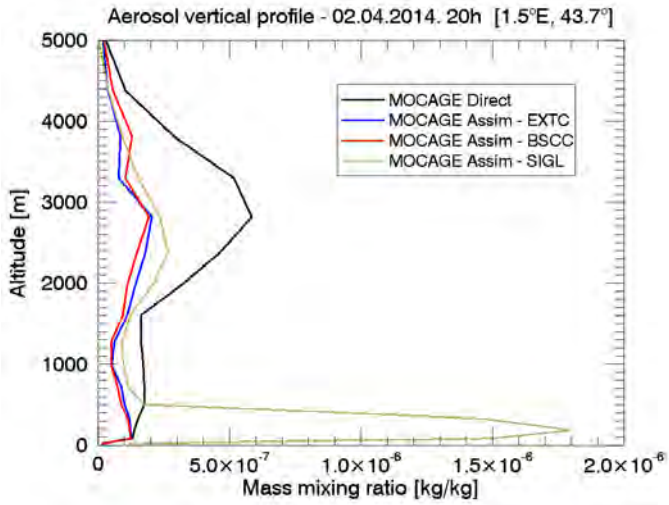

Figure 5.23: The aerosol profile in mixing ratio over Toulouse $[1.5°E, 43.6°N]$ on 02.04.2014 at 20:00 UTC for the direct model run (black), the extinction coefficient (EXTC) assimilation run (blue), the backscatter coefficient (BSCC) assimilation run (red), and the backscatter signal (SIGL) assimilation run (yellow).

5.4.3.4 Discussion and conclusion

While the coherent results of the lidar assimilation using three different physical quantities demonstrated the good performance of our lidar data assimilation system, all assimilated datasets and the assimilation system use a number of assumptions and approximations. This can contribute to some differences in the produced analyses.

The aerosol extinction coefficient and the backscatter coefficient are physical quantities that are retrieved from the backscatter signal. As an inverse problem, their retrieval contains assumptions. Probably the most important is the predefined lidar ratio. The present-day systems can correct the initially presumed lidar ratio using information from the measured signal from different channels. The RMAN-510 lidar system, which we used in our test, with information from polarisation and Raman channels, can estimate the type of aerosols at a certain level, and correct the lidar ratio of that level with the predefined value for the exact aerosol type. The lidar system can distinguish between five different aerosol types. Also, using the Raman channel in inversion can help the retrieval to more accurately separate the molecular and aerosol part of the signal. The Raman inversion is activated only when the Raman channel has a strong and clean signal, which is usually the case during the

night time. The model uses a similar assumption to calculate the backscatter coefficient and the backscatter signal, by predefining the lidar ratio of all aerosol types used in the model.

During the normalization of the observed signal, we assumed that the system constant K (Eq. (5.27)) does not vary with altitude. This is not true for the heights closest to the instrument where the instrument geometry has an important influence on the received signal. This is why we did not take into account the signal for the first couple of model levels. But this gap has an accumulating effect, and all the signal attenuation occurring in the gap was attributed to the first level. By using the vertical correlation, this gap effect was smoothed in the results.

The reference height for the signal normalization is chosen just below the clouds present in the signal. In no-cloud conditions, the reference height would be set to a higher altitude where we would be more certain that the aerosol backscatter part was absent in the signal. Also, we used the parallel polarization channel signal instead of the total signal. Strictly speaking, first, the signal depolarization should have been performed (Rolf, 2012). In this test we assumed that it was adequate to use directly the parallel channel, which was confirmed a posteriori by the test outcomes. Only, it should be aware that in certain cases with desert dust aerosols, lidar radios of the parallel channel can differ from the total signal lidar ratio.

In our test, the observation error covariance matrices in the test should have the same weight compared to the model error covariance matrix in all three assimilation runs. We estimated the error of the signal that would correspond to the error of the backscatter and extinction coefficients. The backscatter signal between different levels is correlated, and for a more rigorous approach, in the case of the assimilation of more lidar signal profiles, the observation error covariance matrix has to be divided into diagonal bloc matrices. These matrices would correspond to different lidar profiles, and their covariance members would describe the correlation between different levels. As a first approximation, other studies also used the diagonal matrix $\mathbf{R}$ with zero covariance members (Wang et al., 2014). The backscatter and extinction coefficients when considered as the physical quantities are not correlated between different levels, and being inverted from the lidar signal they can be approximated to be decorrelated during the inversion.

By assimilating the lidar profile data with above mentioned characteristics and obtaining coherent results we proved the good performance of the lidar assimilation system. A continuous assimilation would be necessary to evaluate the full impact of the lidar assimilation on the forecasted modelled field. The data could include one ground lidar system, a network of the ground lidars or

the spaceborne lidar CALIOP. Wang et al. (2013) studied the impact of the (synthetic) lidar network over the western Europe on the forecast of aerosols. Their results indicate that the data assimilation from a network of only 12 lidar over the western Europe would have a powerful impact on improving of the forecasts of the PM_{10} surface concentrations.

In order to make a continuous backscatter signal assimilation model run, it is necessary to automate the choice of the reference height used for the signal normalization. The reference height could be set high enough where no aerosols are expected only in the case when we are sure that no clouds are between the lidar and the reference height. For this, the lidar system should have the cloud detection detect capability to exclude the profiles with clouds. Also, it should be noted that the uncertainties in the lidar signal are bigger further from the lidar, and ideally, the reference height for normalization would be at the beginning of the molecular zone.

The first assimilation tests with the CALIOP lidar data, showed its high potential by having large spatial coverage along its orbital path. Assimilation of the data from CALIOP does not require the signal normalization because the signal is range-corrected and instrument-calibrated with the system constant K set to unity (Hostetler et al., 2006).

The lidar assimilation of a volcanic plume could be an useful technique to improve the vertical profile of a volcanic plume in the model. It could lower model uncertainties (Chazette et al., 2012a), improve prediction potential and provide insights of the plume evolution for which lidars already proved to be well adapted (Chazette et al., 2012b).

Conclusion and perspectives

With their very diverse properties, aerosols have an extensive influence on the atmosphere, the environment and the human health and activities. To better understand their properties, effects and evolution, significant efforts are put into aerosol modelling and observation systems. Models, as powerful tools to study aerosols, require that aerosols are realistically characterized and represented. The main objective of this thesis is to improve the aerosol representation within the chemical transport model (CTM) MOCAGE. The CTM MOCAGE is a global CTM model with a semi-lagrangian advection scheme simulating gases and primary aerosols. To achieve the objective of the thesis, as a first step, the aerosol representation in the CTM MOCAGE was directly modified by reexamining and improving different schemes, and implementing the more physically detailed schemes and parameterizations. The processes that underwent the important improvements were sea salt and desert dust aerosol emissions, wet deposition and sedimentation. We evaluated the impacts of these changes and compared the modelled fields to observations. The implemented updates significantly enhanced the model agreement with the observations and the AeroCom inter-model comparison data. Comparison with the one year MODIS aerosol optical depth (AOD) data over the oceans showed that the model updates decreased the mean bias (from 0.032 to 0.002) and improved the correlation (from 0.062 to 0.322) in the model. The results confirmed that large uncertainties in models can come from the use of different parameterizations. The discrepancies with the observations are significantly reduced in the model, but some differences remain, particularly in the regions where secondary aerosols could have an important impact. These findings are presented in Sič et al. (2015).

As a complimentary approach to achieve the main objective of the thesis, we also implemented the aerosol data assimilation in the CTM MOCAGE. The assimilation system is able to assimilate AOD observations, as well as different quantities of lidar measurements. The application of data assimilation to aerosols is a relatively new field which got more attention in recent years. Our system uses total 3D aerosol concentration as the control variable, and implements the 3D-FGAT method to assimilate observations into the model. The

AOD observations from space are available from various instruments located on different satellites, which can provide very good spatial and temporal coverage and resolution. We made an extensive validation of the AOD assimilation system by assimilating data from the spaceborne MODIS instrument for the periods of the TRAQA (TRAnsport à longue distance et Qualité de l'Air dans le bassin méditerranéen) and ChArMEx (CHemistry and AeRosol MEditerranean EXperiment) campaigns in the summer 2012 and 2013, respectively, and comparing the directly modelled and assimilated fields with observations. When comparing with independent AOD observations, the assimilated fields have superior statistical indicators compared to the direct model. For example, the comparison with AERONET (AErosol RObotic NETwork) data showed that the assimilation decreased the bias (from 0.050 to 0.006) and increased the correlation (from 0.74 to 0.88). The TRAQA and ChArMEx in-situ measurements were also used to evaluate the impact of the AOD assimilation on other modelled aerosol parameters. The AOD assimilation was able to positively impact the modelled aerosol concentrations, but, as expected, it could not change the size distribution. Also, it was shown that, although a single assimilation cycle cannot change the aerosol vertical profile, the continuous assimilation was able to correct, in some cases, also the modelled vertical profile. The results showed that the AOD assimilation proved to be a very efficient tool to improve the model performance in terms of AOD, but also other aerosol parameters such as concentration.

Lidar observations provide valuable aerosol information on the vertical. We implemented a lidar profile assimilation system for the measurements from an elastic backscatter lidar. The system is able to assimilate the backscatter signal and the retrieved quantities: the backscatter and extinction coefficients. The first performed tests with data from a ground-based lidar showed the promising and coherent impact on the model by assimilating separately all three implemented types of lidar measurements.

During the work on this thesis, we implemented volcanic ash aerosols in the model with their typical point-source emissions. Our developments were based on the rapid-response MOCAGE-Accident system. This addition made possible to evaluate the impact of the AOD assimilation on the model predicting capabilities of a volcanic ash plume. The biggest uncertainties in the volcanic ash modelling come from the source emission term, because of difficulties to obtain information on eruption emitted mass, column height, its vertical distribution, and the particle size distribution. Hypothesis made in the volcanic ash modelling often lead to the overestimation of the volcanic aerosols. We assimilated MODIS data for the 2010 Eyjafjöll eruption in the Iceland. Data assimilation was able to improve the model volcanic plume prediction, even in the period of relatively high cloud cover in the northern Atlantic, correcting

both the modelled intensity and extent of the plume. This suggested that the aerosol assimilation system would be a very useful complimentary tool for the volcanic ash predictions.

Perspectives The CTM MOCAGE, as a state-of-the-art aerosol model, requests continuous developments of the model components. In this sense, the introduction of the secondary aerosols in the model is already in the process (as the objective of the Jonathan Guth's PhD thesis in CNRM, Toulouse). This will, without doubts, have a positive impact on the direct model performance. It is also expected to improve the assimilation OmF (Observation minus Forecast) statistics and likely the assimilation fields.

The lidar assimilation system showed a high potential to improve the modelled aerosol fields. But, it is necessary to validate the analyses with more extensive datasets. Also, an experiment with continuous lidar assimilation, using ground or spaceborne lidar observations, should be conducted to evaluate the full impact of the lidar assimilation on the model.

The aerosol assimilation system has to be carefully and rigorously defined in order to achieve an optimal performance and full potential. This includes a more rigorous definition of the matrices $\mathbf{B}$ and $\mathbf{R}$ which involves the estimation of the best adapted method for the calculation of variances of the matrices $\mathbf{B}$ and $\mathbf{R}$, and correlation lengths of the matrix $\mathbf{B}$. To give an example, if assimilating lidar profiles from the spaceborne CALIOP lidar, possibly it would be advantageous to use anisotropic correlation lengths whose values would be estimated according to the orbit path geometry, data resolution, modelled field, etc. The other pathways towards the optimally defined assimilation system involve tuning of the system parameters, careful preparation of the observations for assimilation, implementation of the bias correction if necessary, etc. All this generally has to be adjusted for each observation dataset assimilated in the model.

AOD and lidar observations can be measured in various wavelengths. By assimilating one type of observations, but measured in different wavelengths, we could also get some information on the size, or the aerosol type. Then, with this information we could modify the size distribution and aerosol bin distribution in the model. To achieve and implement this in the system, it would be necessary to study the relationship and sensitivity between the size and bin distribution in MOCAGE and the aerosol Angstrom exponent obtained from multi-wavelength measurements.

Assimilation of one type of aerosol observation has a positive impact on the model. But, assimilation of both AOD and lidar profiles together simultaneously could be even more powerful technique to improve the modelled aerosol

field. AOD observations have a high spatial coverage, and lidar profiles bring valuable aerosol vertical profile information. Together, their assimilation could have a more complete impact on the model aerosol representation. However, both datasets are subject to various approximations during their retrieval and assimilation. It is necessary to investigate how much these datasets are coherent, how to prepare the system and datasets for their simultaneous assimilation and how much benefit it could bring.

To add an additional value to the work and results achieved in this thesis, we have in the prospect to make publications on AOD and lidar aerosol data assimilation.

Applicative perspectives of the achieved developments include also their incorporation in the operational code of the model used for the air quality forecasts. The assimilation of the volcanic ash could be used in the predictions of Volcanic Ash Advisory Centers as a complimentary tool, which could provide the improved initial conditions for the fast-response systems or constrain probabilities of already made predictions.

The developments made during this thesis act as a necessary stepping stone for conducting research studies on many current aerosol topics. The improved aerosol representation leads to better results and their confidence, which brings better understanding of the aerosols and their effects. The CTM MOCAGE is well adapted to bring advances in research themes like aerosol long-range transport or air quality. It is advantageous for the future air quality studies to able to assimilate surface in-situ aerosol measurements in the model. Scientific applications of the model developments are also closely linked with participation in the international scientific projects, where models play an important role. In the ongoing ChArMEx project (http://charmex.lsce.ipsl.fr/), the goal is to assess the present and future atmospheric state and impacts in the Mediterranean basin. The CTM MOCAGE contributes to the aerosol modelling efforts in the project. These efforts involve the assessment of: regional aerosol budget, seasonal and long term trends over the Mediterranean basin and their impact on air quality, the future evolution, contribution of the regional emissions and long-range transport in the basin, impact on air quality of long-range transport, etc. The participation of the CTM MOCAGE, with both the direct model and assimilation runs, is expected to contribute to the advances in these questions.

Conclusion et perspectives (en français)

Du fait de leurs nombreuses propriétés, les aérosols ont une influence particulièrement importante sur l'atmosphère, l'environnement et la santé humaine. Afin de mieux comprendre leurs propriétés, effets et évolutions, des efforts considérables sont fait actuellement autour de la modélisation des aérosols et des systèmes d'observations. Les modèles, en tant qu'outils performants pour étudier les aérosols, ont besoin d'une caractérisation et d'une représentation réalistes des aérosols. L'objectif général de cette thèse est d'améliorer la représentation des aérosols dans le modèle de chimie-transport (en anglais, Chemical Transport Model, CTM) MOCAGE. Le CTM MOCAGE est un modèle CTM global et régional prenant en compte un schéma d'advection semi-lagrangien et permettant de simuler les gaz et aérosols primaires. Pour parvenir aux objectifs de la thèse, dans un premier temps, nous avons modifié directement la représentation des aérosols dans le CTM MOCAGE en réexaminant et améliorant les différents processus déjà présents via la prise en compte de schémas et de paramétrisations plus détaillés. Les processus ayant subi les améliorations les plus importantes sont les schémas d'émissions des aérosols du type sel marin et poussière désertique, le dépôt humide et la sédimentation. Nous avons évalué les impacts de ces changements et comparé les champs modélisés avec des observations. Les modifications implémentées ont permis d'améliorer significativement l'accord entre modèle et observations et du modèle avec les données de la comparaison inter-modèle AeroCom. Une comparaison avec un an d'observations MODIS de l'épaisseur optique des aérosols (en anglais, Aerosol Optical Depth, AOD) au dessus des océans a montré que le modèle modifié permet de diminuer le biais moyen (de 0,032 à 0,002) et d'améliorer la corrélation (de 0,062 à 0,322). Les différences entre observations et modèle sont fortement réduites grâce au modèle modifié, cependant ils restent encore des différences, en particulier dans les régions où les aérosols secondaires vont avoir un impact important. Ces résultats sont présentés dans Sič et al. (2015).

Comme approche complémentaire pour répondre à l'objectif de cette thèse, nous avons également implémenté dans le CTM MOCAGE l'assimilation de données des aérosols. Le système d'assimilation de données est capable d'assimiler les observations d'AOD, ainsi que les différentes grandeurs obtenues

par les mesures lidar. Les applications liées à l'assimilation de données des aérosols sont des sujets d'investigations particulièrement récents mais en plein développement. Notre système utilise la concentration totale des aérosols en 3D comme variable de contrôle, et implémente la méthode 3D-FGAT pour assimiler les observations dans le modèle. Les observations d'AOD depuis l'espace sont obtenues à partir de différents instruments à bord de différents satellites qui permettent d'obtenir une très bonne couverture et résolution spatio-temporelle. Nous avons réalisé une validation rigoureuse du système d"assimilation de l'AOD en assimilant les données de l'instrument MODIS pour les périodes correspondant aux campagnes de mesure des projets TRAQA (TRAnsport à longue distance et Qualité de l'Air dans le bassin méditerranéen) et ChArMEx (CHemistry and AeRosol MEditerranean EXperiment) durant l'été 2012 et 2013, respectivement, et en comparant les champs directement modélisés et assimilés avec les observations. Lorsque l'on compare ces résultats avec des observations indépendantes d'AOD, les champs assimilés ont des indicateurs statistiques meilleurs que ceux du modèle direct. Par exemple, la comparaison avec les données AERONET (AErosol ROBotic NETwork) a montré que l'assimilation permettait de diminuer le biais (de 0,050 à 0,006) et d'améliorer la corrélation (de 0,74 à 0,88). Les mesures in-situ de TRAQA et ChArMEx ont également été utilisées pour évaluer l'impact de l'assimilation des AOD sur d'autres paramètres des aérosols modélisés. L'assimilation des mesures d'AOD a permis d'améliorer les concentrations d'aérosols modélisés, mais, comme l'on pouvait s'y attendre, n'a pas permis de modifier leur distribution en taille. De plus, on a montré que, bien qu'un cycle unique d'assimilation ne permette pas de changer le profil vertical des aérosols, l'assimilation en continu était capable de corriger, dans certains cas, le profil vertical modélisé. Les résultats ont montré que l'assimilation des AOD était un outil particulièrement efficace pour améliorer les performances du modèle en terme d'AOD, mais également pour les autres paramètres des aérosols tel que la concentration.

Les observations lidar fournissent des informations importantes sur la répartition des aérosols sur la verticale. Nous avons implémenté un système d'assimilation des profils lidar mesurés pour des lidars à rétrodiffusion élastique. Le système est capable d'assimiler le signal de rétrodiffusion et les grandeurs inversées: les coefficients de rétrodiffusion et d'extinction. Les premiers tests effectués à partir des données d'un lidar au sol ont montré un impact cohérent sur le modèle en assimilant séparément les 3 différents types de mesures lidars implémentées. Ceci constitue un premier résultat très prometteur.

Durant ce travail de thèse, nous avons implémenté les aérosols de type cendres volcaniques dans le modèle avec leurs sources ponctuelles d'émissions typiques. Pour cela, nos développements ont été basés sur le système de réponse rapide MOCAGE-Accident. Cet ajout a permis d'évaluer l'impact de

l'assimilation des AOD sur la capacité du modèle à prédire l'évolution d'un panache de cendre volcanique. La plus grande incertitude dans la modélisation des cendres volcaniques provient des sources d'émissions, pour lesquelles il est difficile d'obtenir les informations concernant l'éruption comme la masse émise, la hauteur de la colonne, la distribution verticale, et la distribution en tailles des particules. Les hypothèses faîtes dans la modélisation des cendres volcaniques conduisent souvent à une surestimation des aérosols volcaniques. Nous avons assimilé les données MODIS relatives à l'éruption en 2010 du volcan Eyjafjöll en Islande. L'assimilation de données a été capable d'améliorer la modélisation du panache volcanique, même lors des périodes de couverture nuageuse relativement importante dans la région de l'atlantique nord, corrigeant à la fois l'intensité modélisée et l'étendue du panache. Cela suggère que le système d'assimilation des aérosols devrait être un outil particulièrement utile et complémentaire pour la prévision et l'évolution des panaches de cendres volcaniques.

Perspectives Le CTM MOCAGE, en tant que modèle de l'état de l'art pour les aérosols, nécessite des développements continus. En ce sens, l'introduction d'aérosols secondaires dans le modèle est en cours (à travers les travaux de thèse de Jonathan Guth, CNRM, Toulouse). Cela devrait, sans aucun doute, avoir un impact positif sur les performances du modèle direct. On s'attend également à ce que cela améliore les champs assimilés.

Le système d'assimilation des lidars a montré un réel potentiel pour améliorer les champs d'aérosols modélisés. Mais, il est nécessaire de valider les analyses avec des jeux de données plus importants. Également, une expérience prenant en compte l'assimilation continue des lidars, en utilisant des observations de lidars au sol ou satellite, devrait être conduite pour évaluer l'impact complet de l'assimilation des lidars sur le modèle.

Le système d'assimilation des aérosols doit être défini de façon précise et rigoureuse dans le but d'atteindre des performances optimales et son plein potentiel. Cela comprend une définition plus rigoureuse des matrices $\mathbf{B}$ et $\mathbf{R}$ qui implique le choix de la méthode la plus adaptée pour le calcul des variances des matrices $\mathbf{B}$ et $\mathbf{R}$, et de la longueur de corrélation pour la matrice $\mathbf{B}$. Par exemple, si on assimile les profils lidars de l'instrument CALIOP, il serait avantageux d'utiliser des longueurs de corrélation anisotrope dont les valeurs seraient estimées en prenant en compte la géométrie de l'orbite, la résolution des données, les champs modélisés, etc. D'autres manières d'atteindre un système d'assimilation défini comme optimal impliquent l'ajustement des paramètres du système, la préparation précise des observations pour l'assimilation, l'implémentation d'une correction du biais si

nécessaire, etc. Tout cela nécessite d'être ajusté pour chaque jeu de données assimilés dans le modèle.

L'AOD et les observations lidars peuvent être mesurées dans différentes longueurs d'onde. En assimilant un type d'observations, mais mesuré dans différentes longueurs d'onde, nous permettraient d'obtenir des informations sur la taille ou le type d'aérosols. Ensuite, avec ces informations nous pourrions modifier la distribution en taille des aérosols dans le modèle. Pour y parvenir et l'implémenter dans le système, il serait nécessaire d'étudier la relation et la sensibilité entre les distributions de taille dans MOCAGE, et l'exposant d'Angstrom des aérosols obtenus à partir des mesures multi-longueurs d'onde.

L'assimilation d'un type d'observations d'aérosols a un impact positif sur le modèle. Mais, l'assimilation de l'AOD et des profils lidars en simultanée pourrait être une technique encore plus performante pour améliorer le champ d'aérosol modélisé. Les observations d'AOD ont une très grande couverture spatiale, et les profils lidar apportent une information importante sur le profil vertical des aérosols. Ensemble, leur assimilation devrait avoir un impact plus complet sur la représentation des aérosols dans le modèle. Cependant, les deux jeux de données sont sujets à diverses approximations durant l'inversion et l'assimilation. Il serait nécessaire ainsi d'investiguer plusieurs éléments: la cohérence des jeux de données entre eux, la préparation du système et des jeux de données pour leur assimilation simultanée, et enfin la quantification du bénéfice apporté par cette double assimilation.

Des perspectives possibles d'application sont d'inclure les développements réalisés dans cette thèse dans le modèle opérationnel utilisé pour la prévision de la qualité de l'air. L'assimilation des cendres volcaniques pourrait également être utilisée comme outil complémentaire dans les prévisions données par les VAAC (en anglais, Volcanic Ash Advisory Centers), en améliorant les conditions initiales pour le système de prévision de la dispersion des cendres en temps réel, ou encore en contraignant les probabilités des prévisions déjà réalisées.

Les développements faits durant cette thèse ont permis d'atteindre un nouveau jalon pour les études/recherches conduites autour des aérosols. La représentation améliorée des aérosols a permis d'obtenir de meilleurs résultats et de contraindre les hypothèses, ce qui a permis d'améliorer notre compréhension des aérosols et de leurs effets. Le CTM MOCAGE est bien adapté pour fournir des avancées dans des thèmes de recherche comme le transport à longue distance des aérosols, ou encore la qualité de l'air. L'assimilation des mesures de surfaces in-situ dans le modèle offrent des perspectives pour réaliser des études de la qualité de l'air. Les applications scientifiques des développements réalisés dans le modèle sont également étroitement liées à la participation dans les projets scientifiques internationaux, où les modèles jouent un rôle important. Dans le

projet ChArMEx (http://charmex.lsce.ipsl.fr/), en cours, l'objectif général est d'évaluer l'état atmosphérique, présent et futur, et son impact sur le bassin méditerranéen. Le CTM MOCAGE contribue à l'effort de modélisation des aérosols dans le projet. Ces efforts impliquent l'évaluation du bilan régional d'aérosols, des tendances saisonnières et à long-termes au dessus du bassin méditerranéen et leurs impacts sur la qualité de l'air, les évolutions futures, la contribution des émissions régionales et du transport à longue distance dans le bassin, l'impact du transport à longue distance sur la qualité de l'air, etc. On s'attend à ce que le CTM MOCAGE, avec à la fois le modèle direct et les simulation avec assimilation, contribue à des avancées autour de ces différents thèmes.

Bibliography

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

R. F. Adler, G. J. Huffman, A. Chang, R. Ferraro, P.-P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, and E. Nelkin. The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *Journal of Hydrometeorology*, 4(6):1147–1167, 2003. doi: 10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

S. C. Alfaro, A. Gaudichet, L. Gomes, and M. Maillé. Mineral aerosol production by wind erosion: Aerosol particle sizes and binding energies. *Geophysical Research Letters*, 25(7): 991–994, 1998.

P. Alpert and E. Ganor. Sahara mineral dust measurements from TOMS: Comparison to surface observations over the Middle East for the extreme dust storm, march 14–17, 1998. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 106(D16):18275–18286, 2001.

M. O. Andreae and A. Gelencsér. Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols. *Atmospheric Chemistry and Physics*, 6(10):3131–3148, 2006. doi: 10.5194/acp-6-3131-2006.

E. L. Andreas. A review of the sea spray generation function for the open ocean. *Advances in Fluid Mechanics*, 33:1–46, 2002.

C. Andronache. Estimated variability of below-cloud aerosol removal by rainfall for observed aerosol size distributions. *Atmospheric Chemistry and Physics*, 3(1):131–143, 2003. doi: 10.5194/acp-3-131-2003.

C. Andronache. Diffusion and electric charge contributions to below-cloud wet removal of atmospheric ultra-fine aerosol particles. *Journal of aerosol science*, 35(12):1467–1482, 2004.

C. Andronache, T. Grönholm, L. Laakso, V. Phillips, and A. Venäläinen. Scavenging of ultra-fine particles by rainfall at a boreal site: observations and model estimations. *Atmospheric Chemistry and Physics*, 6(12):4739–4754, 2006. doi: 10.5194/acp-6-4739-2006.

M. D. Anguelova and F. Webster. Whitecap coverage from satellite measurements: A first step toward modeling the variability of oceanic whitecaps. *Journal of Geophysical Research: Oceans (1978–2012)*, 111(C3), 2006.

A. Ansmann and D. Müller. Lidar and atmospheric aerosol particles. In C. Weitkamp, editor, *Lidar*, volume 102 of *Springer Series in Optical Sciences*, pages 105–141. Springer New York, 2005. ISBN 978-0-387-40075-4. doi: 10.1007/0-387-25101-4_4. URL `http://dx.doi.org/10.1007/0-387-25101-4_4`.

P. Arason, G. Petersen, and H. Bjornsson. Observations of the altitude of the volcanic plume during the eruption of Eyjafjallajökull, April–May 2010. *Earth System Science Data Discussions*, 4(1):1–25, 2011.

P. Argall and R. Sica. {LIDAR} | atmospheric sounding introduction. In J. R. Holton, editor, *Encyclopedia of Atmospheric Sciences*, pages 1169 – 1176. Academic Press, Oxford, 2003. ISBN 978-0-12-227090-1. doi: http://dx.doi.org/10.1016/B0-12-227090-8/00203-7.

A. Asmi, M. Collaud Coen, J. Ogren, E. Andrews, P. Sheridan, A. Jefferson, E. Weingartner, U. Baltensperger, N. Bukowiecki, H. Lihavainen, et al. Aerosol decadal trends–part 2: In-situ aerosol particle number concentrations at GAW and ACTRIS stations. *Atmospheric Chemistry and Physics*, 13(2):895–916, 2013.

J. Barré, V.-H. Peuch, J.-L. Attié, L. E. Amraoui, W. Lahoz, B. Josse, M. Claeyman, and P. Nédélec. Stratosphere-troposphere ozone exchange from high resolution MLS ozone analyses. *Atmospheric Chemistry and Physics*, 12(14):6129–6144, 2012. doi: 10.5194/acp-12-6129-2012.

L. A. Barrie. Arctic air pollution: An overview of current knowledge. *Atmospheric Environment (1967)*, 20(4):643–663, 1986.

P. Bechtold, E. Bazile, F. Guichard, P. Mascart, and E. Richard. A mass-flux convection scheme for regional and global models. *Quarterly Journal of the Royal Meteorological Society*, 127(573):869–886, 2001.

H. Bencherif, L. E. Amraoui, N. Semane, S. Massart, D. V. Charyulu, A. Hauchecorne, and V.-H. Peuch. Examination of the 2002 major warming in the southern hemisphere using ground-based and Odin/SMR assimilated data: stratospheric ozone distributions and tropic/mid-latitude exchange. *Canadian Journal of Physics*, 85(11):1287–1300, 2007.

H. Bencherif, L. E. Amraoui, G. Kirgis, J. Leclair De Bellevue, A. Hauchecorne, N. Mzé, T. Portafaix, A. Pazmino, and F. Goutail. Analysis of a rapid increase of stratospheric ozone during late austral summer 2008 over Kerguelen (49.4° S, 70.3° E). *Atmospheric Chemistry and Physics*, 11(1):363–373, 2011.

A. Benedetti, J.-J. Morcrette, O. Boucher, A. Dethof, R. Engelen, M. Fisher, H. Flentje, N. Huneeus, L. Jones, J. Kaiser, et al. Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: 2. data assimilation. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 114(D13), 2009.

A. Bentamy, P. Queffeulou, Y. Quilfen, and K. Katsaros. Ocean surface wind fields estimated from satellite active and passive microwave instruments. *Geoscience and Remote Sensing, IEEE Transactions on*, 37(5):2469–2486, 1999.

S. Berthier, P. Chazette, P. Couvert, J. Pelon, F. Dulac, F. Thieuleux, C. Moulin, and T. Pain. Desert dust aerosol columnar properties over ocean and continental Africa from Lidar in-Space Technology Experiment (LITE) and Meteosat synergy. *Journal of Geophysical Research: Atmospheres*, 111(D21), 2006. doi: 10.1029/2005JD006999.

D. C. Blanchard. The production, distribution, and bacterial enrichment of the sea-salt aerosol. In *Air-sea exchange of gases and particles*, pages 407–454. Springer, 1983.

C. Böckmann, U. Wandinger, A. Ansmann, J. Bösenberg, V. Amiridis, A. Boselli, A. Delaval, F. De Tomasi, M. Frioud, I. V. Grigorov, et al. Aerosol lidar intercomparison in the framework of the EARLINET project. 2. aerosol backscatter algorithms. *Applied Optics*, 43(4):977–989, 2004.

K. Bogumil, J. Orphal, T. Homann, S. Voigt, P. Spietz, O. Fleischmann, A. Vogel, M. Hartmann, H. Kromminga, H. Bovensmann, et al. Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230–2380 nm region. *Journal of Photochemistry and Photobiology A: Chemistry*, 157(2):167–184, 2003.

C. F. Bohren. *Clouds in a glass of beer: simple experiments in atmospheric physics*. Courier Dover Publications, 2001.

T. C. Bond, D. G. Streets, K. F. Yarber, S. M. Nelson, J.-H. Woo, and Z. Klimont. A technology-based global inventory of black and organic carbon emissions from combustion. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 109(D14), 2004.

T. C. Bond, E. Bhardwaj, R. Dong, R. Jogani, S. Jung, C. Roden, D. G. Streets, and N. M. Trautmann. Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850–2000. *Global Biogeochemical Cycles*, 21(2):GB2018, 2007. doi: 10.1029/2006GB002840.

T. C. Bond, S. J. Doherty, D. W. Fahey, P. M. Forster, T. Berntsen, B. J. DeAngelo, M. G. Flanner, S. Ghan, B. Kärcher, D. Koch, S. Kinne, Y. Kondo, P. K. Quinn, M. C. Sarofim, M. G. Schultz, M. Schulz, C. Venkataraman, H. Zhang, S. Zhang, N. Bellouin, S. K. Guttikunda, P. K. Hopke, M. Z. Jacobson, J. W. Kaiser, Z. Klimont, U. Lohmann, J. P. Schwarz, D. Shindell, T. Storelvmo, S. G. Warren, and C. S. Zender. Bounding the role of black carbon in the climate system: A scientific assessment. *Journal of Geophysical Research: Atmospheres*, 118(11):5380–5552, 2013. doi: 10.1002/jgrd.50171.

O. Boucher. *Les aérosols atmosphériques*. Springer, 2012.

C. Bouet, G. Cautenet, G. Bergametti, B. Marticorena, M. C. Todd, and R. Washington. Sensitivity of desert dust emissions to model horizontal grid spacing during the Bodélé Dust Experiment 2005. *Atmospheric Environment*, 50:377–380, 2012.

M. A. Bourassa, D. M. Legler, J. J. O'Brien, and S. R. Smith. SeaWinds validation with research vessels. *Journal of Geophysical Research: Oceans (1978–2012)*, 108(C2), 2003.

N. Bousserez, J. Attié, V. Peuch, M. Michou, G. Pfister, D. Edwards, L. Emmons, C. Mari, B. Barret, S. Arnold, A. Heckel, A. Richter, H. Schlager, A. Lewis, M. Avery, G. Sachse, E. V. Browell, and J. W. Hair. Evaluation of the mocage chemistry transport model during the ICARTT/ITOP experiment. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 112(D10):D10S42, 2007. doi: 10.1029/2006JD007595.

H. Bovensmann, J. Burrows, M. Buchwitz, J. Frerick, S. Noël, V. Rozanov, K. Chance, and A. Goede. SCIAMACHY: Mission objectives and measurement modes. *Journal of the Atmospheric Sciences*, 56(2):127–150, 1999.

J. W. Boylan and A. G. Russell. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. *Atmospheric Environment*, 40(26): 4946–4959, 2006.

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

C. S. Bristow, K. A. Hudson-Edwards, and A. Chappell. Fertilizing the Amazon and equatorial Atlantic with West African dust. *Geophysical Research Letters*, 37(14), 2010.

D. Brooks. Monitoring solar radiation and its transmission through the atmosphere. *Drexel Univ., USA*, 2006. URL `http://www.pages.drexel.edu/~brooksdr/DRB_web_page/papers/UsingTheSun/using.htm`.

R. A. Brost, J. Feichter, and M. Heimann. Three-dimensional simulation of ^{7}Be in a global climate model. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 96(D12): 22423–22445, 1991.

P. P. Brown and D. F. Lawler. Sphere drag and settling velocity revisited. *Journal of Environmental Engineering*, 129(3):222–231, 2003. doi: 10.1061/(ASCE)0733-9372(2003)129:3(222).

S. Buis, A. Piacentini, and D. Déclat. PALM: a computational framework for assembling high-performance computing applications. *Concurrency and Computation: Practice and experience*, 18(2):231–245, 2006.

S. Burton, R. Ferrare, M. Vaughan, A. Omar, R. Rogers, C. Hostetler, and J. Hair. Aerosol classification from airborne HSRL and comparisons with the CALIPSO vertical feature mask. *Atmospheric Measurement Techniques*, 6(5):1397–1412, 2013.

Y. Cai, J. Snider, and P. Wechsler. Calibration of the passive cavity aerosol spectrometer probe for airborne determination of the size distribution. *Atmospheric Measurement Techniques Discussions*, 6(3):4123–4152, 2013.

Y. Callot, B. Marticorena, and G. Bergametti. Geomorphologic approach for modelling the surface features of arid environments in a model of dust emissions: application to the Sahara desert. *Geodinamica Acta*, 13(5):245–270, 2000.

S. Carey and R. Sparks. Quantitative models of the fallout and dispersal of tephra from volcanic eruption columns. *Bulletin of Volcanology*, 48(2-3):109–125, 1986.

D. Carrer, J.-L. Roujean, O. Hautecoeur, and T. Elias. Daily estimates of aerosol optical thickness over land surface based on a directional and temporal analysis of SEVIRI MSG visible observations. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 115 (D10):D10208, 2010. doi: 10.1029/2009JD012272.

R. Chakrabarty, H. Moosmüller, L.-W. Chen, K. Lewis, W. Arnott, C. Mazzoleni, M. Dubey, C. Wold, W. Hao, and S. Kreidenweis. Brown carbon in tar balls from smoldering biomass combustion. *Atmospheric Chemistry and Physics*, 10(13):6363–6370, 2010.

W. Chameides and A. Stelson. Aqueous-phase chemical processes in deliquescent sea-salt aerosols: A mechanism that couples the atmospheric cycles of S and sea salt. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 97(D18):20565–20580, 1992.

D. Chate. Study of scavenging of submicron-sized aerosol particles by thunderstorm rain events. *Atmospheric Environment*, 39(35):6608–6619, 2005.

P. Chazette, C. David, J. Lefrere, S. Godin, J. Pelon, and G. Mégie. Comparative lidar study of the optical, geometrical, and dynamical properties of stratospheric post-volcanic aerosols, following the eruptions of el chichon and mount pinatubo. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 100(D11):23195–23207, 1995.

P. Chazette, M. Bocquet, P. Royer, V. Winiarek, J.-C. Raut, P. Labazuy, M. Gouhier, M. Lardier, and J.-P. Cariou. Eyjafjallajökull ash concentrations derived from both lidar and modeling. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 117(D20), 2012a.

P. Chazette, A. Dabas, J. Sanak, M. Lardier, and P. Royer. French airborne lidar measurements for Eyjafjallajökull ash plume survey. *Atmospheric Chemistry and Physics*, 12(15): 7059–7072, 2012b.

D. B. Chelton. The impact of SST specification on ECMWF surface wind stress fields in the eastern tropical Pacific. *Journal of climate*, 18(4):530–550, 2005.

D. B. Chelton and M. H. Freilich. Scatterometer-based assessment of 10-m wind analyses from the operational ECMWF and NCEP numerical weather prediction models. *Monthly Weather Review*, 133(2):409–429, 2005.

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

M. Claeyman, J. Attié, L. El Amraoui, D. Cariolle, V. Peuch, H. Teyssèdre, B. Josse, P. Ricaud, S. Massart, A. Piacentini, et al. A linear CO chemistry parameterization in a chemistry-transport model: evaluation and application to data assimilation. *Atmos. Chem. Phys*, 10:6097–6115, 2010.

M. Collaud Coen, E. Andrews, A. Asmi, U. Baltensperger, N. Bukowiecki, D. Day, M. Fiebig, A. M. Fjaeraa, H. Flentje, A. Hyvärinen, A. Jefferson, S. G. Jennings, G. Kouvarakis, H. Lihavainen, C. Lund Myhre, W. C. Malm, N. Mihapopoulos, J. V. Molenar, C. O'Dowd, J. A. Ogren, B. A. Schichtel, P. Sheridan, A. Virkkula, E. Weingartner, R. Weller, and P. Laj. Aerosol decadal trends – part 1: In-situ optical measurements at gaw and improve stations. *Atmospheric Chemistry and Physics*, 13(2):869–894, 2013. doi: 10.5194/acp-13-869-2013.

R. Collis and P. Russell. Lidar measurement of particles and gases by elastic backscattering and differential absorption. In *Laser monitoring of the atmosphere*, pages 71–151. Springer, 1976.

P. Courtier, J.-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519):1367–1387, 1994.

P. Courtier, E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher, and J. Pailleux. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1783–1807, 1998.

B. Croft, U. Lohmann, R. Martin, P. Stier, S. Wurzler, J. Feichter, R. Posselt, and S. Ferrachat. Aerosol size-dependent below-cloud scavenging by rain and snow in the ECHAM5-HAM. *Atmospheric Chemistry and Physics*, 9(14):4653–4675, 2009. doi: 10.5194/acp-9-4653-2009.

S. Crumeyrolle, P. Tulet, L. Gomes, L. Garcia-Carreras, C. Flamant, D. J. Parker, A. Matsuki, P. Formenti, and A. Schwarzenboeck. Transport of dust particles from the bodélé region to the monsoon layer — AMMA case study of the 9–14 June 2006 period. *Atmospheric Chemistry and Physics*, 11(2):479–494, 2011. doi: 10.5194/acp-11-479-2011.

N. Daget. *Estimation d'ensemble des paramètres des covariances d'erreur d'ébauche dans un système d'assimilation variationnelle de données océaniques.* PhD thesis, Université de Toulouse, Université Toulouse III-Paul Sabatier, 2008.

R. Daley. *Atmospheric data analysis.* Number 2 in Atmospheric and Space Science Series. Cambridge university press, 1993.

H. M. Davenport and L. Peters. Field studies of atmospheric particulate concentration changes during precipitation. *Atmospheric Environment (1967)*, 12(5):997–1008, 1978.

G. De Leeuw, S. Kinne, J.-F. Léon, J. Pelon, D. Rosenfeld, M. Schaap, P. J. Veefkind, B. Veihelmann, D. M. Winker, and W. von Hoyningen-Huene. Retrieval of aerosol properties. In *The remote sensing of tropospheric composition from space*, pages 259–313. Springer, 2011.

A. De Meij, A. Pozzer, and J. Lelieveld. Trend analysis in aerosol optical depths and pollutant emission estimates between 2000 and 2009. *Atmospheric Environment*, 51:75–85, 2012.

D. Dee and S. Uppala. *Variational bias correction in ERA-Interim.* ECMWF, 2008.

R. Delmas, G. Mégie, and V.-H. Peuch. *Physique et chimie de l'atmosphère.* Belin, 2005.

P. J. DeMott, K. Sassen, M. R. Poellot, D. Baumgardner, D. C. Rogers, S. D. Brooks, A. J. Prenni, and S. M. Kreidenweis. African dust aerosols as atmospheric ice nuclei. *Geophysical Research Letters*, 30(14), 2003.

F. Dentener, S. Kinne, T. Bond, O. Boucher, J. Cofala, S. Generoso, P. Ginoux, S. Gong, J. Hoelzemann, A. Ito, et al. Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom. *Atmospheric Chemistry and Physics*, 6 (12):4321–4344, 2006.

G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3385–3396, 2005.

O. Dubovik, B. Holben, T. F. Eck, A. Smirnov, Y. J. Kaufman, M. D. King, D. Tanré, and I. Slutsker. Variability of absorption and optical properties of key aerosol types observed in worldwide locations. *Journal of the Atmospheric Sciences*, 59(3), 2002.

R. A. Duce. On the source of gaseous chlorine in the marine atmosphere. *Journal of Geophysical Research*, 74(18):4597–4599, 1969.

R. A. Duce, W. Stumm, and J. M. Prospero. Working symposium on sea-air chemistry: Summary and recommendations. *Journal of Geophysical Research*, 77(27):5059–5061, 1972. doi: 10.1029/JC077i027p05059.

A. Dufour, M. Amodei, G. Ancellet, and V.-H. Peuch. Observed and modelled "chemical weather" during ESCOMPTE. *Atmospheric research*, 74(1):161–189, 2005.

F. Dulac and P. Chazette. Airborne study of a multi-layer aerosol structure in the eastern mediterranean observed with the airborne polarized lidar ALEX during a STAAARTE campaign (7 june 1997). *Atmospheric Chemistry and Physics*, 3(5):1817–1831, 2003.

T. Duncan Fairlie, D. J. Jacob, and R. J. Park. The impact of transpacific transport of mineral dust in the united states. *Atmospheric Environment*, 41(6):1251–1266, 2007.

L. B. Ebert. Is soot composed predominantly of carbon clusters? *Science*, 247(4949): 1468–1471, 1990.

S. Eckermann. Hybrid σ-p coordinate choices for a global model. *Monthly Weather Review*, 137(1):224–245, 2009.

S. Eckhardt, A. Prata, P. Seibert, K. Stebel, and A. Stohl. Estimation of the vertical profile of sulfur dioxide injection into the atmosphere by a volcanic eruption using satellite column measurements and inverse transport modeling. *Atmospheric Chemistry and Physics*, 8 (14):3881–3897, 2008.

L. El Amraoui, V.-H. Peuch, P. Ricaud, S. Massart, N. Semane, H. Teyssedre, D. Cariolle, and F. Karcher. Ozone loss in the 2002–2003 Arctic vortex deduced from the assimilation of Odin/SMR O_3 and N_2O measurements: N_2O as a dynamical tracer. *Quarterly Journal of the Royal Meteorological Society*, 134(630):217–228, 2008a.

L. El Amraoui, N. Semane, V.-H. Peuch, and M. Santee. Investigation of dynamical processes in the polar stratospheric vortex during the unusually cold winter 2004/2005. *Geophysical Research Letters*, 35(3), 2008b.

L. El Amraoui, J.-L. Attié, N. Semane, M. Claeyman, V.-H. Peuch, J. Warner, P. Ricaud, J.-P. Cammas, A. Piacentini, B. Josse, et al. Midlatitude stratosphere–troposphere exchange as diagnosed by MLS O_3 and MOPITT CO assimilated fields. *Atmospheric Chemistry and Physics*, 10(5):2175–2194, 2010. doi: 10.5194/acp-10-2175-2010.

R. G. Ellingson and J. C. Gille. An infrared radiative transfer model. part 1: Model description and comparison of observations with calculations. *Journal of the Atmospheric Sciences*, 35(3):523–545, 1978.

E. Emili, B. Barret, S. Massart, E. Le Flochmoen, A. Piacentini, L. El Amraoui, O. Pannekoucke, and D. Cariolle. Combined assimilation of IASI and MLS observations to constrain tropospheric and stratospheric ozone in a global chemical transport model. *Atmospheric Chemistry and Physics*, 14(1):177–198, 2014.

E. Eriksson. The yearly circulation of chloride and sulfur in nature; meteorological, geochemical and pedological implications. part 1. *Tellus*, 11(4):375–403, 1959.

T. Fan and O. Toon. Modeling sea-salt aerosol in a coupled climate and sectional microphysical model: mass, optical depth and number concentration. *Atmospheric Chemistry and Physics*, 11(9):4587–4610, 2011. doi: 10.5194/acp-11-4587-2011.

G. Feingold, W. R. Cotton, S. M. Kreidenweis, and J. T. Davis. The impact of giant cloud condensation nuclei on drizzle formation in stratocumulus: Implications for cloud radiative properties. *Journal of the atmospheric sciences*, 56(24):4100–4117, 1999.

P. Field, O. Möhler, P. Connolly, M. Krämer, R. Cotton, A. Heymsfield, H. Saathoff, and M. Schnaiter. Some ice nucleation characteristics of Asian and Saharan desert dust. *Atmospheric Chemistry and Physics*, 6(10):2991–3006, 2006.

M. Fisher and E. Andersson. *Developments in 4D-Var and Kalman filtering*. European Centre for Medium-Range Weather Forecasts, 2001.

J. W. Fitzgerald. Marine aerosols: A review. *Atmospheric Environment. Part A. General Topics*, 25(3):533–545, 1991.

P. H. Flamant, L. Vincent, J. Cuesta, A. Dabas, M. Olivier, and D. Huber. ADM-Aeolus L2A algorithm theoretical baseline document – particle spin-off products. Technical Report AE-TN-IPSL-GS-001, ESA, 2013.

J. Flemming and A. Inness. Volcanic sulfur dioxide plume forecasts based on uv satellite retrievals for the 2011 Grímsvötn and the 2010 Eyjafjallajökull eruption. *Journal of Geophysical Research: Atmospheres*, 118(17):10–172, 2013.

L. G. Franzen, M. Hjelmroos, P. Kållberg, A. Rapp, J. O. Mattsson, and E. Brorström-Lundén. The Saharan dust episode of south and central Europe, and northern Scandinavia, march 1991. *Weather*, 50(9):313–318, 1995.

K. Fung. Particulate carbon speciation by mno_2 oxidation. *Aerosol science and technology*, 12(1):122–127, 1990.

G. Gangale, A. J. Prata, and L. Clarisse. The infrared spectral signature of volcanic ash determined from high-spectral resolution satellite measurements. *Remote sensing of environment*, 114(2):414–425, 2010.

B. Gary. Pluto occultation of 2007.03.18, 2007. URL `http://brucegary.net/PlutoOcc/x.htm`.

H. E. Gerber. Relative-humidity parameterization of the Navy Aerosol Model (NAM). Technical Report NRL Report 8956, Naval Research Laboratory, 1985.

F. Giorgi and W. L. Chameides. Rainout lifetimes of highly soluble aerosols and gases as inferred from simulations with a general circulation model. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 91(D13):14367–14376, 1986.

A. H. Goldstein and I. E. Galbally. Known and unexplored organic constituents in the earth's atmosphere. *Environmental Science & Technology*, 41(5):1514–1521, 2007.

L. Gomes and D. A. Gillette. A comparison of characteristics of aerosol from dust storms in central Asia with soil-derived dust from other regions. *Atmospheric Environment. Part A. General Topics*, 27(16):2539–2544, 1993.

S. Gong. A parameterization of sea-salt aerosol source function for sub-and super-micron particles. *Global Biogeochemical Cycles*, 17(4):1097, 2003. doi: 10.1029/2003GB002079.

S. Gong, L. Barrie, and J.-P. Blanchet. Modeling sea-salt aerosols in the atmosphere: 1. model development. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 102 (D3):3805–3818, 1997.

R. Goody. A statistical model for water-vapour absorption. *Quarterly Journal of the Royal Meteorological Society*, 78(336):165–169, 1952.

A. Goudie and N. Middleton. Saharan dust storms: nature and consequences. *Earth-Science Reviews*, 56(1):179–204, 2001.

J. Gras. Climatology of tropospheric aerosols. In J. R. Holton, editor, *Encyclopedia of Atmospheric Sciences*, pages 13 – 20. Academic Press, Oxford, 2003. ISBN 978-0-12-227090-1. doi: 10.1016/B0-12-227090-8/00051-8.

M. Griggs. Measurements of atmospheric aerosol optical thickness over water using ERS-1 data. *J. Air Pollut. Conlrol. Assoc.*, 25:622–626, 1975.

A. Guenther, C. N. Hewitt, D. Erickson, R. Fall, C. Geron, T. Graedel, P. Harley, L. Klinger, M. Lerdau, W. McKay, et al. A global model of natural volatile organic compound emissions. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 100(D5):8873–8892, 1995.

C. Guieu, M.-D. Loÿe-Pilot, C. Ridame, and C. Thomas. Chemical characterization of the Saharan dust end-member: Some biogeochemical implications for the western Mediterranean Sea. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 107(D15):ACH–5, 2002.

M. Hallquist, J. C. Wenger, U. Baltensperger, Y. Rudich, D. Simpson, M. Claeys, J. Dommen, N. M. Donahue, C. George, A. H. Goldstein, J. F. Hamilton, H. Herrmann, T. Hoffmann, Y. Iinuma, M. Jang, M. E. Jenkin, J. L. Jimenez, A. Kiendler-Scharr, W. Maenhaut, G. McFiggans, T. F. Mentel, A. Monod, A. S. H. Prévôt, J. H. Seinfeld, J. D. Surratt, R. Szmigielski, and J. Wildt. The formation, properties and impact of secondary organic aerosol: current and emerging issues. *Atmospheric Chemistry and Physics*, 9(14):5155–5236, 2009. doi: 10.5194/acp-9-5155-2009.

J. L. Hatfield and T. J. Sauer. *Soil management: building a stable base for agriculture*. American Society of Agronomy, 2011.

B. Heese and M. Wiegner. Vertical aerosol profiles from raman polarization lidar observations during the dry season AMMA field campaign. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D23), 2008.

W. C. Hinds. *Aerosol technology: properties, behavior, and measurement of airborne particles*. John Wiley & Sons, 2nd edition, 1999.

P. V. Hobbs, L. F. Radke, J. H. Lyons, R. J. Ferek, D. J. Coffman, and T. J. Casadevall. Airborne measurements of particle and gas emissions from the 1990 volcanic eruptions of Mount Redoubt. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 96(D10): 18735–18752, 1991.

A. Hoffer, A. Gelencsér, P. Guyon, G. Kiss, O. Schmid, G. Frank, P. Artaxo, and M. Andreae. Optical properties of humic-like substances (HULIS) in biomass-burning aerosols. *Atmospheric Chemistry and Physics*, 6(11):3563–3570, 2006.

B. Holben, T. Eck, I. Slutsker, D. Tanre, J. Buis, A. Setzer, E. Vermote, J. Reagan, Y. Kaufman, T. Nakajima, et al. AERONET—a federated instrument network and data archive for aerosol characterization. *Remote sensing of environment*, 66(1):1–16, 1998.

C. A. Hostetler, Z. Liu, J. Reagan, M. Vaughan, D. Winker, M. Osborn, W. Hunt, K. Powell, and C. Trepte. CALIOP algorithm theoretical basis document–part 1: Calibration and level 1 data products. *Doc. PC-SCI*, 201, 2006.

N. C. Hsu, S.-C. Tsay, M. D. King, and J. R. Herman. Aerosol properties over bright-reflecting source regions. *Geoscience and Remote Sensing, IEEE Transactions on*, 42(3):557–569, 2004.

N. C. Hsu, S.-C. Tsay, M. D. King, and J. R. Herman. Deep Blue retrievals of Asian aerosol properties during ACE-Asia. *Geoscience and Remote Sensing, IEEE Transactions on*, 44 (11):3180–3195, 2006.

L.-M. w. http://lance modis.eosdis.nasa.gov. Comparison of standard and NRT products. URL `http://lance-modis.eosdis.nasa.gov/data_products/comparison.html`.

V. Huijnen and H. Eskes. Skill scores and evaluation methodology for the MACC II project. Technical report, 2012. Available at `www.gmes-atmosphere.eu/documents/maccii/deliverables/val/MACCII_VAL_DEL_D_85.2_ScoringReport01_20120222.pdf` (last access: 15 October 2014).

C. Ichoku, L. A. Remer, and T. F. Eck. Quantitative evaluation and intercomparison of morning and afternoon Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol measurements from terra and aqua. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 110(D10):D10S99, 2005. doi: 10.1029/2004JD004987.

IPCC. *Climate change 2001: The Scientific Basis: Contribution of Working Group I to the third assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and New York, 2001. ISBN 0521 80767 0.

IPCC. *Climate change 2007: The physical science basis: Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and New York, 2007. ISBN 978-0-521-88009-1.

IPCC. *Climate change 2013: The physical science basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and New York, 2013. ISBN 978-1-107-05799-1.

D. Jacob, H. Liu, C. Mari, and R. Yantosca. Harvard wet deposition scheme for GMI. Technical report, Harvard University Atmospheric Chemistry Modeling Group, 2000. Available at `gmi.gsfc.nasa.gov/models/jacob_wetdep.pdf` (last access: 1 December 2013).

M. Z. Jacobson. A physically-based treatment of elemental carbon optics: Implications for global direct forcing of aerosols. *Geophysical Research Letters*, 27(2):217–220, 2000.

M. Z. Jacobson. Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols. *Nature*, 409(6821):695–697, 2001.

M. Z. Jacobson. *Fundamentals of atmospheric modeling*. Cambridge University Press, 2005.

L. Jaeglé, P. Quinn, T. Bates, B. Alexander, and J.-T. Lin. Global distribution of sea salt aerosols: new constraints from in situ and remote sensing observations. *Atmospheric Chemistry and Physics*, 11(7):3137–3157, 2011.

R. Jaenicke. Tropospheric aerosols. In *Aerosol-cloud-climate interactions*, pages 1–31. Academic Press, Inc, 1993.

E. Jaumouillé, S. Massart, A. Piacentini, D. Cariolle, and V.-H. Peuch. Impact of a time-dependent background error covariance matrix on air quality analysis. *Geoscientific Model Development*, 5(5):1075–1090, 2012.

B. Josse, P. Simon, and V.-H. Peuch. Radon global simulations with the multiscale chemistry and transport model mocage. *Tellus B*, 56(4):339–356, 2004.

C. Junker and C. Liousse. A global emission inventory of carbonaceous aerosol from historic records of fossil fuel and biofuel consumption for the period 1860–1997. *Atmospheric Chemistry and Physics*, 8(5):1195–1207, 2008. doi: 10.5194/acp-8-1195-2008.

R. A. Kahn, B. J. Gaitley, J. V. Martonchik, D. J. Diner, K. A. Crean, and B. Holben. 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 (1984–2012)*, 110(D10), 2005.

J. Kaiser, A. Heil, M. Andreae, A. Benedetti, N. Chubarova, L. Jones, J.-J. Morcrette, M. Razinger, M. Schultz, M. Suttie, et al. Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, 9 (1):527–554, 2012.

E. Kalnay. *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press, 2003.

V. A. Karydis, P. Kumar, D. Barahona, I. N. Sokolik, and A. Nenes. On the effect of dust particles on global cloud condensation nuclei and cloud droplet number. *Journal of Geophysical Research: Atmospheres*, 116(D23), 2011. doi: 10.1029/2011JD016283.

Y. Kaufman and D. Tanré. Satellite remote sensing | aerosol measurements. In J. R. Holton, editor, *Encyclopedia of Atmospheric Sciences*, pages 1941 – 1956. Academic Press, Oxford, 2003. ISBN 978-0-12-227090-1. doi: http://dx.doi.org/10.1016/B0-12-227090-8/00347-X.

Y. J. Kaufman, D. Tanré, and O. Boucher. A satellite view of aerosols in the climate system. *Nature*, 419(6903):215–223, 2002.

W. C. Keene, R. Sander, A. A. Pszenny, R. Vogt, P. J. Crutzen, and J. N. Galloway. Aerosol pH in the marine boundary layer: A review and model evaluation. *Journal of Aerosol Science*, 29(3):339–356, 1998.

E. Kessler. On the distribution and continuity of water substance in atmospheric circulation. *Met. Monograph*, 10:84, 1969. doi: http://dx.doi.org/10.1016/0169-8095(94)00090-Z.

M. D. King, Y. J. Kaufman, D. Tanré, and T. Nakajima. Remote sensing of tropospheric aerosols from space: Past, present, and future. *Bulletin of the American Meteorological Society*, 80(11):2229–2259, 1999.

T. W. Kirchstetter, T. Novakov, and P. V. Hobbs. Evidence that the spectral dependence of light absorption by aerosols is affected by organic carbon. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 109(D21), 2004.

J. D. Klett. Stable analytical inversion solution for processing lidar returns. *Applied Optics*, 20(2):211–220, 1981.

J. F. Kok, E. J. Parteli, T. I. Michaels, and D. B. Karam. The physics of wind-blown sand and dust. *Reports on Progress in Physics*, 75(10):106901, 2012.

A. Kokhanovsky, J. Deuzé, D. Diner, O. Dubovik, F. Ducos, C. Emde, M. Garay, R. Grainger, A. Heckel, M. Herman, et al. The inter-comparison of major satellite aerosol retrieval algorithms using simulated intensity and polarization characteristics of reflected light. *Atmospheric Measurement Techniques*, 3(4):909–932, 2010.

Y. Kondo, N. Oshima, M. Kajino, R. Mikami, N. Moteki, N. Takegawa, R. L. Verma, Y. Kajii, S. Kato, and A. Takami. Emissions of black carbon in East Asia estimated from observations at a remote site in the East China Sea. *Journal of Geophysical Research: Atmospheres*, 116(D16), 2011. doi: 10.1029/2011JD015637.

P. Köpke, M. Hess, I. Schult, and E. Shettle. Global aerosol data set. Technical Report MPI Report No 243, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 1997.

I. Koren, L. A. Remer, Y. J. Kaufman, Y. Rudich, and J. V. Martins. On the twilight zone between clouds and aerosols. *Geophysical Research Letters*, 34(8):L08805, 2007. doi: 10.1029/2007gl029253.

S. Y. Kotchenova, E. F. Vermote, R. Matarrese, F. J. Klemm Jr, et al. Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. part i: Path radiance. *Applied optics*, 45(26):6762–6774, 2006.

N. Kristiansen, A. Stohl, A. Prata, A. Richter, S. Eckhardt, P. Seibert, A. Hoffmann, C. Ritter, L. Bitar, T. Duck, et al. Remote sensing and inverse transport modeling of the Kasatochi eruption sulfur dioxide cloud. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 115(D2), 2010.

N. Kristiansen, A. Stohl, A. Prata, N. Bukowiecki, H. Dacre, S. Eckhardt, S. Henne, M. Hort, B. Johnson, F. Marenco, et al. Performance assessment of a volcanic ash transport model mini-ensemble used for inverse modeling of the 2010 Eyjafjallajökull eruption. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 117(D20), 2012.

B. J. Krueger, V. H. Grassian, A. Laskin, and J. P. Cowin. The transformation of solid atmospheric particles into liquid droplets through heterogeneous chemistry: Laboratory insights into the processing of calcium containing mineral dust aerosol in the troposphere. *Geophysical Research Letters*, 30(3), 2003.

L. Laakso, T. Grönholm, Ü. Rannik, M. Kosmale, V. Fiedler, H. Vehkamäki, and M. Kulmala. Ultrafine particle scavenging coefficients calculated from 6 years field measurements. *Atmospheric Environment*, 37(25):3605–3613, 2003.

C. Lacasse, S. Karlsdóttir, G. Larsen, H. Soosalu, W. Rose, and G. Ernst. Weather radar observations of the hekla 2000 eruption cloud, iceland. *Bulletin of Volcanology*, 66(5): 457–473, 2004.

D. Lack and C. Cappa. Impact of brown and clear carbon on light absorption enhancement, single scatter albedo and absorption wavelength dependence of black carbon. *Atmospheric Chemistry and Physics*, 10(9):4207–4220, 2010.

W. Lahoz, V.-H. Peuch, and J.-L. Attié. *POGEQA White Book. Monitoring, forecasting and managing air quality from space: Part 2. Perspectives for data assimilation.* POGEQA Team, 2011.

W. Lahoz, V.-H. Peuch, J. Orphal, J.-L. Attié, K. Chance, X. Liu, D. Edwards, H. Elbern, J.-M. Flaud, M. Claeyman, et al. Monitoring air quality from space: the case for the geostationary platform. *Bulletin of the American Meteorological Society*, 93(2):221–233, 2012.

J.-F. Lamarque, T. C. Bond, V. Eyring, C. Granier, A. Heil, Z. Klimont, D. Lee, C. Liousse, A. Mieville, B. Owen, et al. Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmospheric Chemistry and Physics*, 10(15):7017–7039, 2010. doi: 10.5194/acp-10-7017-2010.

J.-F. Lamarque, D. T. Shindell, B. Josse, P. Young, I. Cionni, V. Eyring, D. Bergmann, P. Cameron-Smith, W. J. Collins, R. Doherty, et al. The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics. *Geoscientific Model Development*, 6(1), 2013. doi: 10.5194/gmd-6-179-2013.

A. C. Lange and H. Elbern. Lidar data assimilation for improved analyses of volcanic aerosol events. In *EGU General Assembly Conference Abstracts*, volume 16, page 5987, 2014.

B. Langmann, A. Folch, M. Hensch, and V. Matthias. Volcanic ash over Europe during the eruption of Eyjafjallajökull on Iceland, April–May 2010. *Atmospheric Environment*, 48: 1–8, 2012.

J. Langner and H. Rodhe. A global three-dimensional model of the tropospheric sulfur cycle. *Journal of Atmospheric Chemistry*, 13(3):225–263, 1991. doi: 10.1007/BF00058134.

B. Laurent, B. Marticorena, G. Bergametti, and F. Mei. Modeling mineral dust emissions from Chinese and Mongolian deserts. *Global and planetary Change*, 52(1):121–141, 2006.

B. Laurent, B. Heinold, I. Tegen, C. Bouet, and G. Cautenet. Surface wind accuracy for modeling mineral dust emissions: Comparing two regional models in a Bodélé case study. *Geophysical Research Letters*, 35(9):L09804, 2008a. doi: 10.1029/2008GL033654.

B. Laurent, B. Marticorena, G. Bergametti, J. Léon, and N. Mahowald. Modeling mineral dust emissions from the sahara desert using new surface properties and soil database. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D14):D14218, 2008b. doi: 10.1029/2007JD009484.

K. H. Lee, Z. Li, Y. J. Kim, and A. Kokhanovsky. Atmospheric aerosol monitoring from satellite observations: a history of three decades. In *Atmospheric and biological environmental monitoring*, pages 13–38. Springer, 2009.

L. Lee, K. Carslaw, K. Pringle, G. Mann, and D. Spracklen. Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters. *Atmospheric Chemistry and Physics*, 11(23):12253–12273, 2011.

Z. Levin, E. Ganor, and V. Gladstein. The effects of desert particles coated with sulfate on rain formation in the eastern Mediterranean. *Journal of Applied Meteorology*, 35(9): 1511–1523, 1996.

Z. Levin, A. Teller, E. Ganor, and Y. Yin. On the interactions of mineral dust, sea-salt particles, and clouds: A measurement and modeling study from the Mediterranean Israeli Dust Experiment campaign. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 110(D20), 2005.

R. Lewis and E. Schwartz. *Sea salt aerosol production: mechanisms, methods, measurements and models – a critical review*, volume 152. American Geophysical Union, 2004a.

R. Lewis and E. Schwartz. *Sea salt aerosol production: mechanisms, methods, measurements and models — a critical review*, volume 152. American Geophysical Union, 2004b.

J. Li, M. Pósfai, P. V. Hobbs, and P. R. Buseck. Individual aerosol particles from biomass burning in southern Africa: 2, compositions and aging of inorganic particles. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 108(D13), 2003.

Z. Li, X. Zhao, R. Kahn, M. Mishchenko, L. Remer, K.-H. Lee, M. Wang, I. Laszlo, T. Nakajima, and H. Maring. Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: a review and perspective. *Annales Geophysicae*, 27(7):2755–2770, 2009.

H. Liu, D. J. Jacob, I. Bey, and R. M. Yantosca. Constraints from ^{210}Pb and ^{7}Be on wet deposition and transport in a global three-dimensional chemical tracer model driven by assimilated meteorological fields. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 106(D11):12109–12128, 2001.

Z. Liu, Q. Liu, H.-C. Lin, C. S. Schwartz, Y.-H. Lee, and T. Wang. Three-dimensional variational assimilation of MODIS aerosol optical depth: Implementation and application to a dust storm over East Asia. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 116(D23), 2011.

J.-F. Louis. A parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteorology*, 17(2):187–202, 1979.

T. Lurton, J.-B. Renard, D. Vignelles, M. Jeannot, R. Akiki, J.-L. Mineau, and T. Tonnelier. Light scattering at small angles by atmospheric irregular particles: modelling and laboratory measurements. *Atmospheric Measurement Techniques*, 7(4):931–939, 2014.

N. Mahowald, C. Luo, J. del Corral, and C. S. Zender. Interannual variability in atmospheric mineral aerosols from a 22-year model simulation and observational data. *Journal of Geophysical Research: Atmospheres*, 108(D12):4352, 2003. doi: 10.1029/2002JD002821.

N. M. Mahowald, P. J. Rasch, and R. G. Prinn. Cumulus parameterizations in chemical transport models. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 100(D12): 26173–26189, 1995.

W. Malkmus. Random lorentz band model with exponential-tailed s^{-1} line-intensity distribution function. *JOSA*, 57(3):323–329, 1967.

M. Mallet, J. C. Roger, S. Despiau, J. P. Putaud, and O. Dubovik. A study of the mixing state of black carbon in urban zone. *Journal of Geophysical Research: Atmospheres*, 109 (D4), 2004. doi: 10.1029/2003JD003940.

G. Mann, K. Carslaw, C. Reddington, K. Pringle, M. Schulz, A. Asmi, D. Spracklen, D. Ridley, M. Woodhouse, L. Lee, et al. Intercomparison and evaluation of aerosol microphysical properties among AeroCom global models of a range of complexity. *Atmospheric Chemistry and Physics Discussions*, 13(11):30841–30928, 2013.

J. S. Marshall and W. M. K. Palmer. The distribution of raindrops with size. *Journal of meteorology*, 5(4):165–166, 1948.

M. Martet. *Introduction des différentes composantes de l'aérosol atmosphérique dans le modèle MOCAGE*. PhD thesis, University of Toulouse III, 2008.

M. Martet, V.-H. Peuch, B. Laurent, B. Marticorena, and G. Bergametti. Evaluation of long-range transport and deposition of desert dust with the CTM MOCAGE. *Tellus B*, 61(2):449–463, 2009.

B. Marticorena, G. Bergametti, B. Aumont, Y. Callot, C. N'doumé, and M. Legrand. Modeling the atmospheric dust cycle: 2. simulation of Saharan dust sources. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 102(D4):4387–4404, 1997.

S. Massart, C. Clerbaux, D. Cariolle, A. Piacentini, S. Turquety, and J. Hadji-Lazaro. First steps towards the assimilation of IASI ozone data into the MOCAGE-PALM system. *Atmospheric Chemistry and Physics*, 9(14):5073–5091, 2009.

S. Massart, B. Pajot, A. Piacentini, and O. Pannekoucke. On the merits of using a 3D-FGAT assimilation scheme with an outer loop for atmospheric situations governed by transport. *Monthly Weather Review*, 138(12):4509–4522, 2010.

S. Massart, A. Piacentini, and O. Pannekoucke. Importance of using ensemble estimated background error covariances for the quality of atmospheric ozone analyses. *Quarterly Journal of the Royal Meteorological Society*, 138(665):889–905, 2012.

L. Mastin, M. Guffanti, R. Servranckx, P. Webley, S. Barsotti, K. Dean, A. Durant, J. Ewert, A. Neri, W. Rose, et al. A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions. *Journal of Volcanology and Geothermal Research*, 186(1):10–21, 2009.

V. Matthais, V. Freudenthaler, A. Amodeo, I. Balin, D. Balis, J. Bösenberg, A. Chaikovsky, G. Chourdakis, A. Comeron, A. Delaval, et al. Aerosol lidar intercomparison in the framework of the EARLINET project. 1. instruments. *Applied Optics*, 43(4):961–976, 2004.

V. Matthias. The aerosol distribution in europe derived with the community multiscale air quality (cmaq) model: comparison to near surface in situ and sunphotometer measurements. *Atmospheric Chemistry and Physics*, 8(17):5077–5097, 2008. doi: 10.5194/acp-8-5077-2008. URL `http://www.atmos-chem-phys.net/8/5077/2008/`.

L. M. McInnes, D. S. Covert, P. K. Quinn, and M. S. Germani. Measurements of chloride depletion and sulfur enrichment in individual sea-salt particles collected from the remote marine boundary layer. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 99 (D4):8257–8268, 1994.

R. M. Measures. *Laser remote sensing: fundamentals and applications*. Krieger, 1992.

A. I. Medalia and D. Rivin. Particulate carbon and other components of soot and carbon black. *Carbon*, 20(6):481–492, 1982.

F. Mei, X. Zhang, H. Lu, Z. Shen, and Y. Wang. Characterization of MASDs of surface soils in north China and its influence on estimating dust emission. *Chinese Science Bulletin*, 49(20):2169–2176, 2004.

S. Melfi. Remote measurements of the atmosphere using raman scattering. *Applied Optics*, 11(7):1605–1610, 1972.

R. Ménard, S. E. Cohn, L.-P. Chang, and P. M. Lyster. Assimilation of stratospheric chemical tracer observations using a Kalman filter. part i: Formulation. *Monthly Weather Review*, 128(8):2654–2671, 2000.

S. Migliorini. On the equivalence between radiance and retrieval assimilation. *Monthly Weather Review*, 140(1):258–265, 2012.

S. Millington, R. Saunders, P. Francis, and H. Webster. Simulated volcanic ash imagery: A method to compare NAME ash concentration forecasts with SEVIRI imagery for the Eyjafjallajökull eruption in 2010. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 117(D20), 2012.

E. C. Monahan and I. G. O'Muircheartaigh. Whitecaps and the passive remote sensing of the ocean surface. *International Journal of Remote Sensing*, 7(5):627–642, 1986.

C. Moulin, C. E. Lambert, F. Dulac, and U. Dayan. Control of atmospheric export of dust from North Africa by the North Atlantic Oscillation. *Nature*, 387(6634):691–694, 1997.

M. Mozurkewich. Mechanisms for the release of halogens from sea-salt particles by free radical reactions. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 100(D7): 14199–14207, 1995.

M. Mu, J. T. Randerson, G. R. van der Werf, L. Giglio, P. Kasibhatla, D. Morton, G. J. Collatz, R. S. DeFries, E. J. Hyer, E. M. Prins, D. W. T. Griffith, D. Wunch, G. C. Toon, V. Sherlock, and P. O. Wennberg. Daily and 3-hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 116(D24):D24303, 2011. doi: 10.1029/2011JD016245.

T. Murayama, N. Sugimoto, I. Uno, K. Kinoshita, K. Aoki, N. Hagiwara, Z. Liu, I. Matsui, T. Sakai, T. Shibata, et al. Ground-based network observation of Asian dust events of April 1998 in east Asia. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 106 (D16):18345–18359, 2001.

D. Murphy, J. Anderson, P. Quinn, L. McInnes, F. Brechtel, S. Kreidenweis, A. Middlebrook, M. Posfai, D. Thomson, and P. Buseck. Influence of sea-salt on aerosol radiative properties in the Southern ocean marine boundary layer. *Nature*, 392(6671):62–65, 1998a.

D. Murphy, D. Thomson, A. Middlebrook, and M. Schein. In situ single-particle characterization at Cape Grim. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 103 (D13):16485–16491, 1998b.

E.-Y. Nho-Kim, M. Michou, and V.-H. Peuch. Parameterization of size-dependent particle dry deposition velocities for global modeling. *Atmospheric Environment*, 38(13):1933–1942, 2004.

T. Niu, S. Gong, G. Zhu, H. Liu, X. Hu, C. Zhou, and Y. Wang. Data assimilation of dust aerosol observations for the CUACE/dust forecasting system. *Atmospheric Chemistry and Physics*, 8(13):3473–3482, 2008.

C. D. O'Dowd and M. H. Smith. Physicochemical properties of aerosols over the northeast Atlantic: Evidence for wind-speed-related submicron sea-salt aerosol production. *Journal of Geophysical Research: Atmospheres*, 98(D1):1137–1149, 1993. doi: 10.1029/92JD02302.

C. D. O'Dowd, M. H. Smith, I. E. Consterdine, and J. A. Lowe. Marine aerosol, sea-salt, and the marine sulphur cycle: A short review. *Atmospheric Environment*, 31(1):73–80, 1997.

C. D. O'Dowd, M. C. Facchini, F. Cavalli, D. Ceburnis, M. Mircea, S. Decesari, S. Fuzzi, Y. J. Yoon, and J.-P. Putaud. Biogenically driven organic contribution to marine aerosol. *Nature*, 431(7009):676–680, 2004.

O. Pannekoucke and S. Massart. Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, 134(635):1425–1438, 2008.

J. R. Parrington, W. H. Zoller, and N. K. Aras. Asian dust: Seasonal transport to the Hawaiian Islands. *Science*, 220(4593):195–197, 1983.

D. F. Parrish and J. C. Derber. The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8):1747–1763, 1992.

K. D. Perry, T. A. Cahill, R. A. Eldred, D. D. Dutcher, and T. E. Gill. Long-range transport of North African dust to the eastern United States. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 102(D10):11225–11238, 1997.

A. Petzold and B. Kärcher. Aerosols in the atmosphere. In U. Schumann, editor, *Atmospheric Physics: Background – Methods – Trends*, pages 37–53. Springer, 2012.

A. Petzold, M. Gysel, X. Vancassel, R. Hitzenberger, H. Puxbaum, S. Vrochticky, E. Weingartner, U. Baltensperger, and P. Mirabel. On the effects of organic matter and sulphur-containing compounds on the CCN activation of combustion particles. *Atmospheric Chemistry and Physics*, 5(12):3187–3203, 2005.

N. A. Phillips. A coordinate system having some special advantages for numerical forecasting. *Journal of Meteorology*, 14(2):184–185, 1957.

J. B. Pollack, O. B. Toon, and B. N. Khare. Optical properties of some terrestrial rocks and glasses. *Icarus*, 19(3):372–389, 1973.

A. Prata. Observations of volcanic ash clouds in the 10-12 μm window using AVHRR/2 data. *International Journal of Remote Sensing*, 10(4-5):751–761, 1989.

A. Prata. Satellite detection of hazardous volcanic clouds and the risk to global air traffic. *Natural hazards*, 51(2):303–324, 2009.

A. Prata and A. Tupper. Aviation hazards from volcanoes: The state of the science. *Natural hazards*, 51(2):239–244, 2009.

J. Prospero, R. Glaccum, and R. Nees. Atmospheric transport of soil dust from Africa to South America. *Nature*, 1981.

J. M. Prospero. Long-range transport of mineral dust in the global atmosphere: Impact of African dust on the environment of the southeastern United States. *Proceedings of the National Academy of Sciences*, 96(7):3396–3403, 1999.

J. M. Prospero and R. T. Nees. Impact of the north african drought and El Nino on mineral dust in the Barbados trade winds. *Nature*, 1986.

J. M. Prospero, W. M. Landing, and M. Schulz. African dust deposition to Florida: Temporal and spatial variability and comparisons to models. *Journal of Geophysical Research: Atmospheres*, 115(D13):D13304, 2010. doi: 10.1029/2009JD012773.

H. R. Pruppacher and J. D. Klett. *Microphysics of clouds and precipitation*. D. Reidel, Hingham, MA, 1978.

H. R. Pruppacher, J. D. Klett, and P. K. Wang. *Microphysics of clouds and precipitation*. Kluwer Academic Publishers, Dordrecht, 1997.

M. Pósfai, R. Simonics, J. Li, P. V. Hobbs, and P. R. Buseck. Individual aerosol particles from biomass burning in southern Africa: 1. compositions and size distributions of carbonaceous particles. *Journal of Geophysical Research: Atmospheres*, 108(D13), 2003. doi: 10.1029/2002JD002291.

F. Rabier, A. Bouchard, E. Brun, A. Doerenbecher, S. Guedj, V. Guidard, F. Karbou, V.-H. Peuch, L. El Amraoui, D. Puech, et al. The Concordiasi project in Antarctica. *Bulletin of the American Meteorological Society*, 91(1):69–86, 2010.

G. Ramachandran and P. C. Reist. Characterization of morphological changes in agglomerates subject to condensation and evaporation using multiple fractal dimensions. *Aerosol science and technology*, 23(3):431–442, 1995.

D. Randall. *An introduction to atmospheric modeling*. Citeseer, 2004.

P. J. Rasch, J. Feichter, K. Law, N. Mahowald, J. Penner, C. Benkovitz, C. Genthon, C. Giannakopoulos, P. Kasibhatla, D. Koch, H. Levy, T. Maki, M. Prather, D. L. Roberts, G.-J. Roelofs, D. Stevenson, Z. Stockwell, S. Taguchi, M. Kritz, M. Chipperfield, D. Baldocchi, P. Mcmurry, L. Barrie, Y. Balkanski, R. Chatfield, E. Kjellstrom, M. Lawrence, H. N. Lee, J. Lelieveld, K. J. Noone, J. Seinfeld, G. Stenchikov, S. Schwartz, C. Walcek, and D. Williamson. A comparison of scavenging and deposition processes in global models: results from the WCRP Cambridge Workshop of 1995. *Tellus B*, 52(4):1025–1056, 2000.

L. A. Remer, D. Tanre, Y. J. Kaufman, C. Ichoku, S. Mattoo, R. Levy, D. A. Chu, B. Holben, O. Dubovik, A. Smirnov, et al. Validation of MODIS aerosol retrieval over ocean. *Geophysical research letters*, 29(12):MOD3–1, 2002.

L. A. Remer, Y. Kaufman, D. Tanré, S. Mattoo, D. Chu, J. V. Martins, R.-R. Li, C. Ichoku, R. Levy, R. Kleidman, et al. The MODIS aerosol algorithm, products, and validation. *Journal of the atmospheric sciences*, 62(4):947–973, 2005.

L. A. Remer, R. G. Kleidman, R. C. Levy, Y. J. Kaufman, D. Tanré, S. Mattoo, J. V. Martins, C. Ichoku, I. Koren, H. Yu, and B. Holben. Global aerosol climatology from the MODIS satellite sensors. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D14):D14S07, 2008. doi: 10.1029/2007JD009661.

J.-B. Renard, G. Berthet, F. Jégou, M. Jeannot, L. Jourdain, F. Dulac, M. Mallet, J.-C. Dupont, C. Thaury, T. Tonnelier, et al. LOAC (Light Optical Particle Counter): a new small aerosol counter with particle characterization capabilities for surface and airborne measurements. In *EGU General Assembly Conference Abstracts*, volume 15, page 2824, 2013.

R. W. Reynolds, N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang. An improved in situ and satellite SST analysis for climate. *Journal of climate*, 15(13):1609–1625, 2002.

P. Ricaud, J.-L. Attié, H. Teyssedre, L. E. Amraoui, V.-H. Peuch, M. Matricardi, and P. Schluessel. Equatorial total column of nitrous oxide as measured by IASI on MetOp-A: implications for transport processes. *Atmospheric Chemistry and Physics*, 9(12):3947–3956, 2009. doi: 10.5194/acp-9-3947-2009.

C. Rodgers. Some extensions and applications of the new random model for molecular band transmission. *Quarterly Journal of the Royal Meteorological Society*, 94(399):99–102, 1968.

C. Rolf. *Lidar observations of natural and volcanic-ash-induced cirrus clouds*, volume 163. Forschungszentrum Jülich, 2012.

P. Royer, L. Sauvage, A. Bizard, and L. Thobois. Monitoring and characterization of atmospheric aerosols with raman and dual-polarization lidars. In *SPIE Remote Sensing*, pages 88940J–88940J. International Society for Optics and Photonics, 2013.

I. Sano, S. Mukai, M. Yamano, T. Takamura, T. Nakajima, and B. Holben. Calibration and validation of retrieved aerosol properties based on {AERONET} and {SKYNET}. *Advances in Space Research*, 32(11):2159 – 2164, 2003. ISSN 0273-1177. doi: http://dx.doi.org/10.1016/S0273-1177(03)90538-2. Calibration, Characterization of Satellite Sensors, Physical Parameters Derived from Satellite Data.

M. Schnaiter, C. Linke, O. Möhler, K.-H. Naumann, H. Saathoff, R. Wagner, U. Schurath, and B. Wehner. Absorption amplification of black carbon internally mixed with secondary organic aerosol. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 110(D19), 2005.

M. Schroedter-Homscheidt, H. Elbern, and T. Holzer-Popp. Observation operator for the assimilation of aerosol type resolving satellite measurements into a chemical transport model. *Atmospheric Chemistry and Physics*, 10(21):10435–10452, 2010.

G. L. Schuster, M. Vaughan, D. MacDonnell, W. Su, D. Winker, O. Dubovik, T. Lapyonok, and C. Trepte. Comparison of calipso aerosol optical depth retrievals to aeronet measurements, and a climatology for the lidar ratio of dust. *Atmospheric Chemistry and Physics*, 12(16):7431–7452, 2012. doi: 10.5194/acp-12-7431-2012.

N. Schutgens, T. Miyoshi, T. Takemura, and T. Nakajima. Applying an ensemble Kalman filter to the assimilation of AERONET observations in a global aerosol transport model. *Atmospheric Chemistry and Physics*, 10(5):2561–2576, 2010.

C. Seigneur, B. Pun, P. Pai, J.-F. Louis, P. Solomon, C. Emery, R. Morris, M. Zahniser, D. Worsnop, P. Koutrakis, W. White, and I. Tombach. Guidance for the performance evaluation of three-dimensional air quality modeling systems for particulate matter and visibility. *Journal of the Air & Waste Management Association*, 50(4):588–599, 2000.

J. H. Seinfeld and S. N. Pandis. *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley & Sons, New York, USA, 1998.

T. Sekiyama, T. Tanaka, A. Shimizu, and T. Miyoshi. Data assimilation of CALIPSO aerosol observations. *Atmospheric Chemistry and Physics*, 10(1):39–49, 2010.

N. Semane, V.-H. Peuch, L. El Amraoui, H. Bencherif, S. Massart, D. Cariolle, J.-L. Attié, and R. Abida. An observed and analysed stratospheric ozone intrusion over the high Canadian Arctic UTLS region during the summer of 2003. *Quarterly Journal of the Royal Meteorological Society*, 133(S2):171–178, 2007.

A. Simmons and R. Strüfing. Numerical forecasts of stratospheric warming events using a model with a hybrid vertical coordinate. *Quarterly Journal of the Royal Meteorological Society*, 109(459):81–111, 1983.

B. Sič, L. El Amraoui, V. Marécal, B. Josse, J. Arteta, J. Guth, M. Joly, and P. Hamer. Modelling of primary aerosols in the chemical transport model mocage: development and evaluation of aerosol physical parameterizations. *Geoscientific Model Development*, 8(2): 381–408, 2015. doi: 10.5194/gmd-8-381-2015.

S. Slinn and W. Slinn. Predictions for particle deposition on natural waters. *Atmospheric Environment (1967)*, 14(9):1013–1016, 1980.

W. Slinn. Some approximations for the wet and dry removal of particles and gases from the atmosphere. *Water, Air, and Soil Pollution*, 7(4):513–543, 1977.

W. Slinn. Predictions for particle deposition to vegetative canopies. *Atmospheric Environment (1967)*, 16(7):1785–1794, 1982a.

W. Slinn. Predictions for particle deposition to vegetative canopies. *Atmospheric Environment (1967)*, 16(7):1785–1794, 1982b. doi: 10.1016/0004-6981(82)90271-2.

W. Slinn and J. Hales. A reevaluation of the role of thermophoresis as a mechanism of in-and below-cloud scavenging. *Journal of the Atmospheric Sciences*, 28(8):1465–1471, 1971.

A. Smirnov, B. N. Holben, Y. J. Kaufman, O. Dubovik, T. F. Eck, I. Slutsker, C. Pietras, and R. N. Halthore. Optical properties of atmospheric aerosol in maritime environments. *Journal of the Atmospheric Sciences*, 59(3), 2002.

M. H. Smith, I. E. Consterdine, and P. M. Park. Atmospheric loadings of marine aerosol during a hebridean cyclone. *Quarterly Journal of the Royal Meteorological Society*, 115 (486):383–395, 1989. doi: 10.1002/qj.49711548610.

I. N. Sokolik and O. B. Toon. Direct radiative forcing by anthropogenic airborne mineral aerosols. *Nature*, 381(6584):681–683, 1996.

A. Soufiani and J. Taine. High temperature gas radiative property parameters of statistical narrow-band model for H_2O, CO_2 and CO, and correlated-K model for H_2O and CO_2. *International Journal of Heat and Mass Transfer*, 40(4):987–991, 1997.

M. Spada, O. Jorba, C. Pérez García-Pando, Z. Janjic, and J. M. Baldasano. Modeling and evaluation of the global sea-salt aerosol distribution: sensitivity to size-resolved and sea-surface temperature dependent emission schemes. *Atmospheric Chemistry and Physics*, 13(23):11735–11755, 2013. doi: 10.5194/acp-13-11735-2013.

R. S. J. Sparks, M. Bursik, S. Carey, J. Gilbert, L. Glaze, H. Sigurdsson, and A. Woods. *Volcanic plumes*. Wiley, 1997.

B. Sportisse. A review of parameterizations for modelling dry deposition and scavenging of radionuclides. *Atmospheric Environment*, 41(13):2683–2698, 2007.

B. Sportisse. *Pollution atmosphérique: des processus à la modélisation*. Springer, 2008.

A. Staniforth and J. Côté. Semi-Lagrangian integration schemes for atmospheric models-a review. *Monthly Weather Review*, 119(9):2206–2223, 1991.

P. Stier, J. H. Seinfeld, S. Kinne, J. Feichter, and O. Boucher. Impact of nonabsorbing anthropogenic aerosols on clear-sky atmospheric absorption. *Journal of Geophysical Research: Atmospheres*, 111(D18):n/a–n/a, 2006. doi: 10.1029/2006JD007147.

A. Stohl, A. Prata, S. Eckhardt, L. Clarisse, A. Durant, S. Henne, N. Kristiansen, A. Minikin, U. Schumann, P. Seibert, et al. Determination of time-and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull eruption. *Atmospheric Chemistry and Physics*, 11(9):4333–4351, 2011.

L. L. Stowe, A. M. Ignatov, and R. R. Singh. Development, validation, and potential enhancements to the second-generation operational aerosol product at the National Environmental Satellite, Data, and Information Service of the National Oceanic and Atmospheric Administration. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 102(D14):16923–16934, 1997.

J. W. Strapp, W. Leaitch, and P. Liu. Hydrated and dried aerosol-size-distribution measurements from the particle measuring systems FSSP-300 probe and the deiced PCASP-100X probe. *Journal of Atmospheric and Oceanic Technology*, 9(5):548–555, 1992.

W. Su, N. G. Loeb, G. L. Schuster, M. Chin, and F. G. Rose. Global all-sky shortwave direct radiative forcing of anthropogenic aerosols from combined satellite observations and GOCART simulations. *Journal of Geophysical Research: Atmospheres*, 118(2):655–669, 2013. doi: 10.1029/2012JD018294.

T. Takemura, T. Nakajima, O. Dubovik, B. N. Holben, and S. Kinne. Single-scattering albedo and radiative forcing of various aerosol species with a global three-dimensional model. *Journal of Climate*, 15(4), 2002.

O. Talagrand. A posteriori validation of assimilation algorithms. In R. Swinbank, V. Shutyaev, and W. Lahoz, editors, *Data Assimilation for the Earth System*, volume 26 of *NATO Science Series*, pages 85–95. Springer Netherlands, 2003. ISBN 978-1-4020-1593-9. doi: 10.1007/978-94-010-0029-1_8.

O. Talagrand. Variational assimilation. In W. Lahoz, B. Khattatov, and R. Menard, editors, *Data Assimilation*, pages 41–67. Springer Berlin Heidelberg, 2010. ISBN 978-3-540-74702-4. doi: 10.1007/978-3-540-74703-1_3.

I. N. Tang, A. Tridico, and K. Fung. Thermodynamic and optical properties of sea salt aerosols. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 102(D19):23269–23275, 1997.

N. Taylor, D. Collins, C. Spencer, D. Lowenthal, B. Zielinska, V. Samburova, and N. Kumar. Measurement of ambient aerosol hydration state at great smoky mountains national park in the southeastern united states. *Atmospheric Chemistry and Physics*, 11(23):12085–12107, 2011.

I. Tegen. Modeling the mineral dust aerosol cycle in the climate system. *Quaternary Science Reviews*, 22(18):1821–1834, 2003.

C. Textor, M. Schulz, S. Guibert, S. Kinne, Y. Balkanski, S. Bauer, T. Berntsen, T. Berglen, O. Boucher, M. Chin, et al. Analysis and quantification of the diversities of aerosol life cycles within AeroCom. *Atmospheric Chemistry and Physics*, 6(7):1777–1813, 2006.

C. Textor, M. Schulz, S. Guibert, S. Kinne, Y. Balkanski, S. Bauer, T. Berntsen, T. Berglen, O. Boucher, M. Chin, et al. The effect of harmonized emissions on aerosol properties in global models–an AeroCom experiment. *Atmospheric Chemistry and Physics*, 7(17): 4489–4501, 2007.

H. Teyssèdre, M. Michou, H. Clark, B. Josse, F. Karcher, D. Olivié, V.-H. Peuch, D. Saint-Martin, D. Cariolle, J.-L. Attié, et al. A new tropospheric and stratospheric chemistry and transport model MOCAGE-Climat for multi-year studies: evaluation of the present-day climatology and sensitivity to surface processes. *Atmospheric Chemistry & Physics*, 7(22), 2007. doi: 10.5194/acp-7-5815-2007.

F. Thieuleux, C. Moulin, F. M. Bréon, F. Maignan, J. Poitou, and D. Tanré. Remote sensing of aerosols over the oceans using MSG/SEVIRI imagery. *Annales Geophysicae*, 23(12): 3561–3568, 2005. doi: 10.5194/angeo-23-3561-2005. URL `http://www.ann-geophys.net/23/3561/2005/`.

C. Todd. A system for computing ice phase hydrometeor development. Technical Report ARG Report 64, Paper 121, Atmopsheric Research Group, Altadena, 1964.

O. Torres, P. Bhartia, J. Herman, A. Sinyuk, P. Ginoux, and B. Holben. 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.

K. Tørseth, W. Aas, K. Breivik, A. Fjæraa, M. Fiebig, A. Hjellbrekke, C. Lund Myhre, S. Solberg, and K. Yttri. Introduction to the European Monitoring and Evaluation Programme (EMEP) and observed atmospheric composition change during 1972–2009. *Atmospheric Chemistry and Physics*, 12(12):5447–5481, 2012.

K. Tsigaridis, Y. Balkanski, M. Schulz, and A. Benedetti. Global error maps of aerosol optical properties: an error propagation analysis. *Atmospheric Chemistry and Physics Discussions*, 8(4):16027–16059, 2008. doi: 10.5194/acpd-8-16027-2008.

A. Tupper, C. Textor, M. Herzog, H.-F. Graf, and M. S. Richards. Tall clouds from small eruptions: the sensitivity of eruption height and fine ash content to tropospheric instability. *Natural hazards*, 51(2):375–401, 2009.

I. Uno, K. Eguchi, K. Yumimoto, T. Takemura, A. Shimizu, M. Uematsu, Z. Liu, Z. Wang, Y. Hara, and N. Sugimoto. Asian dust transported one full circuit around the globe. *Nature Geoscience*, 2(8):557–560, 2009.

C. R. Usher, A. E. Michel, and V. H. Grassian. Reactions on mineral dust. *Chemical reviews*, 103(12):4883–4940, 2003.

H. C. Van de Hulst. *Light scattering by small particles*. Courier Dover Publications, 1981.

G. R. van der Werf, J. T. Randerson, L. Giglio, G. Collatz, M. Mu, P. S. Kasibhatla, D. C. Morton, R. DeFries, Y. v. Jin, and T. T. van Leeuwen. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric Chemistry and Physics*, 10(23):11707–11735, 2010. doi: 10.5194/acp-10-11707-2010.

E. Vermote, D. Tanré, J. Deuzé, M. Herman, J. Morcrette, S. Kotchenova, and T. Miura. Second simulation of the satellite signal in the solar spectrum (6S), 6S user guide version 3 (november, 2006), 2006.

E. Vignati, M. Karl, M. Krol, J. Wilson, P. Stier, and F. Cavalli. Sources of uncertainties in modelling black carbon at the global scale. *Atmospheric chemistry and physics*, 10(6): 2595–2611, 2010.

X. Wang, L. Zhang, and M. Moran. Uncertainty assessment of current size-resolved parameterizations for below-cloud particle scavenging by rain. *Atmospheric Chemistry and Physics*, 10(12):5685–5705, 2010. doi: 10.5194/acp-10-5685-2010.

Y. Wang, K. Sartelet, M. Bocquet, and P. Chazette. Assimilation of ground versus lidar observations for PM_{10} forecasting. *Atmospheric Chemistry and Physics Discussions*, 13 (9):269–283, 2013.

Y. Wang, K. Sartelet, M. Bocquet, and P. Chazette. Modelling and assimilation of lidar signals over greater paris during the megapoli summer campaign. *Atmospheric Chemistry and Physics*, 14(7):3511–3532, 2014.

S. Watkin. The application of AVHRR data for the detection of volcanic ash in a Volcanic Ash Advisory Centre. *Meteorological Applications*, 10(04):301–311, 2003.

A. Weaver and P. Courtier. Correlation modelling on the sphere using a generalized diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, 127(575):1815–1846, 2001.

H. Webster, D. Thomson, B. Johnson, I. Heard, K. Turnbull, F. Marenco, N. Kristiansen, J. Dorsey, A. Minikin, B. Weinzierl, et al. Operational prediction of ash concentrations in the distal volcanic cloud from the 2010 Eyjafjallajökull eruption. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 117(D20), 2012.

E. J. Welton, J. R. CampbellL, J. D. Spinhirne, and V. Stanley Scott. Global monitoring of clouds and aerosols using a network of micro-pulse lidar systems. In *Lidar Remote Sensing for Industry and Environmental Monitoring - SPIE proceedings series*, pages 151–158. Society of Photo-Optical Instrumentation Engineers, 2001.

M. Wentzel, H. Gorzawski, K.-H. Naumann, H. Saathoff, and S. Weinbruch. Transmission electron microscopical and aerosol dynamical characterization of soot aerosols. *Journal of aerosol science*, 34(10):1347–1370, 2003.

F. M. White. *Viscous fluid flow*. McGraw-Hill New York, 1991.

J. Williams, M. Scheele, P. Van Velthoven, J.-P. Cammas, V. Thouret, C. Galy-Lacaux, and A. Volz-Thomas. The influence of biogenic emissions from Africa on tropical tropospheric ozone during 2006: a global modeling study. *Atmospheric Chemistry and Physics*, 9(15): 5729–5749, 2009. doi: 10.5194/acp-9-5729-2009.

D. L. Williamson and P. J. Rasch. Two-dimensional semi-Lagrangian transport with shape-preserving interpolation. *Monthly Weather Review*, 117(1):102–129, 1989.

D. Winker, J. Pelon, J. Coakley Jr, S. Ackerman, R. Charlson, P. Colarco, P. Flamant, Q. Fu, R. Hoff, C. Kittaka, et al. The CALIPSO mission: A global 3D view of aerosols and clouds. *Bulletin of the American Meteorological Society*, 91(9):1211–1229, 2010.

D. M. Winker, R. H. Couch, and M. McCormick. An overview of LITE: NASA's lidar in-space technology experiment. *Proceedings of the IEEE*, 84(2):164–180, 1996.

D. M. Winker, W. H. Hunt, and M. J. McGill. Initial performance assessment of CALIOP. *Geophysical Research Letters*, 34(19), 2007.

W. J. Wiscombe. *Mie scattering calculations: advances in technique and fast, vector-speed computer codes*. Atmospheric Analysis and Prediction Division, National Center for Atmospheric Research, 1979, revised 1996.

W. J. Wiscombe. Improved Mie scattering algorithms. *Applied optics*, 19(9):1505–1509, 1980.

A. Wonaschütz, R. Hitzenberger, H. Bauer, P. Pouresmaeil, B. Klatzer, A. Caseiro, and H. Puxbaum. Application of the integrating sphere method to separate the contributions of brown and black carbon in atmospheric aerosols. *Environmental science & technology*, 43(4):1141–1146, 2009.

A. W. Woods. The dynamics of explosive volcanic eruptions. *Reviews of geophysics*, 33(4): 495–530, 1995.

A. W. Woods and S. Self. Thermal disequilibrium at the top of volcanic clouds and its effect on estimates of the column height. *Nature*, 355(6361):628–630, 1992.

S. Woodward, D. Roberts, and R. Betts. A simulation of the effect of climate change–induced desertification on mineral dust aerosol. *Geophysical Research Letters*, 32(18), 2005.

K.-M. Xu and D. A. Randall. A semiempirical cloudiness parameterization for use in climate models. *Journal of the atmospheric sciences*, 53(21):3084–3102, 1996.

C. Zehner, editor. *Monitoring volcanic ash from space*, ESA Scientific & Technical Memoranda, 2012.

C. Zender. Particle size distributions: theory and application to aerosols, clouds, and soils. `http://dust.ess.uci.edu/facts/psd/psd.pdf`, 2010.

J. Zhang, J. S. Reid, and B. N. Holben. An analysis of potential cloud artifacts in modis over ocean aerosol optical thickness products. *Geophysical Research Letters*, 32(15):L15803, 2005. doi: 10.1029/2005GL023254.

J. Zhang, J. S. Reid, D. L. Westphal, N. L. Baker, and E. J. Hyer. A system for operational aerosol optical depth data assimilation over global oceans. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D10), 2008.

J. Zhang, J. R. Campbell, E. J. Hyer, J. S. Reid, D. L. Westphal, and R. S. Johnson. Evaluating the impact of multisensor data assimilation on a global aerosol particle transport model. *Journal of Geophysical Research: Atmospheres*, 119(8):4674–4689, 2014.

K. Zhang, D. O'donnell, J. Kazil, P. Stier, S. Kinne, U. Lohmann, S. Ferrachat, B. Croft, J. Quaas, H. Wan, et al. The global aerosol-climate model ECHAM-HAM, version 2: sensitivity to improvements in process representations. *Atmospheric Chemistry and Physics*, 12(19):8911–8949, 2012a. doi: 10.5194/acp-12-8911-2012.

L. Zhang, X. Wang, M. Moran, and J. Feng. Review and uncertainty assessment of size-resolved scavenging coefficient formulations for below-cloud snow scavenging of atmospheric aerosols. *Atmospheric Chemistry and Physics*, 13(19):10005–10025, 2013. doi: 10.5194/acp-13-10005-2013.

X. Zhang, Y. Wang, T. Niu, X. Zhang, S. Gong, Y. Zhang, and J. Sun. Atmospheric aerosol compositions in China: spatial/temporal variability, chemical signature, regional haze distribution and comparisons with global aerosols. *Atmospheric Chemistry and Physics*, 12(2):779–799, 2012b.

T. Zhao, S. L. Gong, X. Zhang, A. Abdel-Mawgoud, and Y. Shao. An assessment of dust emission schemes in modeling east Asian dust storms. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 111(D5):D05S90, 2006. doi: 10.1029/2004JD005746.

Y. Zhu, W. C. Hinds, S. Kim, S. Shen, and C. Sioutas. Study of ultrafine particles near a major highway with heavy-duty diesel traffic. *Atmospheric Environment*, 36(27):4323–4335, 2002.

Part III

Appendices

A

Example of tangent-linear and adjoint code construction

A tangent-linear and adjoint coding is illustrated in this simple example. We start from a simple function:

$$f(a,b) = 2ab. \tag{A.1}$$

The tangent-linear code is obtained by differentiation of Eq. (A.1):

$$\begin{aligned}\delta f &= \frac{\partial f}{\partial a}\delta a + \frac{\partial f}{\partial b}\delta b \\ &= 2b\delta a + 2a\delta b.\end{aligned}$$

This equation can be writen in its matrix form:

$$\begin{bmatrix}\delta a \\ \delta b \\ \delta f\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 \\ 0 & 1 & 0 \\ 2b & 2a & 0\end{bmatrix} \times \begin{bmatrix}\delta a \\ \delta b \\ \delta f\end{bmatrix},$$

which has to be transposed to obtain the adjoint form:

$$\begin{bmatrix}\delta a^* \\ \delta b^* \\ \delta f^*\end{bmatrix} = \begin{bmatrix}1 & 0 & 2b \\ 0 & 1 & 2a \\ 0 & 0 & 0\end{bmatrix} \times \begin{bmatrix}\delta a^* \\ \delta b^* \\ \delta f^*\end{bmatrix}.$$

From here we can write adjoint expressions corresponding to Eq. (A.1):

$$\begin{aligned}\delta a^* &= \delta a^* + 2b\delta f^* \\ \delta b^* &= \delta b^* + 2a\delta f^* \\ \delta f^* &= 0.\end{aligned}$$

Every discrete operation in the tangent linear operator (and non-linear forward operator) has a corresponding operation in the adjoint operator, but the order of execution is reversed.

B

Sensitivity test of the implementation of model variables during assimilation

As explained in Section 5.3, model variables, other than the control variable, that are necessary to the tangent linear and the adjoint operator during the minimisation of the cost function, has to be approximated by one of their characteristics values: their value at the beginning of the assimilation cycle, or at the end of the cycle, or their average during the cycle.

We made sensitivity tests to determine which choice is the best, and to which degree each approximation is accurate. The sensitivity tests are composed of passing the tangent linear test (Eq. (5.9)) with varying the size of the assimilation cycle and varying the approximation method. The concerned fields are relative mass contributions of bins, pressure and specific humidity. For the tangent linear test execution, the model state x is taken at the beginning of the cycle at 0 h. The perturbed model states are taken at the end of cycle of different lengths: 1 h, 3 h, 6 h and 12 h. The tangent linear operator $\mathbf{H}$ is run with the perturbation with its relative mass contributions of bins:

- calculated from the perturbation itself, in which case the tangent linear test should always be passed with the exact value of one. This is only the control case.
- calculated from the model state x at the beginning of the cycle,
- calculated from the perturbed model state $x + \delta x$ at the end of the cycle,
- averaged over the cycle, taking into account the model state in each time slot.

Table B.1: Percentage of times when the tangent linear test returned the value closest to one by using one of three tested choices for the relative mass contributions of bins. The bin contributions are calculated at the beginning of the assimilation cycle, at the end of the cycle, or averaged over the cycle. Tests were performed for different cycle lengths.

	Relative contributions of bins calculated at		
	the cycle beginning	the cycle end	averaged
1 h cycle	36%	32%	32%
3 h cycle	34%	31%	35%
6 h cycle	31%	31%	38%
12 h cycle	30%	30%	40%

Taking necessary meteorological fields from different times in the cycle, or averaging them, do not produce any important effect on the results of the tangent linear test. Results of the tests where we differ relative contributions of bins as described above, are presented in Table B.1. We looked which way of calculating the bin contributions would give the value the closest to 1 in the tangent linear test.

We concluded that our linearisation choice (using the background relative mass contributions of bins also for the perturbation) has a much more important effect than the way of calculating those contributions. The differences produced by calculating them at different times or averaging them were minor and did not clearly make any of the calculating methods preferable. In the end, in our assimilation system we decided to average them over the assimilation cycle, because for longer cycles averaging gave the slight improvement.

C

Validation of the AOD assimilation during the ChArMEx campaign

To validate the AOD assimilation in MOCAGE, for the ChArMEx campaign we performed similar analysis as we did for the TRAQA campaign.

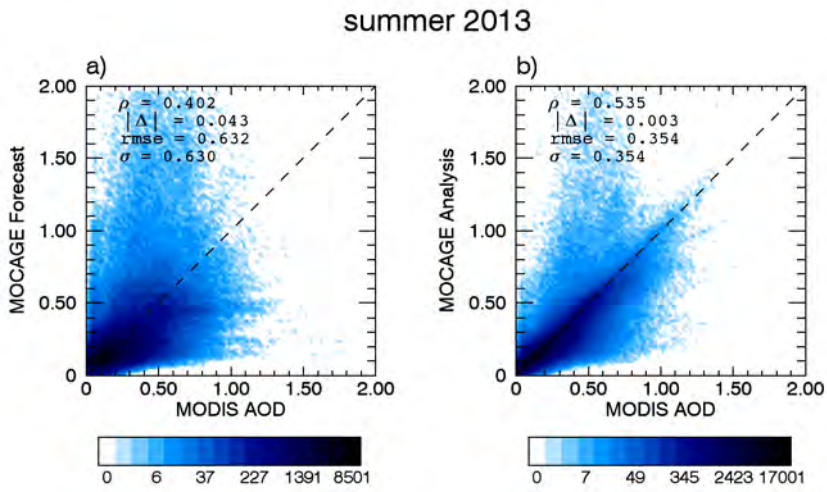

Figure C.1: Scatterplots of aerosol optical depths of assimilated MODIS observations and: the forecast **(a)**, and analysis **(b)**. In each panel, correlation (ρ), absolute bias (Δ), root mean square error (RMSE) and standard deviation (σ) are noted. The assimilated data correspond to the period of the ChArMEx campaign from 12.06.2013 until 05.07.2017., and covers the MEDI02 domain (Fig. 5.4).

OmF and OmA analysis We evaluated the histograms of OmF (Observation minus Forecast) and OmA (Observation minus Analysis) in Fig. C.2. The OmF and OmA differences in the ChArMEx assimilation experiment gave the same conclusions as for TRAQA. The assimilation narrowed bias between the model and observation to the similar degree as in Fig. 5.6. Statistics of the forecast and analysis with assimilated MODIS observations shows that the assimilation improved all considered statistical parameters Fig. C.1.

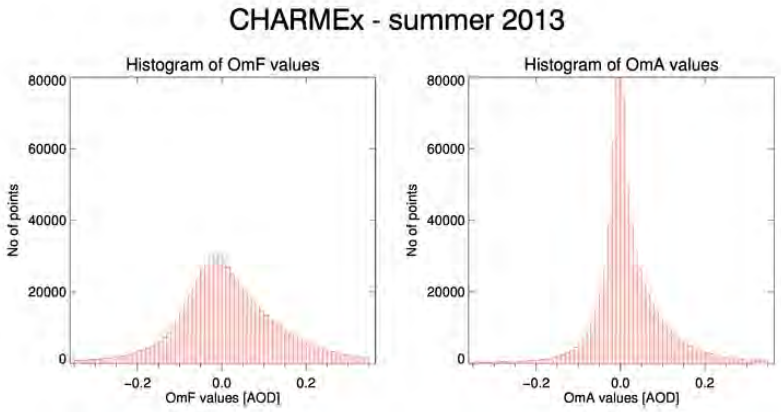

Figure C.2: Histograms of (**left**) OmF (Observation minus Forecast), and (**rigth**) OmA (Observation minus Analysis) for the period of the ChArMEx campaign from from 12.06.2013 until 05.07.2013.

Comparison with SEVIRI When compared with the independent observations from SEVIRI on the scatterplot (Fig. C.3), the direct model run often overestimated AOD, in the contrast to the run for the TRAQA period. We see also some underestimation in the model, but present to a lesser degree than the overestimated values. The assimilation reduced this disagreement, with better performance in reducing the overestimated points than correcting underestimated values. Statistical parameters of the assimilation model run compared with SEVIRI observations were slightly worse than for the TRAQA period (Fig. 5.8).

The timeseries comparison with SEVIRI observations shows the AOD variability during the summer 2013 over the western Mediterranean (Fig. C.4). The first part of the period is marked by a desert dust outbreak from the Northern Africa lasting from 15. June to 21. June. This event is overestimated in the direct model run. The assimilation well corrected the modeled AOD in the assimilation model run. In the second part of the period, there is a case of a long-range transport of biomass burning from North America. It lasted

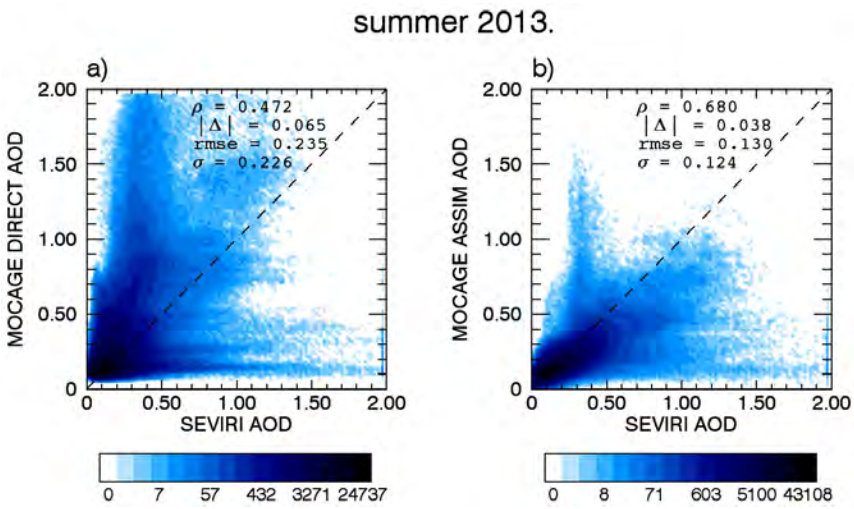

Figure C.3: Scatterplots of aerosol optical depths from the independent observation dataset (SEVIRI) and the simulations: the direct model run **(a)** and the assimilation model run **(b)**. In each panel, correlation (ρ), absolute bias (Δ), root mean square error (RMSE) and standard deviation (σ) are noted. The included data correspond to the period of the ChArMEx campaign from 12.06.2013 until 05.07.2013., and covers the whole MEDI02 domain (Fig. 5.4).

Table C.1: Correlation (ρ), absolute bias (Δ) and root mean square error (RMSE) between SEVIRI observations and MOCAGE direct/assimilation run for the western Mediterranean during the ChArMEx campaign between 12.06.2013 and 05.07.2013. Mean number of SEVIRI observations per hour is also given. The region which is considered is presented in Fig. C.5, and data correspond to Fig. C.4.

		MOCAGE direct			MOCAGE assimilation		
	$\overline{N_{obs}}$ [h^{-1}]	ρ	Δ	RMSE	ρ	Δ	RMSE
SEVIRI - ChArMEx	20828	0.28	0.04	0.14	0.65	0.02	0.07

a half dozen days over the Mediterranean basin from 24. June to 29. June. In the direct model, its duration is well simulated thanks to GFAS emissions, but its strength was underestimated during the second, stronger part of the event. The assimilation had an positive effect on that period, but the SEVIRI observations were still giving higher AOD values than the model. In the end of the analysed period, there was a case of another desert dust event. This event had a weaker strength than the first one and it was correctly simulated in both direct model run and in the assimilation model run. Table C.1 shows the statistics of Fig. C.4. The mean number of the available observations per

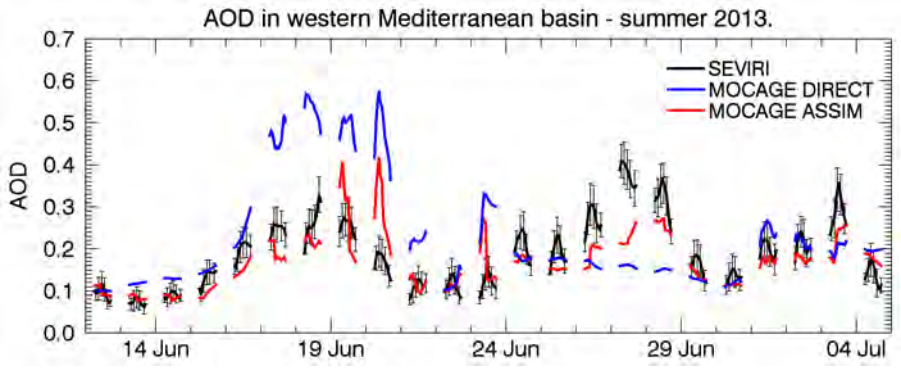

Figure C.4: Time series of aerosol optical depth at 550 nm of SEVIRI data, the direct model and the assimilation model run over the western Mediterranean (0–10° E, 35–45° N) for the period of the ChArMEx campaign from 12.06.2013 until 05.07.2013. The considered region is also marked in Fig. C.5 by box. Correlation, bias and root mean square error for both the direct model and the assimilation model run as compared to SEVIRI data are given in Table C.1.

hour is similar as for the TRAQA period, but the correlation of the model to observations is lower. However, assimilation of the MODIS data significantly improved the performance of the model.

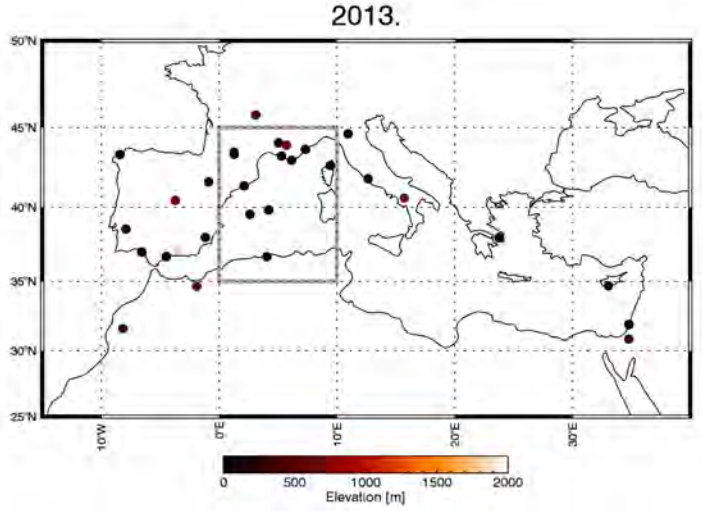

Figure C.5: Positions of AERONET stations used in this study for the period of ChArMEx campaign in summer 2013. The black box mark the region from which we considered SEVIRI data used in Fig. C.4, Table C.1.

Comparison with AERONET In the summer 2013 we used data from 30 AERONET stations situated in or around the Mediterranean basin to compare them with the model direct run and assimilation model run (Fig. C.5). Timeseries plots for ten stations are presented in Fig. C.7, and the statistics for all stations in Table C.2. During the ChArMEx period, in the AERONET data we see numerous events that elevated AOD values. But, they have a less generalized effect over the region than the events during the TRAQA campaign, and events are seen in data usually from a smaller number of stations. This indicates that the scales of these events are smaller, suggesting the more changing conditions in that period.

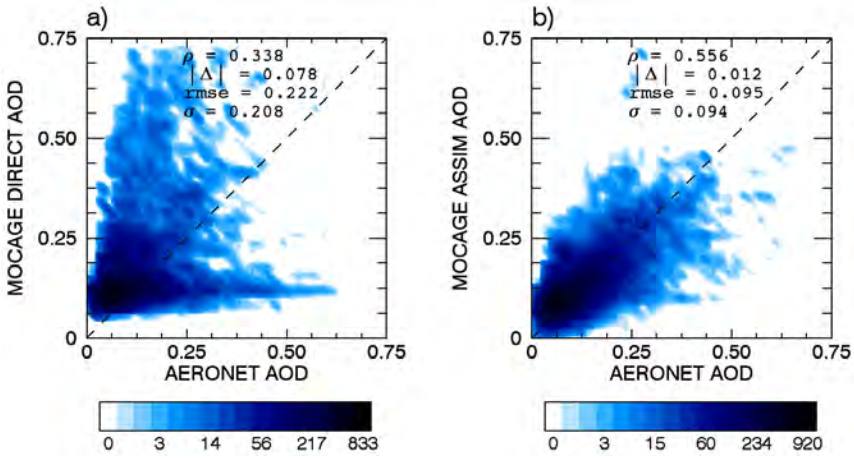

Figure C.6: Scatterplots of aerosol optical depths from the independent observation dataset (AERONET) and the simulations: the direct model run **(a)** and the assimilation model run **(b)**. In each panel, correlation (ρ), absolute bias (Δ), root mean square error (RMSE) and standard deviation (σ) are noted. The included data correspond to the period of the ChArMEx campaign from 12.06.2013 until 05.07.2013., and covers all stations presented in (Fig. C.5).

The first desert dust event present in SEVIRI data (Fig. C.4), is seen in AERONET data how it moves from Morocco, along Spanish coast, over Balearic Islands and Italy, but it is not seen in others regions. The biomass burning event is seen in a larger number of stations. In the direct model run, it is underestimated over the western Mediterranean and West Spain, but overestimated at stations in South Spain. It is simulated reasonably well considering its trans-atlantic transport.

In the eastern Mediterranean, in the second part of the considered ChArMEx period, less aerosols were observed than what was simulated in the direct model run. Also, observations did not show such large variations in AOD as the model did. In all of these cases, the data assimilation adjusted

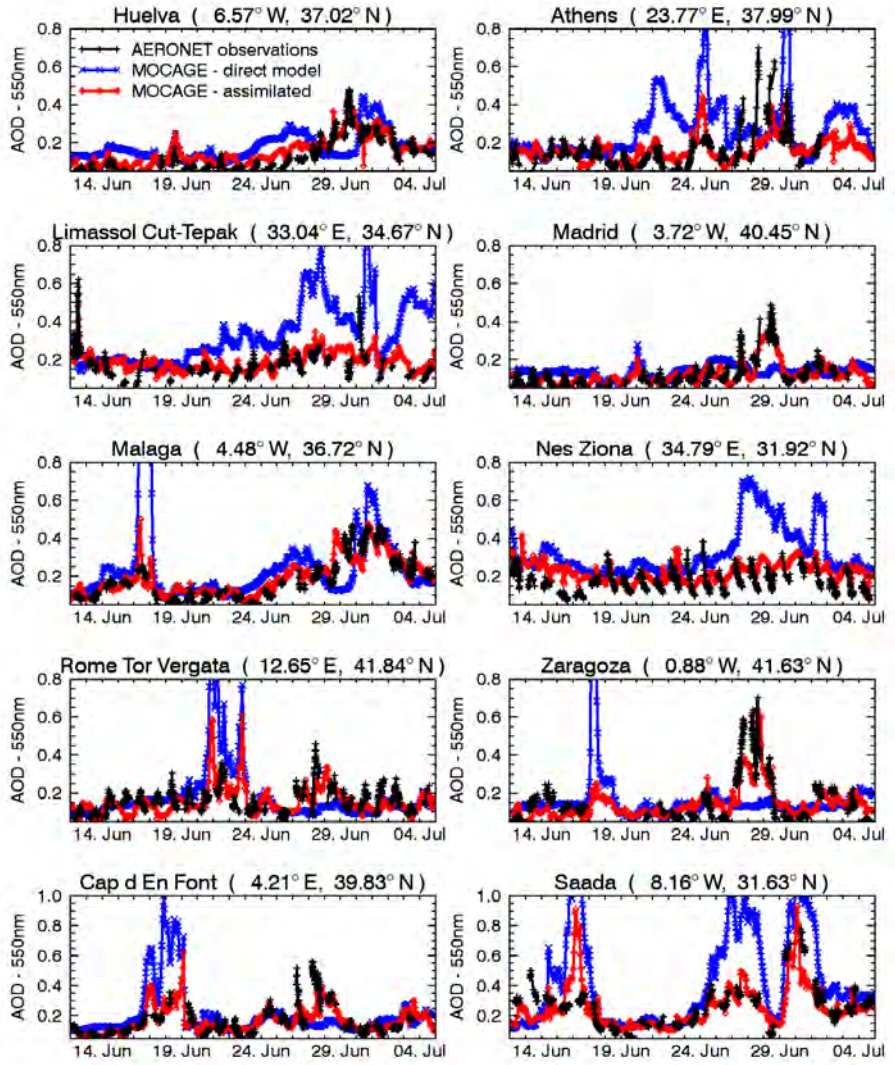

Figure C.7: Time series of aerosol optical depth at 550 nm of AERONET data, the direct model and the assimilation model run for the period of the ChArMEx campaign from 12.06.2013 until 05.07.2013. The presented AERONET data are from ten stations: Huelva (ESP), Athens (GRE), Limassol (CYP), Madrid (ESP), Malaga (ESP), Nes Ziona (ISR), Rome Tor Vergata (ITA), Zaragoza (ESP), Capd'En Font (ESP), Saada (MAR). Correlation, bias and root mean square error for both the direct model and the assimilation model run as compared to AERONET data are given in Table 5.3.

Table C.2: Correlation (ρ), absolute bias (Δ) and root mean square error (RMSE) between AERONET observations and MOCAGE direct/assimilation run for the period of the ChArMEx campaign between 12.06.2013 and 05.07.2013. AERONET site locations are presented in Fig. C.5.

		MOCAGE direct			MOCAGE assimilation		
	N_{obs}	ρ	Δ	RMSE	ρ	Δ	RMSE
Athens	921	0.377	0.170	0.266	0.563	0.014	0.080
Aubiere	353	0.491	0.066	0.227	0.338	0.010	0.058
Barcelona	252	0.268	0.026	0.091	0.820	0.024	0.057
Cap d'en Font	988	0.222	0.024	0.126	0.730	0.015	0.069
Carpentras	509	0.501	0.069	0.162	0.699	0.006	0.046
Cerro Poyos	1030	0.514	0.108	0.220	0.846	0.045	0.063
Coruna	612	-0.14	0.030	0.124	0.733	0.020	0.081
Evora	1083	0.239	0.014	0.107	0.752	0.013	0.073
Frioul	758	0.347	0.039	0.112	0.687	0.011	0.051
Huelva	1303	0.393	0.061	0.110	0.899	0.031	0.050
Le Fauga	388	0.133	0.008	0.074	0.748	0.002	0.055
Limassol	697	0.092	0.175	0.240	0.677	0.022	0.059
Madrid	1011	0.209	0.021	0.076	0.840	0.003	0.040
Malaga	1131	0.348	0.096	0.309	0.874	0.014	0.058
Modena	884	0.096	0.041	0.162	0.252	0.068	0.117
Montesoro Bastia	918	0.325	0.032	0.129	0.541	0.013	0.051
Murcia	1128	0.380	0.057	0.221	0.759	0.017	0.071
Nes Ziona	861	0.195	0.145	0.189	0.380	0.047	0.075
OHP Observatoire	935	0.240	0.050	0.138	0.524	0.007	0.066
Oujda	805	0.451	0.250	0.579	0.387	0.045	0.286
Palma de Mallorca	1003	0.336	0.091	0.258	0.720	0.021	0.062
Porquerolles	629	0.580	0.068	0.182	0.678	0.000	0.054
Potenza	837	0.016	0.076	0.171	0.146	0.008	0.052
Rome Tor Vergata	1007	0.322	0.022	0.166	0.555	0.016	0.059
Sede Boker	1134	0.073	0.201	0.226	0.322	0.107	0.121
Saada	854	0.558	0.171	0.305	0.574	0.001	0.107
Seysses	416	-0.04	0.005	0.084	0.608	0.002	0.066
Tizi Ouzou	434	0.276	0.251	0.482	0.283	0.039	0.281
Villefranche	484	0.229	0.072	0.210	0.377	0.009	0.067
Zaragoza	979	-0.09	0.003	0.124	0.770	0.002	0.076
All sites	24344	0.338	0.078	0.221	0.556	0.012	0.095

AOD values to correlate better with observations. Statistics of comparison with AERONET data confirm this strong positive effect Table C.2.

The scatterplot of MOCAGE and AERONET data from all considered stations (Fig. C.6) confirms that the direct model in the ChArMEx period showed stronger overestimation than underestimation of AOD, and that the data assimilation was able to strongly improve the performance of the model. The smaller number of observations per station per day during ChArMEx compared to the TRAQA period suggests that the ChArMEx period could have been more turbulent, provoking also a weaker correlation between the model and observations than during the TRAQA campaign.

In-situ aircraft concentration measurements In Fig. C.8 we compare the modeled aerosol number concentrations with aerosol concentrations measured by the PCASP instrument onboard of the ATR aircraft. During the ChArMEx period, flights with the research aircraft ATR-42 flew over the whole western Mediterrean basin during 13 different days, carrying the PCASP onboard. In Fig. C.8 we presented three flights: flight D of 12.06.2013 from Toulouse to Cagliari (Fig. C.8a), flight E of 14.06.2013 from Cagliari toward Marseille and back (Fig. C.8b), and flight F of 17.06.2014 from Granada to Minorca (Fig. C.8c).

Following the path of the Flight D (Fig. C.8a) we first notice the pollution from Toulouse agglomeration, and then low concentrations during the flight. The assimilation had a certain positive effect on the form of the timeseries curve, but both modeled curves simulated well concentrations during the flight.

During the Flight E (Fig. C.8b) aerosol concentrations were low. By decreasing altitude of the aircraft near Marseille coast, concentrations increased, and this was better simulated in the assimilation model run. The direct model run appeared to have too many aerosols near Sardinia.

In the Flight F data (Fig. C.8c), the aircraft data showed higher aerosol concentrations by flying through the desert dust plume. The assimilation model run improve the number concentrations by lowering the overestimated concentrations in the direct model run. The assimilation did not change significantly the shape of the timeseries curve, suggesting lowering of AOD by similar degree over large area that the aircraft flight covered.

In-situ balloon concentration measurements During ChArMEx, Minorca was a launching base for meteorological balloons on which were also flying LOAC instruments. Figure C.9 shows four flights from 15. June to 18. June.

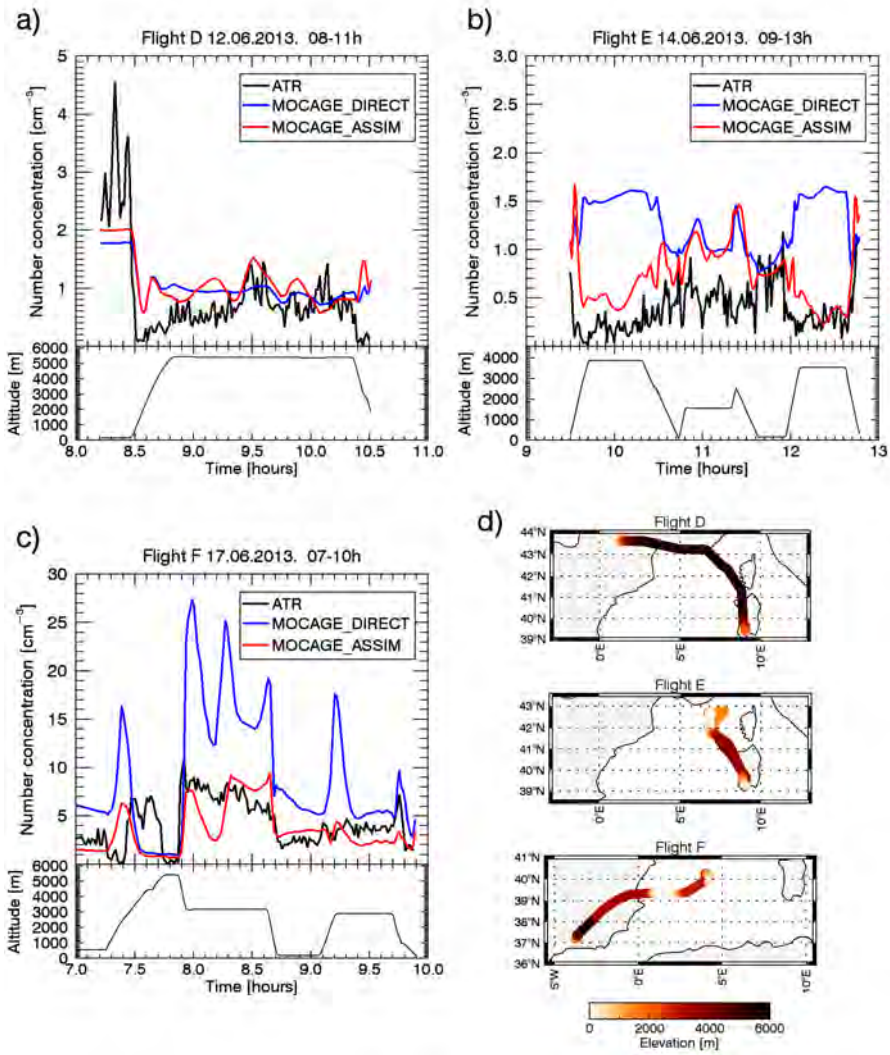

Figure C.8: Aerosol number concentration from the PCASP instrument onboard of the ATR aircraft for three different flights: Flight D of 12.06.2013 **(a)**, Flight E of 14.06.2013 **(b)**, and Flight F of 17.06.2013 **(c)**. The altitude of the aircraft is also given for all three flights. Also, the maps of flight tracks are presented **(d)**.

All flights show plumes of desert dust aerosols. The first flight data (Fig. C.9a) showed that both direct model run and assimilation model simulated well the beginning of the desert dust event over Minorca, with the assimilation bringing some improvement on the profile shape. The assimilated field has a better profile shape which resembles more to the observations. This profile is an another example of AOD assimilation improving the aerosol vertical profile

shape by having aerosols coming from different locations where they were previously assimilated. The other three LOAC flights show an evolution of the dust event which was overestimated in the direct model run. The assimilation model significantly corrected concentrations, always with preserving the profile shape. The third flight (Fig. C.9b) coincided with the Flight F (Fig. C.8c). The Flight F, during its descent to Minorca shows the measured values which were between the direct model run and the assimilation model run. When compared, Fig. C.9b confirms aircraft measurements. The data assimilation tried to match the satellite AOD by lowering aerosol concentrations, but having preserved the profile shape, the part of the plume remained underestimated. Overall, LOAC flights confirm the assimilation capability to improve the aerosol representation.

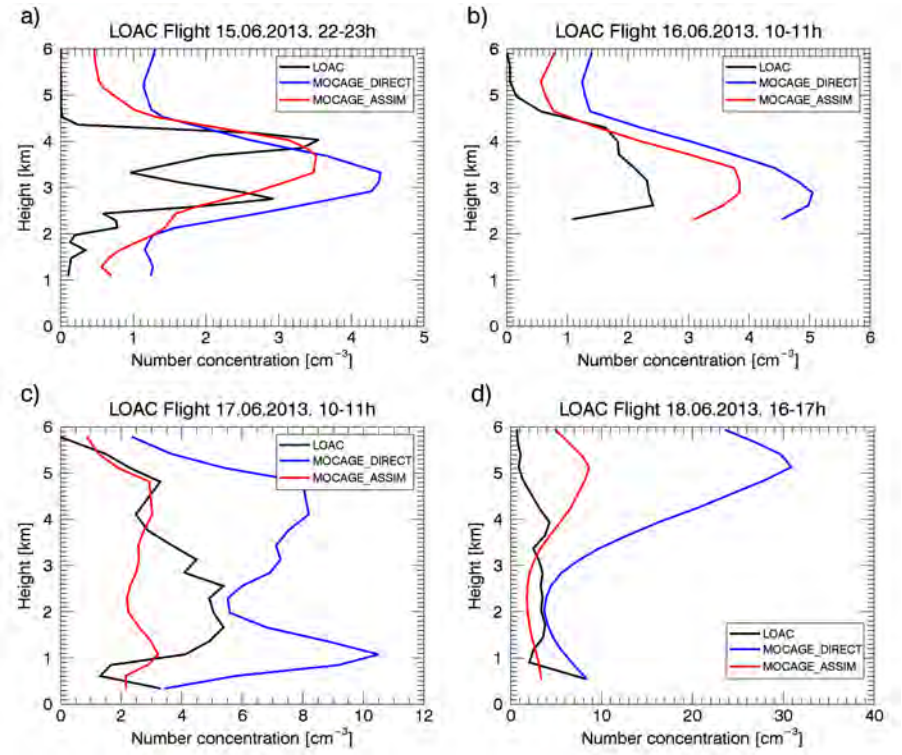

Figure C.9: Aerosol number concentration from the LOAC instrument onboard on meteorological sounding balloons launched from Minorca. The presented flights are performed: on 15.06.2013 **(a)**, on 16.06.2013 **(b)**, on 17.06.2013 **(c)**, and on 18.06.2013 **(c)**. LOAC measurements are compared with the direct model run and the assimilation model run.