THESIS FOR THE DEGREE OF DOCTORATE OF PHILOSOPHY

Use of remote sensing and in situ observations of the atmosphere in chemical transport models

EMMA WARD

CHALMERS

Department of Space, Earth and Environment
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Use of remote sensing and in situ observations of the atmosphere in chemical transport models
EMMA WARD
ISBN: 978-91-7905-133-4
Series number: 4600

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie (ISSN 0346-718X)

Department of Space, Earth and Environment
Chalmers University of Technology
SE - 412 96 Gothenburg, Sweden

Printed by Chalmers Reproservice
Chalmers University of Technology
Gothenburg, Sweden 2019

Use of remote sensing and in situ observations of the atmosphere in chemical transport models

EMMA WARD
Department of Space, Earth and Environment
Chalmers University of Technology

Abstract

Measuring and simulating atmospheric chemical compositions helps us to better understand the impact from long-range transport of different pollutants. Aerosols and tropospheric ozone are examples of pollutants with life times long enough to get transported between continents. Thus, observing and modelling this transport and its interaction with the surrounding atmosphere is vital to providing information about the fate of these pollutants. The main subject of this thesis is to make combined use of, and exploit the advantages of both sources of information. This is done in two ways; (i) by evaluating model results with the help of observations, and (ii) by combining observations with model results by data analysis/assimilation.

The thesis is divided into four studies, where the first study sets the stage by providing a methodology for evaluating lateral boundary conditions (LBCs) in a regional chemical transport model (CTM). The methodology is applied to ozone (O_3) and carbon monoxide (CO), and includes a direct evaluation of LBCs at the boundary of a CTM, as well as an indirect evaluation of a model run. The results show that a combined direct and indirect evaluation give a more complete picture of how well a given set of LBCs perform, compared to using only a direct or indirect comparison.

To prepare for using the methodology from study one with aerosols, we needed an aerosol optical model that can map aerosol concentration fields onto aerosol optical properties. Therefore, the second study of this thesis investigates the impact of implementing a newly developed model with a realistic description of aerosol optical properties. The results show that realistic aerosol descriptions in an aerosol optics model impact radiometric quantities and radiative forcing to the same degree as including or omitting aerosol dynamics in a CTM.

Further, the aim is also to derive LBCs based on CTM results constrained by satellite retrieved aerosol optical properties. This requires us to know how much information we can use from the retrievals to constrain model variables. The main question of the third study therefore is, how much information do extinction and backscattering measurements contain about the chemical composition of atmospheric aerosol? The information content of extinction and backscattering measurements was analysed with a singular-value decomposition and was shown to increase with the number of observations and decrease with the observation error. We also found that the model variable best constrained with these types of measurements, was PM_{10}.

The last and fourth study invokes the new aerosol optics model and the information content analysis, to constrain LBCs obtained from a hemispheric version of the CTM MATCH by use of CALIPSO extinction retrievals at a wavelength of 532 nm. An indirect evaluation with ground based observations, shows a bias reduction for PM_{10}.

Keywords: Chemical transport modelling, Satellite observations, Ground based observations, Air quality, Climate

Appended Papers

The thesis is based on the following papers:

- **E. Andersson**, M. Kahnert, A. Devasthale: Methodology for evaluating lateral boundary conditions in the regional chemical transport model MATCH (v5.5.0) using combined satellite and ground-based observations, *Geosci. Model Dev.* , 8, 3747-3763, doi:10.5194/gmd-8-3747-2015, 2015
- **E. Andersson** and M. Kahnert: Coupling aerosol optics to the MATCH (v5.5.0) chemical transport model and the SALSA (v1) aerosol microphysics module, *Geosci. Model Dev.*, 9 1803-1826, doi: 10.5194/gmd-9-1803-2016, 2016.
- M. Kahnert and **E. Andersson**: How much information do extinction and backscattering measurements contain about the chemical composition of atmospheric aerosol? *Atmos. Chem. Phys.*, 17, 3423-3444, doi:10.5194/acp-17-3423-2017, 2017.
- **E. Ward** and M. Kahnert: Construction and implementation of lateral boundary conditions for aerosol particles from satellite observations of aerosol extinction profile, a prepared manuscript, 2019.

Contents

Acknowledgements

I am grateful for have been given the opportunity to conduct my doctoral studies at Chalmers in collaboration with the Swedish Meteorological and Hydrological Institute (SMHI). This was made possible through funding from the Swedish National Space Board within the OSCES project (no. 101/13).

Without a doubt, I am most grateful to my supervisor, Michael Kahnert, for all his help and guidance. Michael has been a great support and encouragement throughout these years of my doctoral studies. Also, I would like to thank Donal Murtagh, my examiner and boss at Chalmers, who always seems to be able to lift the spirit at the workplace. I would also like to thank David Simpson at Chalmers, and Abhay Devasthale, Manu Thomas, Lennart Robertsson and Robert Bergström at SMHI for their valuable input.

Last, but not least, I want to thank my family and beloved husband for being by my side supporting and helping me through both good and bad. Thank you Marston, without you I would be lost.

1

Background and motivation

Atmospheric chemical composition is of special interest to both climate and air quality research and has, as a result of increasing emissions of anthropogenic pollutants, gained more recognition. Tropospheric ozone and aerosols (solid and or liquid particles suspended in air) are examples of pollutants, with clear anthropogenic signals, that affect both the radiative balance of the Earth as well as human health and the environment. Tropospheric ozone contribute to crop loss by reducing a plant's ability to photosynthesise and sequester carbon (Roshchina and Roshchina 2003, ch. 5), as well as absorbs infrared radiation. Aerosols absorb and scatter radiation, and are a key component in cloud microphysics. These pollutants also have a relative long life time (days to weeks before being removed from the atmosphere) making it possible for them to get transported over long distances ($\sim$10 km to $\sim$1000 km), thus affecting large areas of the globe. By observing and modelling the atmospheric chemical composition, better understanding is gained of the dispersion and impact different pollutants have on the globe. This knowledge can further provide a scientific basis for political decisions regarding mitigation strategies.

Simulating the dispersion of the atmospheric chemical composition requires knowledge about emissions, deposition, and transport as well as chemical and physical transformations. To mathematically represent these processes, a complex and highly coupled system with a large number of balanced and interactive processes needs to be defined. A major problem for the modelling community is deciding which processes to include, which ones to describe in a simplified manner, and which ones to excluded, for example, how many chemical reactions should be taken into accounted, how do we represent long-range transport in a smaller scale model, and how do we best represent aerosol size distributions. These questions need to be addressed and the choice of representation needs to be evaluated.

This thesis focuses on the importance of capturing long-range transport in a regional chemical transport model, which is represented by lateral boundary conditions. Long-range transport of pollutants strongly impacts the background concentration of pollutants, and can have episodic impact on regional and local air quality at all levels of the troposphere. Several studies have pointed out the significance of having well represented lateral boundary conditions. For example, Jiménez et al. (2007) concluded that ground level ozone concentrations on

the Iberian Peninsula are strongly affected by the boundary conditions of both ozone and its precursors. Tang et al. (2007) studied different spatio-temporal resolutions of lateral boundary conditions and how they compared to air-craft measurements. The conclusion was that using dynamic lateral boundary conditions better correlates with measurements at the boundary.

Confronting model simulations, and a model's different input data, with observations is essential for model development. Observations from both in situ and remote sensing instruments are available for this purpose. Large ground based observation networks provide continuous data-sets of, for examples, remotely sensed ozone and aerosol optical properties, as well as in situ measurements of particle number and mass concentrations. Artificial satellites from large space agencies, such as NASA (The National Aeronautics and Space Administration) or ESA (European Space Agency), are used for monitoring atmospheric conditions using remote sensing techniques. Thus providing global data coverage of for example retrieved vertical profiles of ozone and carbon monoxide, and aerosol optical properties.

This thesis aims to use both in situ and remote sensing observations to evaluate and constrain a regional chemical transport modelling system. More specifically, satellites retrieved quantities are used for evaluating and constraining the model's lateral boundary fields, while both satellite and ground based observations are employed for evaluating the model's performance. The first introductory chapters of this thesis, therefore, include descriptions of (i) the terrestrial atmosphere, (ii) observations of the atmosphere, (iii) modelling atmospheric composition, and (iv) making use of observation types and model results.

2

Atmosphere of the Earth

Earth's atmosphere stretches up to about 100 km; it contains 78% nitrogen (N_2), 21% oxygen (O_2) and 1% trace gases and other air pollutants such as liquid or solid particles suspended in the air (aerosols). Despite the relative small concentrations of trace gases and aerosols they contribute significantly to the radiative balance of the Earth and the loss of good air quality close to the surface.

Atmospheric density and pressure decrease exponentially with height, and can be described by $p = p_0 \cdot \exp(-\frac{z}{H})$. Here p_0 is the surface pressure, z the height from the surface, and $H = 7$ to 8 km corresponds to the scale height, which is defined as the vertical height in which the density and pressure decrease by a factor of $\frac{1}{e}$. This equation is derived by assuming a hydro-static equilibrium in the atmosphere, and by using the ideal gas law to describe the pressure, density, and temperature relationship (Wallace and Hobbs 2006). This chapter will give an overview of the atmosphere as a chemical and dynamical system.

2.1 Structure of the Atmosphere

The vertical structure of the atmosphere can be divided into layers of common properties. The most common way to divide the atmosphere is according to how the temperature changes with altitude. Figure 2.1 depicts this temperature change with height and the different layers: troposphere, stratosphere, mesosphere, and thermosphere. The troposphere, which is closest to the surface, is defined by a decreasing temperature with height (a so called negative lapse rate). Air in the troposphere is heated from below due to absorption of solar radiation by the surface of the planet, which then remits this energy as IR radiation, to which the atmosphere is not transparent. This heating energy is distributed through convection, latent and sensible heat fluxes, and radiative processes. At the tropopause, which is the transition zone between the first and the second layer, the negative lapse rate becomes either neutral or positive. In the stratosphere there is a general positive lapse rate due to absorption of ultra-violet radiation (< 400 nm) by ozone (O_3). As one continues upwards, the temperature trend is again inverted at the stratopause and starts to decrease in the mesosphere, where ozone absorption of the solar UV-radiation no longer

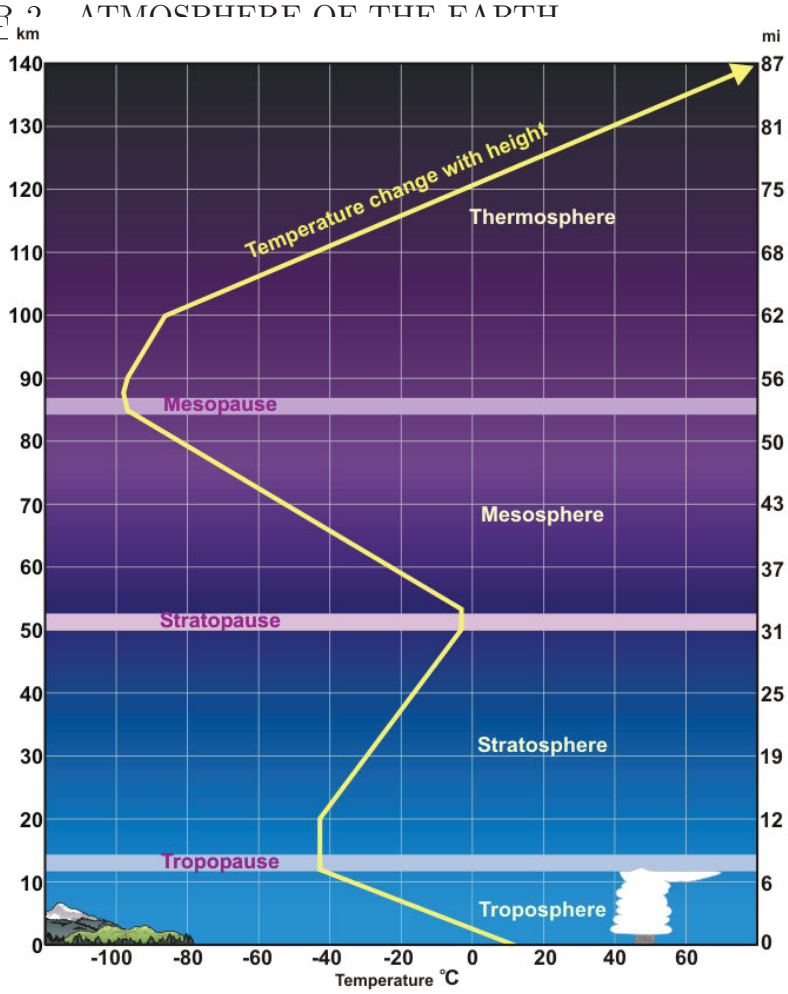


Figure 2.1: Structure of the atmosphere when dividing it into layers according to the vertical temperature structure . Image taken from Wikimedia commons *https://upload.wikimedia.org/wikipedia/commons/2/28/Atmprofile.jpg* .

plays a significant role, and where carbon dioxide (CO_2) acts as a cooling agent (by absorbing infrared radiation from below and radiating it out into space). The uppermost layer reaches out into space and is characterised by an increase in temperature (the thermosphere) due to the absorption of solar radiation that photo-disassociates and ionises air molecules.

This thesis focuses on the troposphere, its chemistry, trace gases, aerosols, and aerosol micro-physics.

2.2 The Troposphere

Approximately 75% of all the air in the atmosphere resides in the troposphere which ranges from 8 km to 16 km, depending on the latitude (highest in the tropics and decreasing pole-ward). The troposphere can further be divided into the atmospheric, or planetary, boundary layer (ABL), which is closest to the surface, and the free troposphere. The ABL is formed as a consequence of the interaction of the atmosphere and the planetary surface; it is characterised by its strong turbulent mixing. In air quality research the height of the ABL is often referred to as the mixing height. It represents the height in which surface or near-surface emitted air pollutants become well-mixed. The height of the ABL can stretch from several tens of meters to several kilometers (Arya 2001) over land surfaces. It varies less over oceans,and is highly dependent upon the

diurnal heating and cooling processes. The free troposphere is a more stratified layer, has generally a negative lapse rate, and is also the part of the troposphere in which long-range transport of air pollutants takes place.

2.2.1 Tropospheric chemistry

Atmospheric chemistry is concerned with the study of the composition and evolution of the atmosphere over various time scales. The composition is influenced by surface emissions, vertical and horizontal transport, deposition, and transformation processes.

Trace gases are transformed by chemical reactions in the gas or liquid phases. Aerosols undergo transformation of their size distribution by microphysical processes, such as nucleation of new aerosols from the gas phase, condensation of vapour onto existing particles, as well as coagulation among particles. Aerosols and their microphysics will be discussed in Sect. 2.2.2. This section focuses on gaseous chemical species and their emission, deposition and reactions, specifically tropospheric ozone and carbon monoxide. These two trace gases are both relatively long-lived species (days to weeks), and they affect the temperature of the planet as well as air quality (see the following paragraphs). Their residence times in the atmosphere also make them good candidates for evaluating chemical transport models.

Tropospheric ozone (O_3)

Ozone is one of the most important oxidisers in the lower part of the atmosphere and is harmful to both plant and animal life, where it affects respiratory system and creates toxins making it hard to breathe (Kampa and E. 2008). Ozone absorbs UV radiation at around 400 nm. As a result, stratospheric ozone effectively attenuates this part of the solar spectrum, which would be harmful to plant and animal life. However, ozone also has absorption bands in the thermal infrared at wavelengths near 9.6 µm and 14.3 µm. Thus elevated concentrations of tropospheric ozone lead to increased absorption of surface thermal radiation, which contributes to the warming effect of the atmosphere. Ozone is not directly emitted into the atmosphere, but photochemically produced from precursor emissions such as volatile organic compounds (VOC), carbon monoxide CO and nitrogen oxides (NO_2 and NO). A chemical reaction scheme of the net production of ozone in the troposphere is shown in [I], where R represents a nonspecific organic compound or hydrocarbon.

$$\begin{aligned}
\mathrm{VOC} + \mathrm{OH} &\xrightarrow{\mathrm{O_2}} \mathrm{RO_2} + \mathrm{H_2O} \\
\mathrm{CO} + \mathrm{OH} &\xrightarrow{\mathrm{O_2}} \mathrm{HO_2} + \mathrm{CO_2} \\
\mathrm{RO_2} + \mathrm{NO} &\xrightarrow{\mathrm{O_2}} \text{secondary VOC} + \mathrm{HO_2} + \mathrm{NO_2} \\
\mathrm{HO_2} + \mathrm{NO} &\longrightarrow \mathrm{OH} + \mathrm{NO_2} \\
\mathrm{NO_2} + \mathrm{h}\nu &\longrightarrow \mathrm{NO} + \mathrm{O} \\
\mathrm{O} + \mathrm{O_2} &\xrightarrow{\mathrm{M}} \mathrm{O_3}
\end{aligned} \qquad [\mathrm{I}]$$

The production of ozone is balanced with several sinks, but the largest sink of tropospheric ozone comes from a reaction with a catalyst such as the hydroxyl

radical OH, nitrogen oxide NO, chlorine, Cl or Bromine Br, according to reaction scheme [II].

$$\begin{aligned} X + O_3 &\longrightarrow XO + O_2 \\ XO + O &\longrightarrow X + O_2 \end{aligned} \qquad \text{[II]}$$

The residence time of ozone in the troposphere ranges from a day to a week.

Carbon monoxide (CO)

Carbon monoxide formed from incomplete combustion, is emitted into the atmosphere from mainly burning fossil fuel and biomass. Its adverse effect on health is well known. CO is a precursor to tropospheric ozone and CO_2 as well as a major sink for the strongest oxidant of the atmosphere, the hydroxyl radical OH (Finlayson-Pitts and Pitts Jr. 2000, Fig. 6.29). In areas where there are elevated CO concentrations, other trace gas concentrations which are dependent upon being removed by OH, such as CH_4, become indirectly affected. The lifetime of carbon monoxide in the troposphere is about a month, making CO a good tracer for transport in models.

2.2.2 Aerosols

Aerosols play a vital role in the atmosphere, since they affect the radiative balance of the Earth, provide a surface for chemical reactions, and impact cloud formation. Aerosols affect the radiative budget directly by scattering short-wave radiation from the sun and by absorbing long-wave radiation from the Earth. They also affect the radiative balance indirectly through their impact on cloud formation by water condensing onto aerosols. These water-covered aerosols can grow to form cloud or fog droplets. The water condensed onto aerosols further provides a medium for liquid-phase chemical reactions to take place.

The size of aerosols can range from ≈0.001 µm to 100 µm. Each aerosol particle is characterised by its chemical composition, size and morphology. When the particles are non-spherical, there is no unique way to characterise the particles' size. In chemical transport modelling, the size of aerosols is usually expressed as the radius (or diameter) of a sphere of equivalent volume (or, equivalently, as the radius of a sphere of equivalent mass and with the same mass density). Aerosols with diameters up to 2.5 µm are commonly referred to as $PM_{2.5}$, which includes the modes of ultra-fine (≈0.001 µm to 0.1 µm) and fine particles (≈0.1 µm to 2.5 µm). This size range is particularly studied when investigating health impacts of aerosols, since these small particles strongly affect respiratory systems (Shiraiwa et al. 2017). Aerosols with diameters up to 10 µm are referred to as PM_{10} and also includes the coarse mode of particles, which is defined as the size range from 2.5 µm to 10 µm. Particles with even larger diameters quickly settle out of the atmosphere. Also, they do not penetrate deeply into the lungs. Therefore they are commonly not included when studying aerosols.

Aerosols in the atmosphere have three main sources, dispersion of materials from Earth's surface, emissions form high temperature processes such as biomass burning and volcanoes, and chemical reactions (Zellner 1999)). The two former sources emit aerosols directly into the atmosphere; those particles are referred to as primary particles. The chemical reactions produce secondary particles; this

includes particles formed by gas-to-particle conversion. Examples of primary particles are dust, sea-salt and black carbon. Dust particles are generated by wind erosion on surfaces, where deserts are dominant sources for mineral dust. Sea-salt is carried into the atmosphere by wave breaking over oceans, causing air bubbles to be formed which brake at the ocean surface, ejecting small drops of sea water into the air. The drops can subsequently dry out and produce sea-salt particles. Black carbon is emitted from combustion of biomass and fossil fuel. Secondary particles generated from gas-to-particle conversion emerge when the gaseous products have a low enough vapour pressure to condense. An example of such an aerosol formation is the oxidation of sulfur dioxide to ammonium sulfate, as shown in reaction scheme [III],

$$\mathrm{OH} + \mathrm{SO_2} \longrightarrow \mathrm{HOSO_2}$$

$$\mathrm{HOSO_2} + \mathrm{O_2} \longrightarrow \mathrm{SO_3} + \mathrm{HO_2}$$

$$\mathrm{SO_3} + \mathrm{H_{20}} \longrightarrow \mathrm{H_2SO_4}$$

$$\mathrm{H_2SO_4(g)} + \mathrm{NH_3(g)} \longrightarrow \mathrm{(NH_4)HSO_4(s)} \qquad \text{[III]}$$

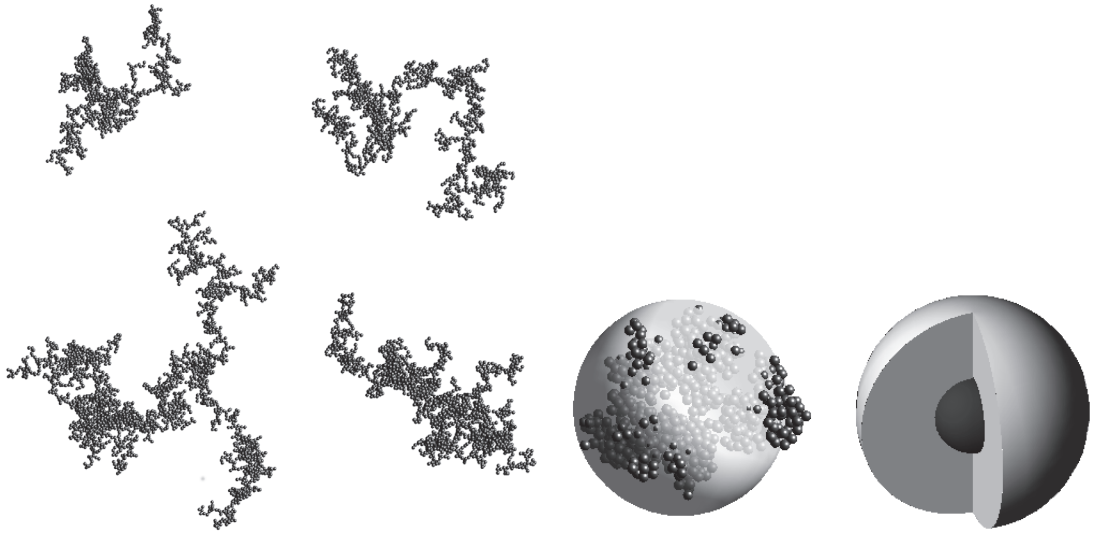
where sulfur dioxide (SO_2) oxidizes with a hydroxyl radical OH to form sulfuric acid H_2SO_4 in the presence of oxygen and water, and finally reacts with ammonia to form ammonium sulfate.

After being emitted or chemically formed in the air, aerosols undergo multiple dynamical, or micro-physical changes, which effect their size, shape, and chemical composition. The formation of aerosols by gas-to-particle conversion is also called homogeneous nucleation (formation of new phase), and is an example of a micro-physical change, where a phase change occurs spontaneously with the involved products. Heterogeneous nucleation occurs at a surface of typically another aerosol. Nucleation forms new particles with sizes smaller than 0.01 µm. As aerosols age in the atmosphere they change both chemically and physically. As a stable particle is formed, surrounding gas can condense onto the surface, causing it to increase in size and change its chemical composition. Further growth from collisions of particles are called either coagulation, when the collided particles merge together, or agglomeration when they collide and stick to each other.

A part of this thesis includes incorporating a more detailed description of black carbon and its micro-physical, into a chemical transport model and an aerosol optical model. Therefore, the following section describes black carbon and its sources, sinks and dynamical changes the atmosphere.

Black carbon (BC)

Black carbon (BC) is a primary carbonaceous aerosol emitted from incomplete combustion of hydrocarbons. Freshly emitted BC particles from combustion typically have shapes of fractal-like aggregates of carbon monomers (Kahnert 2010a, and references therein) as in Fig. 2.2a. Their sizes lie in both the ultra-fine and fine mode. The aggregates are hydrophobic upon emission, but as they age, they get oxidised and become more hydrophillic. Subsequently, vapors condense onto the aggregates and form liquid-phase coatings consisting of sulfate, organic material, and water.



(a) (b)

Figure 2.2: Black carbon as a fractal-like aggregate with four different numbers of monomers is depicted in 2.2a. A more detailed description of these particles can be found in Kahnert (2010a). Black carbon coated with some weakly absorbing liquid, such as sulfate, water, organic matter or dissolved sea-salt, is depicted in 2.2b, both as a realistic particle (left) and as a simplified core-shell geometry. See Kahnert et al. (2013) for more details.

BC aerosols do not make up a dominant part of all the total mass concentration of aerosols emitted into the atmosphere. According to an emission inventory study done by Klimont et al. (2017), 15% of global anthropogenic $PM_{2.5}$ emissions comes from black carbon, globally. However, they have a significant impact on the radiative warming of the atmosphere due to the high absorption cross section of BC aerosols. BC is therefore often referred to as light-absorbing carbon (LAC). Compared to molecules, it is more difficult to determine the optical properties of a particle, due to their varying chemical composition and complex morphologies. Comprehensive research is therefore dedicated to deduce dielectric properties (refractive index) and optical properties, such as the mass absorption cross section of BC. A review of modelling and measuring optical and dielectric properties of light absorbing carbonaceous particles is given by Bond and Bergstrom (2006).

2.3 Atmospheric Radiative Transfer

Atmospheric radiative transfer aims to describe the interaction between molecules, aerosols, or hydrometeors with the electromagnetic (EM) radiation propagating through the atmosphere. EM radiation is classically described by propagating electromagnetic waves of various wavelengths λ and frequencies ν, related by $\nu = c\lambda$, where c = $2.998 \cdot 10^8$ m/s denotes the speed of light. The classical picture of EM waves is employed when we describe, e.g., the interaction of radiation with particles, such as aerosols and cloud droplets. However, this classical models breaks down when attempting to explain phenomena such as the spectrum of black-body radiation. Such phenomena can be explained by

the quantum picture of EM radiation, in which the energy of the electromagnetic field can only be a multiple of elementary energy quanta $E = h\nu$, where $h = 6.626 \cdot 10^{-34}$ J $\cdot$ s is called Plank's constant. The quanta of the EM field are known as photons. The quantum nature of EM radiation is also essential when we consider the interaction of radiation with atoms and molecules. On the other hand, when we consider macroscopic media in which the scale of the system is much larger than the wavelength of light, then we can employ a much simpler model. We can describe EM radiation as travelling in straight lines, known as rays. Radiative transfer theory is based on this picture.

The main source of energy on Earth is provided by the sun and its incident radiation. An emission spectra of the sun can be described with Plank's law and black body radiation, with a temperature of about 6000 K, as seen at the top of Fig. 2.3. The largest part of the solar spectrum lies between 0.1 µm to 4 µm, and 90% of the total incident solar flux consists of visible (0.4 µm to 0.75 µm) and infrared (0.75 µm to 4 µm) wavelengths. The incident solar radiation is partly absorbed by the surface and the atmosphere and heats the planet. On the other hand, the surface and the atmosphere emit long-wave radiation in the thermal infrared at a range of 4 µm to 200 µm. The terrestrial spectrum can also, very roughly, be described as a black body curve with a temperature of 250 K (in reality the emission spectra from the Earth is highly variable and depends on latitude and altitude of present clouds), as seen in Fig. 2.3. The atmosphere is fairly transparent to a large part of the incident solar flux, whereas it is opaque to a large spectral portion of the outgoing terrestrial radiation. This creates the so called "greenhouse warming" effect. The transmission properties of the atmosphere are determined by scattering and absorption from gases, clouds, and aerosols. Absorption from important trace gases is also shown in Fig. 2.3 at ground level and at 11 km. The total attenuation of radiation, caused by the interaction with the atmosphere, is called extinction; it is the combined effect of absorption and scattering. The following subsections will describe the atmospheric absorption, emission and scattering in more detail.

2.3.1 Atmospheric Absorption and Emission

Absorption of electromagnetic energy by atoms, molecules or aerosols depends on their molecular structure or chemical composition (for aerosols). The energy an atom or a molecule can absorb or emit consists of both continuous bands and discrete lines. The total energy of an atom is made up of its kinetic energy and the potential energy of its electrons due to the Coulomb interaction among the electrons and the nucleus. Molecules have, in addition, vibrational and rotational energy states. While the kinetic energy is continuous, electronic, vibrational and rotational energies are discrete, resulting in spectra with discrete absorption and emission lines. These energies are specific to each atom and molecule and depend on their electron configurations and distinct vibration and rotation levels. The CO molecule is a diatomic molecule, which has one mode of vibration (symmetric vibration) and two rotation modes.Two of these energy levels corresponds to energies given by photons with wavelength 4.67 µm and 2.34 µm (infrared part of the solar spectrum). A tri-atomic molecule, such as O_3 has three vibrational and three rotational modes, giving rise to a larger set of wavelengths at which it can absorb radiation (Liou 2002). Atoms and molecules may undergo a transition to a higher energy state by absorbing electromagnetic

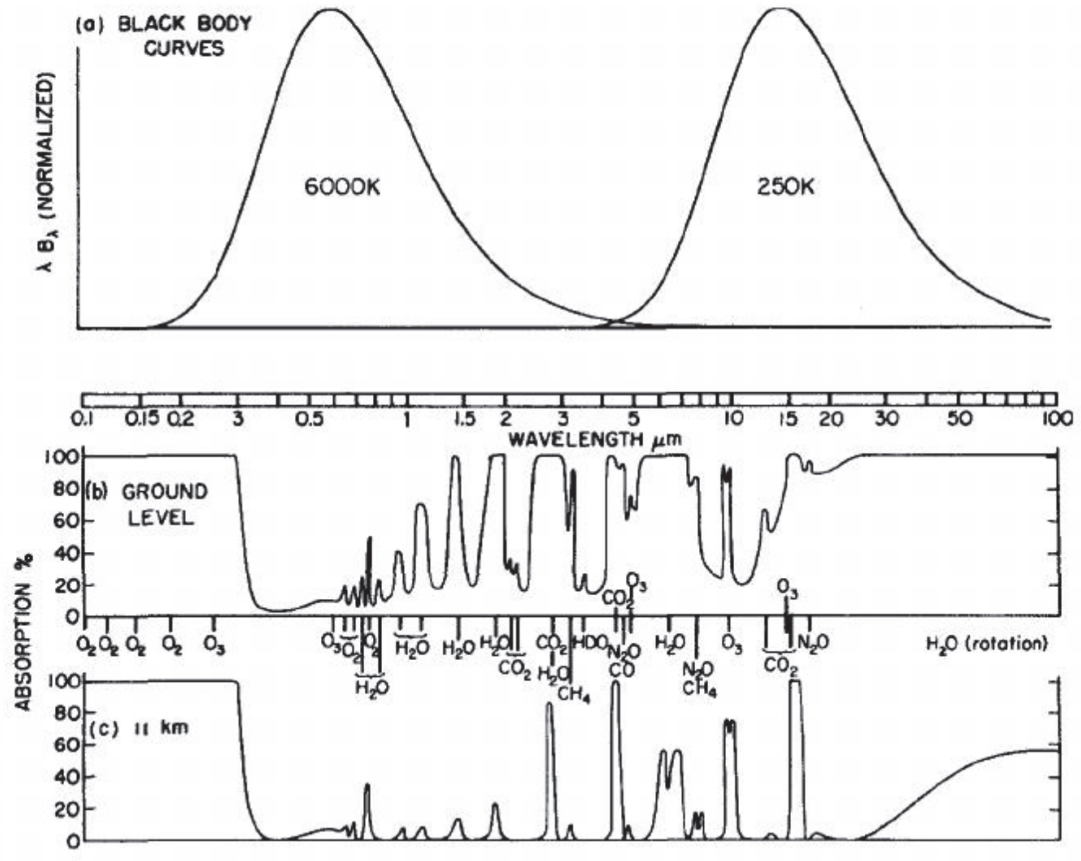


Figure 2.3: (a) Normalized black body spectra of the sun and Earth. (b) and (c) show mono-chromatic absorption of different chemical species at ground level and at 11km, respectively. Figure taken from Gooby and Yung (1995) with permission given by the Oxford Press University Inc.

energy. Similarly they may transition to a lower energy state, emitting the corresponding radiation. Apart from exciting a new energy state within an atom or molecule, radiation can also photo-dissociate (split the molecule into its constituent atoms) and photo-ionize (remove outer electron(s)) a molecule or atom, as illustrated in Fig. 2.4. To photo-dissociate or photo-ionize requires more energy than exciting a new energy level within a molecule or atom, which entails photons with shorter wavelengths as from the ultra-violet part of the solar spectrum.

Due to the discrete nature of absorption and emission of atoms and molecules, an absorption spectra could be expected to give a distinct line, but because of broadening processes, the line has a finite width. These broadening processes are cause by natural broadening (finite life time of an excited energy state), Doppler broadening due to the thermal motion of molecules, and pressure broadening (collisions between absorbing/emitting molecules/atoms inducing transitions between energy levels). The latter two are dominant in the atmosphere, Doppler broadening prevails in the upper atmosphere and pressure broadening in the lower atmosphere.

Aerosols differ from molecules and atoms in that their interaction with radiation can not solely be described by their constituent atoms, since they are in liquid or solid form. Instead, each particle can be treated as a macroscopic

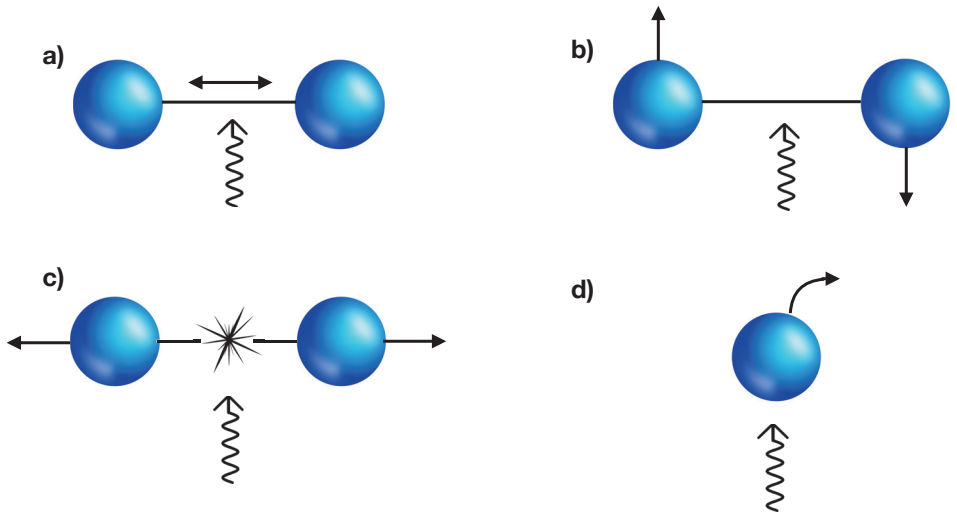


Figure 2.4: Illustration of how radiation (curly arrow) might interact with a molecule or atom. a) radiation is absorbed and excites vibration, b) radiation is absorbed and excites rotation, c) radiation photo-dissociates a molecule, and d) radiation photo-ionize an atom.

medium and the interaction be described with classical electromagnetic theory. The response of the medium to the incident electromagnetic field is then described using bulk dielectric properties, expressed by a complex refractive index ($m = m_r + i \cdot m_i$). Where a non-absorbing medium has a imaginary part of the refractive index $m_i = 0$ and the real part describes scattering. The refractive index for aerosols is however difficult to obtain from fundamental theories, such as solid-state physics. Therefore, it is common to derive refractive indices form observations. For atmospheric applications is is important to ensure that refractive indices taken from lab measurements have been obtained for aerosols representative for atmospheric aerosols. For instance, refractive index measurement for black carbon aerosols have been comprehensively discussed in this regard in Bond and Bergstrom (2006).

2.3.2 Atmospheric Scattering

Scattering is the process by which incident electromagnetic energy interacts with a molecule, or an aerosol, and is being re-radiated in all directions. This process is conservative, i.e., no net energy is lost during the interaction. Scattering is often described by a scattering cross section, σ_{s}, having the unit of m^2, representing the amount of incident radiation that has been removed from the original direction, in terms of an effective geometrical area. Atmospheric absorption can be described in a similar manner with an absorption cross section σ_{a}.

The total scattering cross section and the angular distribution of the scattered radiation depend on the size, shape and chemical composition of an aerosol, as well as the wavelength of the incident radiation. If the comparative size of a particle (molecule or aerosol) is smaller than the incident wavelength, then it scatters the same amount of energy into the forward as into the backward direction. This type of scattering is referred to as Rayleigh scattering and is illustrated in Fig. 2.5 together with Lorenz-Mie scattering for when the scat-

terer is about the same size and larger than the incident wavelength. The larger the scatterer becomes in relation to the wavelength, the more it scatters in the forward direction. When describing the scattering event, the scatterer is often assumed to be a homogeneous sphere. This is a good approximation for Rayleigh scattering, where the particles are small compared to the wavelength, but for larger particles, numerical methods need to be applied using parameters for the size, refractive index and the shape of the particle.

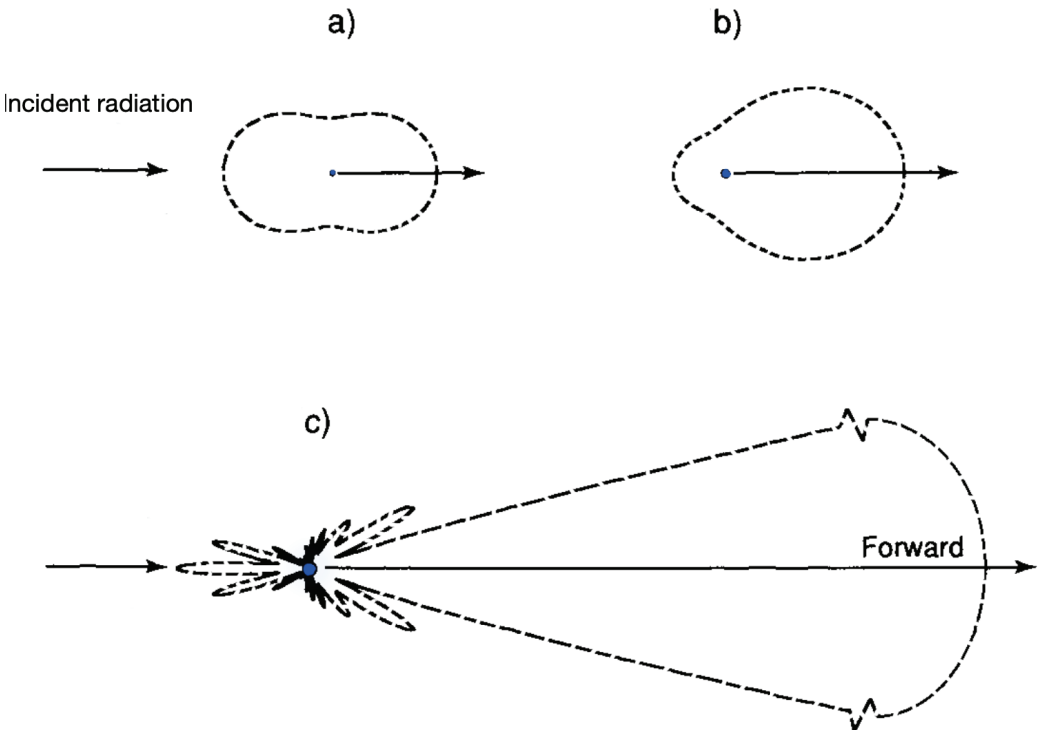


Figure 2.5: Scattering patterns from three different cases, where the particle (blue dot) a) has a much smaller size compared to the incident wavelength, and where the size of the particle is about the same size, b), and larger, c), than the wavelength . Illustration adapted from Liou (2002).

3

Observing the atmosphere

There are two main ways of collecting information about the atmosphere, either by using remote sensing techniques or by in situ observations.

Remote sensing techniques measure radiation that is emitted, absorbed, or scattered by the atmosphere. The measured radiative quantities carry information about atmospheric properties that can be retrieved from the measurement using inverse methods. A remote sensing instrument can either measure radiation in a passive or an active manner. Passive remote sensing methods detect scattered or reflected sunlight or thermal radiation emitted by the planetary surface or by the atmosphere. Instruments for active remote sensing have their own radiation source. To remotely sense the atmosphere and the Earth began in the 1820s with the invention of the camera obscura. Since then, aerial photographs have been used to map the Earth and more advanced instruments have been developed utilising other parts of the radiation spectrum, such as infrared and ultra-violet to map the atmospheric composition around the globe.

In situ measurement techniques collect information about atmospheric properties at a location where an instrument is located. Examples of such instruments are anemometers, barometers, thermometers and instruments taking air samples for compositional analysis.

This chapter continues by focusing on remotely sensed measurements from satellites and measurements done at the surface of the Earth and will specifically focus on the observations used in this thesis.

3.1 Satellite observations

Observing the atmosphere from space using satellites provides good spatial and temporal coverage, but is an expensive venture. The first satellite mission was launched 1957 with the satellite Sputnik 1, sent up by the Soviet Union to broadcast radio signals. This mission provided density information about the upper atmosphere by studying the drag, as well as information about the ionosphere from studying the path of the radio signals. The Sputnik 1 satellite orbited the Earth for three weeks before its batteries died. Since 1957, a myriad of satellite missions have been launched to study the atmosphere, using various

techniques to collect information about, e.g., trace gases, water vapour, precipitation, clouds, and aerosols. Modern satellites typically remain operational in orbit for several years.

The measuring technique, viewing geometry and orbital parameters of a satellite are chosen according to the desired measurand, resolution and ground coverage are, respectively. The following four examples present instruments and satellites measuring ozone, carbon monoxide and aerosol optical properties. Their measuring techniques, viewing geometries and orbital parameters will be explained. Data from all four satellite instruments are used in this thesis and belong to NASA's Earth observing system (EOS) developed for long-term monitoring of Earth's environment.

AIRS/Aqua The Aqua satellite was launched in 2002 and has six different instruments on board. The satellite orbit is sun-synchronous, near polar orbiting at an altitude of 705 km. These orbital parameters correspond to a low Earth orbit (LEO) and allows the spacecraft to get good ground coverage, one orbit around the Earth takes about 99 min and it covers the Earth in 16 days. Aqua was designed to operate for 6 years, but it is still providing data at the time of writing (Chahine et al. 2006). One of its six instruments is used in this thesis, the Atmospheric Infrared Sounder (AIRS).

AIRS consists of a passive hyper-spectral spectrometer and a photometer, viewing the Earth in a nadir cross-track viewing geometry. The spectrometer collects infrared radiation with a wavelength range from 3.7 µm to 15.4 µm using 2378 channels, and the photometer has four channels in the visible/near-infrared part of the spectrum, 0.1 µm to 4 µm. The broad range of wavelengths makes it possible to retrieve information about the terrestrial surface, clouds, vertical profiles of water vapour, temperature, ozone, carbon monoxide and methane. With a cross-track scanning swath of about 800 km at either side of the ground-track, it provides, for the IR-channels at nadir, a spatial resolution of 13.5 km and the visible/near-infrared a resolution of 2.3 km.

Clouds can have a large impact on what a satellite instrument observes. In the visible part of the spectrum, radiation is scattered by clouds and in the near infrared part of the spectrum, radiation is strongly absorbed by clouds. Longer wavelengths for example microwaves ($\approx$1 cm to 100 cm) can better penetrate cloud and give information about, in and under the cloud. This needs to be taken into account when using satellite retrieved products. AIRS data products have therefore jointly been retrieved by also using data from two microwave instruments on board Aqua, making it possible to filter out some effects from clouds. As a result, only measurements of cloud-free skies are being used to derive the different data products. Effects of clouds in satellite products are not explicitly taken into account in this thesis, since long-term averages have been used. The products used from AIRS in this thesis, are ozone and carbon monoxide concentrations.

MOPITT/Terra The Terra satellite was launched 1999 and has a sun synchronous near polar orbit at 705 km above the surface, revolving the Earth in 98 min. Terra was the first satellite within the EOS mission and it was designed to provide data for about five years, but is still producing data at time of writing. Terra has five instruments on board, one of which are used in this thesis, the

Measurement of Pollution in the Troposphere (MOPITT) instrument. MOPITT observes up-welling infrared radiation through a nadir cross-track scanning instrument at three narrow wavelength bands, 2.2, 2.3 and 4.7 µm. Radiation received by the two latter bands are dedicated to observing carbon monoxide, where the 4.7 µm band is sensitive to mid- and upper atmospheric CO and the 2.3 µm band is sensitive to the whole column of CO. The former band is used to collect information about methane. The instrument uses gas correlation spectroscopy, a technique utilising a cell of the target gas as an optical filter (Drummond et al. 2010). At nadir MOPITT has a spatial resolution of 22×22 km horizontally and about 4 km to 5 km vertically and the swath created by the scanning is about 640 km. The data products used in this thesis is version 6 of monthly mean data (Level 3) of carbon monoxide. This data product is described and evaluated in Deeter et al. (2014).

OMI/Aura The Aura satellite was launched in 2004 and flies in a train formation with the Aqua satellite, 15 minutes behind. Aura has four instruments on board, and one of them is the Ozone monitoring instrument (OMI) that used in this thesis. OMI is a passive, nadir across-viewing non-scanning instrument measuring backscattered UV and visible radiation with two UV bands, 270 nm to 314 nm and 306 nm to 380 nm, and one visible band, 350 nm to 500 nm. The radiation is received in a wide-field telescope, with an angle of 115° and a swath of about 2600 km, feeding two image grating spectrometers giving rise to a resolution of 13×24 km. The data products derived from the measured radiance are a range of important pollutants, mostly related to ozone and its sources and sinks. Gridded orbital data products of ozone, version 3, (OMO3PR) are used in this thesis. These products are described and validated in Kroon et al. (2011).

CALIOP/Calipso The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (Calipso) satellite was launched in 2006 and has similar orbital parameters as Aqua and Aura, since it flies in a train formation with these two, as well as with three other satellites. Calipso carries three instruments, where the instrument Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) has been used in this thesis. CALIOP is an active sensor, meaning it has its own laser source. CALIOP can therefore observe the atmosphere both at daytime and nighttime. It transmits linearly polarised laser light at 532 nm and 1064 nm and measures range-resolved backscattered intensities. Compared to passive instruments which commonly use a cross-track configuration, a space-borne lidar such as CALIOP has a very narrow field of view (130 µrad), resulting in a much smaller swath of 61 km and a horizontal resolution of 1 km. This means it takes much longer time for CALIOP to cover the whole Earth, giving rise to low horizontal resolution for gridded retrieval products. Compared to most passive sensors, CALIOP has a very high vertical resolution of 30 m, due to its tracking of the time of the backscattered laser beam in the atmosphere.

CALIOP retrieves information about clouds and aerosols and with its high vertical resolution, these retrieval products are good candidates for use in data analysis/assimilation studies when vertical information of clouds and aerosols are desired. In this thesis, aerosol extinction at 532 nm is used (Calipso Level 3, version 3). More specifically, a gridded monthly mean is used, which is quality

screened and used for all sky conditions (all cloud occurrences: below optically thin clouds, in clear skies and above clouds) and is averaged based upon daytime data. This data product was recently evaluated against EARLINET data (a European ground based measurement network, measuring aerosol optical properties) by Papagiannopoulos et al. (2016), where the CALIOP aerosol extinction product performed fairly well, both close to the ground and further up. The latest quality screening used in version 3, was also investigated in a study done by Tackett et al. (2018) and showed a bias reduction of 24 % and 31 % on global ocean and land AOD (which uses vertically integrated aerosol extinction) products, respectively.

3.1.1 Retrievals

All of the above mentioned data products rely on a forward model describing the relationship between the physical atmospheric constituents and their observed radiative properties. A forward model is directly used if a mapping of a physical atmospheric state to observed radiance is desired. If the atmospheric state is required, a retrieval algorithm is employed. A retrieval algorithm finds a solution to an ill-posed and often under-constrained inverse problem. The problem is under-constrained, because multiple atmospheric states can give rise to the same observed radiance. Therefore in order to solve this problem, a probabilistic approach is used. Such methods rely on a forward model and on an a priori estimate of the atmospheric state. A priori knowledge of a desired atmospheric state is commonly derived from models or other measurements and serves as a "first guess".

A common retrieval algorithm used by, e.g., the OMI retrievals and MOPITT, is Optimal Estimation, described by Rodgers (2000), which derives an atmospheric state from a conditional probability given the measurement and prior knowledge of the atmospheric state, using Gaussian statistics and Bayes theorem. The retrieval then becomes a weighted average of the a priori and the information contained in the measurement. It does not lie within the context of this thesis to further investigate the different retrieval algorithms, but it is important to mention their impact on the retrieved products used in this thesis. When comparing retrieved data products with for example model outputs, it becomes important to know the relative effect an a priori has on a resulting retrieval, i.e., does the retrieval mostly reflect the measurement or the a priori. If Optimal Estimation is used to retrieve an atmospheric state, Eq. 3.1 describes how much of an a priori is weighted in a retrieval compared to the measured quantity,

$$\mathbf{y}_{rtv} = \mathbf{y}_a + \mathbf{A}(\mathbf{y}_m - \mathbf{y}_a) \tag{3.1}$$

where $\mathbf{y}_{rtv}$ is the retrieval, $\mathbf{y}_a$ is the a priori, $\mathbf{y}_m$ is the measured target and $\mathbf{A}$ is an averaging kernel. Each row of $\mathbf{A}$ corresponds to a function describing how the difference between the measured and a priori information is smoothed vertically. The functions therefore determines the degree of smoothing of a retrieval. This error will be diminished for an ideal retrieval, i.e., the averaging kernel tends towards an identity matrix, or if $\mathbf{y}_m = \mathbf{y}_a$ (Deeter et al. 2012, Kroon et al. 2011). An additional error term ϵ is commonly added to Eq. 3.1 to add the errors from the retrieval process, other than the error arising form smoothing.

This retrieval error arise from measurement uncertainties and is often seen as random and can be removed by averaging, unlike smoothing errors. Further descriptions of uncertainties and errors follow in the next section.

3.1.2 Satellite errors

An error is defined by a deviation from a true state, yet a true state can never be found. Therefore, we have to estimate errors by error statistics. When deriving error statistics, the potential error sources are divided into random and systematic errors. Random errors have zero mean when averaged over a sufficiently long time interval. Systematic errors have a non-zero mean when averaged over time. Quantifying errors is essential for measurements, since it provides a gauge for comparisons, and it is required when using measured and retrieved data in data assimilation/analysis (a common application for satellite remote sensing data). Depending on the instrument and forward model, methods for deriving error estimates vary. For example, MOPITT first performs a radiance bias correction and then calculates an error covariance matrix. The bias correction is done to remove possible systematic measurement errors from instrumental specifications, forward model errors, spectroscopic errors, and geophysical errors. The error covariance matrix is thereafter derived by assuming random errors from the instrument and the forward model. The instrument error is defined as a product between the standard deviation of an un-calibrated measurement and a gain factor. The forward model error is calculated from an expectation value of the difference between the used forward model and another forward model (Francis et al. 2017).

Errors in satellite retrieved data come from two major sources, the instrument itself and the retrieval process. Measurement or instrument errors can arise from instrumental noise, calibration errors, inaccuracies from pointing the instrument's line of sight or degrading instrument parts. For example, CALIOP has been plagued with calibration errors from previous versions of their data products. Kacenelenbogen et al. (2011) investigated the sources of errors in CALIOP's retrieved data and found from an extensive validation study that parts of large errors were likely to come from calibration procedures. CALIOP data products were later improved in their version 4, by Kar et al. (2018). Another example of errors generated from degrading instrument parts was reported for OMI, which is referred to as the "row-anomaly". This anomaly is believed to be caused by a loss of thermal insulation in the instrument's entrance slit, resulting in measurement errors at all wavelengths for a specific cross-track viewing direction (Bak et al. 2015).

Retrieval errors come from the assumed forward model as well as the assumptions made regarding the a priori knowledge used. Most of these errors are systematic, arising from "known un-quantifiable unknowns", such as errors from cloud filtering and assumptions regarding the a priori, as well as "unknown unknowns", exemplified by variability on scales that are smaller than the resolution of the observations (Povey and Grainger 2015). Although this thesis does not explicitly focus on satellite errors, these errors are taken into account when evaluating models and when performing data analysis. Povey and Grainger (2015) reviewed errors and uncertainties in satellite remote sensing data and provides a good overview of how measurement and retrieval errors arise, and how they can be quantified.

3.2 Ground based observations

Observing the atmosphere form the surface of the Earth provides both the satellite and model community with validation and evaluation possibilities. The advantage of ground based observations over satellite observations is that the cost of instruments and installation is much smaller, the maintenance and quality of the data can be assured more easily, while the temporal resolution can compete with that of satellite instruments. Depending on the desired atmospheric constituent, the measurement techniques varies. Both in-situ and remote sensing techniques are used from the surface. For measuring ozone it is common to use spectroscopic methods studying the absorption in the UV or visible part of the solar spectrum. For measuring carbon monoxide one often uses direct sampling of the air followed by an analysis by, e.g., gas chromatography or spectroscopic methods. Aerosol optical properties are commonly retrieved by using sun-photometers, which measure direct sun radiance. Aerosol mass concentrations can be determined by analysing the amount of particles deposited onto a filter. Examples of analysis methods are gravimetric, optical, or microbalance methods. A gravimetric method weighs the filter before and after it has been subjected to a particle sample. An optical method uses attenuation of light to determine the particle amount. Micro-balance methods measures mass concentration by an oscillating collection substrate which will change its oscillation frequency depending on the mass collected on the substrate.

This thesis uses ground based observations to directly evaluate model results. Measurements have been obtained from two major networks, the Global Atmospheric Watch (GAW) and its world data center for greenhouse gases (WDCGG) and the European Monitoring and Evaluation Programme (EMEP).

GAW - WDCGG The GAW programme is a World Meteorological Organisation (WMO) driven programme, providing information about global air pollution by maintaining a global network of long-term observations of atmospheric chemical composition. With several hundred measurement stations spread over the whole globe, it collects data to seven main data centers for aerosols, reactive gases, greenhouse gases and stratospheric ozone and UV, solar radiation, precipitation chemistry, and remote sensing of the atmosphere. In this thesis data from the greenhouse gas data center of surface ozone is being used. All data came were measured with UV absorption spectroscopy.

EMEP EMEP is a co-operative programme for monitoring and evaluating air pollution. Its main focus is on long-range transported air pollutants over Europe. It consists of data bases for emissions, measurements, and model results that can be used to support policy-making for mitigating hazardous effects from these pollutants. Measurement data for total mass of PM_{10} was used in this thesis for 52 European stations, but various instrument types where used to collect this information. Among the most common instrument types where high and low volume sampler, as well as filter 1-pack and 3-pack, which all use a gravimetric analysis of the content. Two other common instrument types among the used stations where beta attenuation, which determines the particle mass by attenuation of β-radiation , and tapered element oscillating microbalance (TEOM), which is based on a microbalance method.

3.2.1 Measurement uncertainties

Just as satellite observations, ground based observations need to report corresponding uncertainties to their measurements. The sources of errors are similar to the ones in satellite instruments. However, instrument errors arising from calibration or instrument failure can easier be detected, since the instrument is readily accessible on the ground.

Uncertainties from individual measurement stations need to be reported according to different standards/guidelines, depending on the network it collaborates with. Measurements of ozone from GAW-WDCGG need to be done by the same instrument type, they need to undergo a given calibration procedure, and they need to apply a common retrieval method (Galbally and Schultz 2013). Guidelines for estimating the uncertainty is also given, in addition to data submission filtering. EMEP measurements for total mass of PM_{10} have similar guidelines, but do not require a common instrument type, even though it is recommended to use an instrument with a microbalance method (Berg et al. 2014). Also, there is a European standard for measuring PM_{10} with the gravimetric method, EN 12341:2014. In this thesis the reported uncertainties were used in terms of flagged data. Flagged data points (non-zero flag) correspond to data with a defined uncertainty, and were filtered out before usage (`https://ebas-submit.nilu.no/Submit-Data/List-of-Data-flags`).

4

Chemical Transport Modelling

A chemical transport model (CTM) is a mathematical representation of the atmospheric chemical composition, which in general is affected by four processes, emission, transport, chemistry and deposition. These processes are numerically represented to simulate the variability of the chemical composition of the atmosphere by solving the continuity equation. As input parameters, the model needs initial concentrations, meteorological data (to represent transport and wet deposition as well as calculate chemical reaction rates), as well as emissions and land use data. Limited-area models also need boundary values.

There are a wide range of applications for CTMs, such as air-quality forecasting, determining regional and global budgets of chemical species in the atmosphere, serve as a comparison to observations in order to describe the underlying processes, or to provide a priori knowledge of the atmosphere to satellite retrievals. CTMs further allow us to map deposition of eutrophying and acidifying components and base cations to land and water surfaces. CTMs are also used to integrate large samples of observations to find a more optimal estimation of the atmospheric state (chemical data assimilation/analysis). This chapter will introduce the concept of a CTM, present its underlying continuity equation, and explain how the different atmospheric processes can be represented. The chapter will also discuss model uncertainties and the specific CTM used in this thesis.

4.1 The continuity equation

The continuity equation describes the mass conservation of a chemical species and is commonly described as a change in number density or concentration (molecules per cm^{-3}) in the context of finding the chemical atmospheric state. The change of concentration of a chemical in a given volume is affected by its transport in and out of the volume, i.e., its flux vector $\mathbf{F}$, and the production and losses inside of the volume. The mass flux $\mathbf{F}$ is affected by wind-driven transport (advection and turbulence) and molecular diffusion. However, in the lower part of the atmosphere, molecular diffusion is very slow compared to the advection

and turbulence, if looking at transport scales larger than 1 cm. It takes about one month for a molecule to get transported 10 m by molecular diffusion, compared to having horizontal wind speeds, $\mathbf{u}$ of 1 m/ sec to 10 m/ sec and vertical wind speeds, $\mathbf{v}$, of 0.1 m/ sec to 1 m/ sec. Therefore having winds speeds of orders of magnitude larger than the molecular diffusion in the lower atmosphere, the mass flux of a chemical species is assumed to only be affected by advection and turbulence. The local production is represented in terms of emissions and chemical production, whereas the loss term is represented by chemical losses and scavenging (wet or dry deposition). If we assume an Eulerian framework, in which we have a fixed frame of reference at which the mass concentration changes with time, the continuity equation can be described as in Eq. 4.1,

$$\frac{\partial c_i}{\partial t} + \underbrace{\nabla \cdot (\mathbf{u} c_i)}_{transport} = \underbrace{R_i(c_1, c_2, \ldots, c_n)}_{chemistry} + \underbrace{E_i}_{emission} - \underbrace{S_i}_{sinks} \tag{4.1}$$

where $\frac{\partial c_i}{\partial t}$ is the concentration of a chemical species varying in time at a fixed position, $\mathbf{u}$ is the horizontal velocity vector ($\mathbf{u} \cdot c_i = \mathbf{F}$), R_i describes transformation among species by chemical reactions and microphysical processes (for aerosols), E_i correspond to emissions and S_i the sinks in terms of gravitational settling (for aerosols) and wet and/or dry deposition, (Seinfeld and Pandis 2016). The continuity equation in its Eulerian form is a first order partial differential equation in space and time. A solution can be found by integration, given initial conditions, spatial boundary conditions and discretization of the problem. Initial conditions can be obtained from interpolation of boundary conditions or given by previous model runs. Spatial boundary conditions will be discussed in Sect. 4.1.5. Discretization is done both spatially and temporally.

Spatial discretization is done by a finite number of gridboxes/gridpoints. In the horizontal directions one commonly employs a latitude-longitude grid. In the vertical direction, one uses a pressure coordinate, typically a σ-coordinate, defined by eq. 4.2,

$$\sigma = \frac{p - p_T}{p_S - p_T} \tag{4.2}$$

where p is the pressure at the level at which σ is defined, p_S and p_T correspond to the surface and top pressure of the model domain, respectively. By definition, the σ-coordinate follows the terrain close to the surface and can attain values between 1 at the surface to 0 at the top of the model domain. However, numerical errors are known to be introduced when using σ-coordinates (Auclair et al. 2000, and references therein). Therefore a hybrid-sigma coordinate system was introduced to minimize this error, which is a linear combination of σ-levels and pure pressure levels, see Eq. 4.3.

$$p_j = \alpha_j + \beta_j \cdot p_S \tag{4.3}$$

where α_j and β_j are derived values corresponding to how quickly the model vertical levels are transformed from σ close to the surface, to isobaric pressure levels higher up. α_j and β_j used in CTM, often corresponds to the same as used in the parent weather prediction model.

To find a solution to the continuity equation, the equation is split into operators, solving the transport, chemistry, emission and deposition term separately.

This is done under the assumption that the coupling between the different terms can be neglected over certain time steps. Where the operator associated with the shortest time step acts first on the initial concentration, and the operator with the longest time step acts last. For example, the largest time step is associated with the reading of the meteorological data and boundary conditions, which commonly is read every 3^{rd} to 6^{th} hour and interpolated to every 1 hour. Then the mean advection time step can be in the range of 10 min to 60 min and the chemistry time step around 2 min. There are, however, also other methods for applying operator splitting, but the subject of how numerically solve the continuity equation is not covered in this thesis.

Figure 4.1 gives a schematic overview and a flow chart of how the different processes in a CTM are coupled together in solving the continuity equation. The transport operator and its implications for solving the continuity equation is described in Sect. 4.1.1, and the chemical reaction term R_i, which describes the chemical production and losses of chemical species, is described in Sect. 4.1.2. The emission source term and the deposition sink term are described in Sect. 4.1.3 and Sect. 4.1.4, respectively.

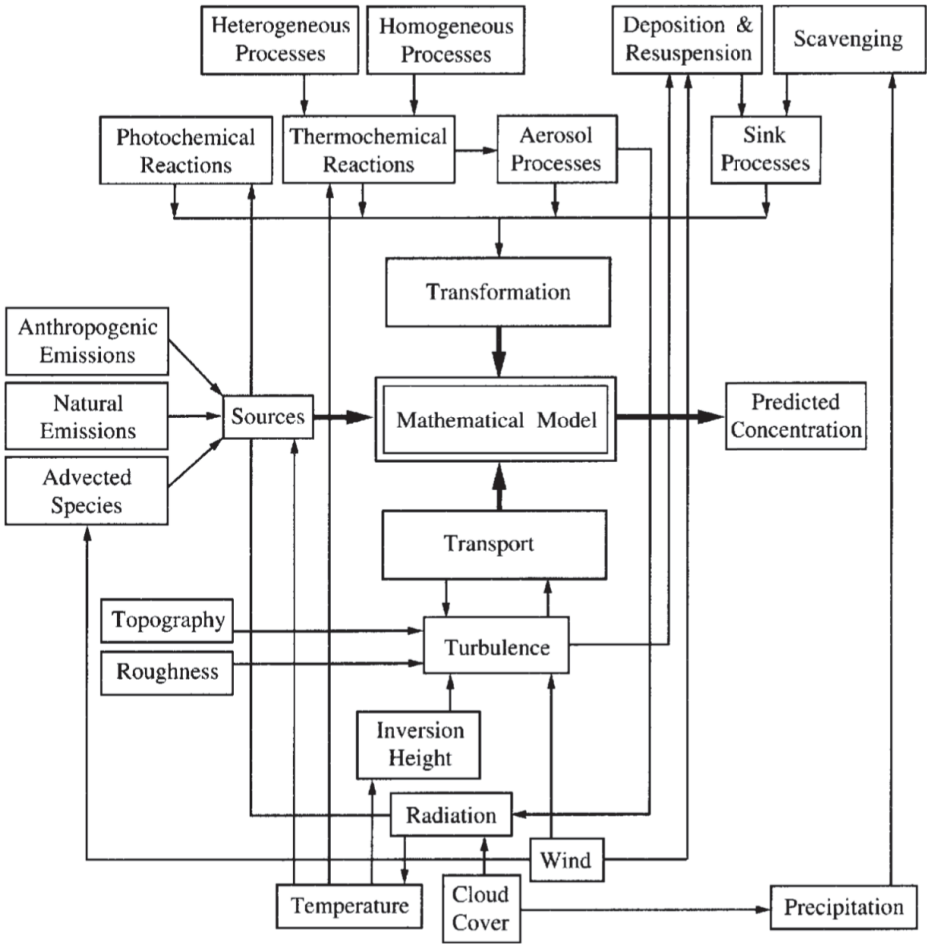


Figure 4.1: Schematic overview of processes involved in a mathematical chemical transport model. Originally illustrated by Seinfeld and Pandis (2016) and permission to use was given by John Wiley and Sons Ltd, Copyright(2016)

4.1.1 Transport

To solve the transport part of the continuity equation, information about the wind speed is needed. Wind speeds characterise the movement of the atmosphere; the fluid motion can often be highly chaotic, or turbulent. To account for turbulence in the wind speed, it is divided into two components, a mean and a turbulent component. Information about the mean wind component is given by meteorological input data, frequently derived from a numerical weather prediction model, while the turbulent component is derived from a parameterisation scheme.

For turbulent flow within a gridbox, an eddy diffusion parameterisation is used, which assumes turbulent motion to be random; it is simulated in the same manner as molecular diffusion, with Fick's law. Hence the mean concentration of a chemical species is proportional to a turbulent diffusion parameter, $\mathbf{K}$, which describes the strength of the turbulence and depends on the height from the ground, mean wind shear, and surface heating by the sun (Wallace and Hobbs 2006). Turbulent energy motion is generally much larger in the vertical and is therefore often assumed to be the only contributor to the turbulent transport of chemical species in a CTM. Since eddy diffusion only describes a change in concentration locally, it can not be used for wet convective transport, which is caused by a release of latent heat from a stable air parcel, causing rapid cloud updrafts and affecting several vertical gridboxes. If accounted for, a convective parameterisation scheme needs to be used for this type of turbulent motion, commonly adapted from the parent numerical weather prediction model.

4.1.2 Chemical and physical transformations

A chemical reaction scheme is defined by a set of coupled differential equations, including photo-chemical reactions, gas- and aqueous-phase chemistry, as well as aerosol micro-physical processes. The change of each included chemical component is described as a differential equation depending on other species concentrations proportional to each reaction rate. Reaction rates for gas- and aqueous-phase chemistry are commonly derived from laboratory experiments and may depend on temperature and pressure. Photo-chemical reaction rates are, on the other hand, derived by calculations using absorption cross-sections and quantum yields. If aerosol micro-physical processes are included in a model, each chemical component (e.g., black carbon, organic carbon, sea-salt, sulphate, nitrate) have to be represented with a size distribution. Two common methods of representing aerosol size distribution in 3D chemical transport models are the modal and sectional method (Seinfeld and Pandis 2016). A modal method describes a a size distribution with a set of modes, one mode for each sub-size range. A mode is a separate size distribution function, described by a log-normal distribution. A sectional method divides the size distribution into discrete size bins.

The chemical operator of the continuity equation for each size bin/mode and chemical component is then represented in a similar manner as each gas component, except for addition of terms for nucleation, condensation, and coagulation. Section 4.3 will exemplify a chemical reaction scheme from the CTM used in this thesis.

4.1.3 Emission

A defined spatial grid of a CTM has lateral, top and surface boundaries that need specified in- and outflows in order to solve the continuity equation. The lateral and top boundary conditions will be discussed separately in Sect. 4.1.5 and the outflow to the surface, referred to as deposition will be discussed in Sect. 4.1.4. The inflow boundary at the surface are characterised by emissions from the surface. Nevertheless, emissions are not only limited to the surface, for example, air traffic and chimney emissions injects pollutants above the surface.

Emissions can be treated both as a pure input to a model or be calculated online. Input emissions are often given as a total amount of a chemical component (gas or aerosol) per sector and time period, where a sector corresponds to a group of emission sources, such as agriculture, shipping, road traffic or combustion plants. Reported time periods are often months or years, and emission inventories are commonly reported by individual countries. Some emissions depend upon meteorology or land use and need to be calculated within a CTM. Emissions of, e.g., wind-blown dust, volatile organic compounds (VOCs), and sea-salt, are dependent upon meteorology, while biogenic emissions and emissions of NH_3 from livestock depend also upon the land use.

The accuracy of emission inventories vary depending on the source region and on the geographical area from which the emissions are reported. For example, large point source emissions are commonly subjected to monitoring, and have therefore well characterised emissions. Most other emission sources are highly dispersed, spatially and/or temporally. Emissions from road traffic might be derived from well-known road networks and based on good approximations regarding the amount of exhaust, but have large uncertainties associated with estimating the flow of cars. Natural emissions from wild-fires and anthropogenic emissions from residential burning are exceedingly variable and hence difficult to estimate. Some geographical areas also have no or little jurisdiction to report on certain emissions, giving rise to large discrepancies in their reported inventories. As a result, emissions are one of the largest contributors to model uncertainties.

4.1.4 Deposition

Deposition is divided into dry and wet deposition. Dry deposition describes the uptake of trace gases and aerosols at the surface, and wet deposition describes the scavenging of water soluble gases and aerosols by precipitation.

Dry deposition can be implemented in a model by either a one- or two-way exchange, i.e., as an irreversible or reversible process where a chemical component disappears from the model or can get re-emitted into the model domain. Both exchanges are defined by a flux rate towards the surface, which is dependent upon surface concentration and the reactivity of the surface. A one-way exchange depends on the vertical turbulent motion and the characteristics of the surface, while a two-way exchange also depends on the conductivity of the ground. The above surface flux, is modelled by assuming a surface layer height in which the vertical turbulent motion and the reactivity of the surface is characterised by a dry deposition velocity, assuming a dry deposition flux rate linearly dependent upon the number density of a chemical component. The dry deposition velocity is parameterised and commonly described by resistances, an aerodynamic resistance based upon the vertical turbulent motion, a boundary

layer resistance dependent upon molecular diffusion and a surface or canopy resistance characterised by the surface reactivity. A two-way exchange would, in a similar manner, describe a flux from its ground reservoir with a flux rate dependent upon a certain depth and conductivity.

Wet deposition is separated between different precipitation processes as well as in- and below-cloud. If a CTM has a convective parameterisation scheme or if the meteorological input data contains information about the convective precipitation, wet deposition form convective scavenging can be derived. Large scale precipitation involves scavenging in and below clouds and depends upon the precipitation rate and the solubility of the chemical components.

4.1.5 Boundary conditions

Lateral and top boundary conditions are important to represent large-scale transport of trace gases and aerosols in regional scale models. The impact on in-domain concentrations can be large depending on the life time of a chemical species. Boundary conditions can be derived from other model simulations and observations from satellites, air-crafts or ground based measurement stations. A common model setup is to use prescribed boundary conditions (from either models or observations) which are interpolated and scaled spatially and temporally. However, this setup is known to introduce uncertainties, since it can not capture episodic emission events outside the model domain, such as wild-fires, volcanic emissions and dust-storms. To better capture temporal and spatial variations outside the model domain, larger scale models or satellite observations are used, but this will cause the regional model to inherit potential biases from the parent model or the satellite observations. The choice of boundary conditions is therefore vital in the pursuit of improving overall model performance.

4.2 Model errors

As stated in section 3.1.2, an error is a deviation from an unknown true state. Wherefore, model errors have to be estimated by analysing assumptions and inputs given to a model. Errors in a CTM come from mainly three sources, i) parameterisation schemes that describe sub-grid processes, such as turbulence, or processes which are described by empirical relations rather than from first principles, such as biogenic emissions or dry deposition of trace gases, ii) numerical approximations arising from different solvers, grid sizes, time steps or amount of chemical species, iii) input parameters such as emission inventories, boundary conditions, and meteorology. Due to the non-linearity of a CTM it is hard to determine the effect of each source on the model output. An overall model performance can be quantified by comparing model results to observations, but does not give you any information about how the discrepancy is coupled to a potential error source (model evaluations are discussed more in section 5.1). Further, sensitivity studies can be performed to discern the effect of different input parameters, solvers, or parameterisations on the model result, and to obtain an estimate of the model errors. Another method of finding possible error sources is to study ensembles of chemical transport modelling systems and compare them with observations. The derivation of a model error covariance matrix, that is needed for a data analysis/assimilation is briefly mentioned

in section 5.2.1.

4.3 MATCH

The Multi-scale Atmospheric Transport and CHemistry (MATCH) model is a three-dimensional Eulerian chemical transport model developed at the Swedish Meteorological and Hydrological Institute (SMHI) (Robertson et al. 1999). MATCH is used to forecast air quality, study effects from different emissions scenarios and can be used on multiple scales for local, regional and hemispheric chemical transport studies. MATCH also includes a module for 3D variational data assimilation/analysis (see Sect. 5.2.1), and an aerosol dynamics model describing the aging of aerosols in the atmosphere, called Sectional Aerosol module for Large Scale Applications (SALSA) (Kokkola et al. 2008).

Transport is described by meteorological input and by a vertical turbulent parameterisation scheme assuming eddy diffusion or a first-order approximation of the flux intensity based upon mixing-length theory. A wet convective transport scheme based upon the work done by Tiedtke (1989) is implemented, but not used in this thesis. The numerical solver for the advection (mean wind and the turbulent flow) is of Bott-type (Bott 1989). A full description of the transport can be found in Robertson et al. (1999).

The gas-phase chemistry is adopted from the European Monitoring and Evaluation Programme Meteorological Synthesizing Centre West (EMEP MSC-W) which is described in Simpson et al. (1993, 2012), with modifications regarding the isoprene chemistry. Aqueous-phase oxidation of SO_2 is also included together with a few heterogeneous reaction for nitrogen compounds and secondary organic aerosol formation from monoterpenes and isoprene. In total the chemical reaction scheme includes about 140 thermal, wet and photo-chemical reactions involving about 60 different chemical species. The aerosol dynamics model SALSA is adopted to MATCH and is a sectional model calculating mass and number concentrations of aerosols. The set-up of SALSA used in this thesis consists of 20 size bins. Each size bin represents a range of particle volume equivalent radii, having a constant volume ratio. The total number of size ranges are ten, reaching from 1.5 nm to 5000 nm. The reason why the size ranges are different from the amount of size bins, is because SALSA accounts for both external and internal mixing of particles. A more detailed description of the chemistry within MATCH-SALSA is found in Andersson et al. (2015).

Emission inventories correspond to both biogenic and anthropogenic emissions sources. Sea-salt and isoprene emissions are calculated online, whereas the rest of the emissions are given as input. The model setup used in this thesis use emission inventories from EMEP which are given for each species, such as black carbon, organic carbon, dust, CO, etc., per month and sector. 12 sectors are included in the model, see table 4.1.

Deposition is modelled as a one-way exchange. Dry deposition of trace gases and particles are modelled with a resistance approach. Effective dry deposition velocities are calculated with seasonal and diurnal varying prescribed velocities together with aerodynamic resistance and land types. Particles also have a an additional term for gravitational settling. There exist different schemes for wet deposition in MATCH, but in for MATCH-SALSA, most trace gases are assumed to have the same proportionality to precipitation in and below clouds. But SO_2,

Table 4.1: Table of the 12 different emission sectors used in MATCH.

Sector	Name
1	Combustion in energy and transformation industries
2	Non-industrial combustion plants
3	Combustion in manufacturing industry
4	Production processes
5	Extraction and distribution of fossil fuels and geothermal energy
6	Solvents and other product use
7	Road transport
8	Other mobile sources and machinery
9	Waste treatment and disposal
10	Agriculture
11	DMS (Dimethyl sulfide)
12	Shipping

O_3 and H_2O_2 has an in-cloud we scavenging depending upon their solubility. Particle wet scavenging varies in and below a cloud, and is dependent upon the size of the particle, as well as if it is activated (to serve as a cloud condensation nuclei) in a cloud.

5

Atmospheric observations and CTMs

Observations and models both provide useful information about the atmosphere. Models provide a complete picture of the atmospheric state, but their errors grow rapidly in time, unless constraints are provided. Observations provide an incomplete description of the atmosphere, but the uncertainty in the observations is generally lower than that in the model results. A main subject of this thesis is to make combined use and exploit the advantages of both sources of information. This is done in two ways, (i) by evaluating models with observations, and (ii) combining observations and model results by use of data analysis.

5.1 Model evaluation

Assessing performance of models is important for quantifying uncertainties and for improving our knowledge of different processes. Evaluation of models can be done by comparing model results to observations. In this thesis observations from both satellite and ground-based observations are being used for model evaluation.

A large challenge in comparing models with observations is to make sure the data sets represent the same quantity. Models have a finite grid resolution and time step, whereas observations are done at a set time and geographical place. Therefore, the data sets need to be collocated in both space and time. This can be difficult depending upon the compared quantity and the spatio-temporal resolution, which can affect the representativeness of a comparison. Constraining criteria have to be defined for whether and when an observation and a model can be compared, in order to make a robust analysis. First, collocation of model results and observations need to be done by minimizing the difference in resolution (horizontal, vertical and temporal). This is commonly solved by averaging a domain with a finer spatial and temporal resolution to match a coarser space-time domain. Secondly, collocation needs to make sure that the difference between a point measurement (ground-based) and a grid box size is small in relationship to the spatio-temporal variability of the compared quantity.

Evaluating an overall model performance can be done by, for example, time-series analysis at certain grid points to study bias and correlation in time, or by comparing larger areas to analyse the spatial bias and correlation. In this thesis we try to analyse the impact from lateral boundary conditions by studying both time-series from comparisons with rural ground-based measurement stations within the model, as well as large spatial areas form comparisons with satellite observations. The statistical metrics for bias and correlation used in this thesis are defined according to Eq. 5.1 and Eq. 5.2.

$$\text{bias} = \frac{1}{N}\sum_{i=0}^{N}\frac{x_{m,i} - x_{o,i}}{x_{o,i}} \cdot 100\% \tag{5.1}$$

$$\text{correlation} = \sum_{i=1}^{N}\frac{(x_{m,i} - x_{m,avg})(x_{o,i} - x_{o,avg})}{\sigma_m \cdot \sigma_o} \tag{5.2}$$

where N is the total number of compared data pairs, $x_{m,i}$ and $x_{o,i}$ correspond to the modelled and observed data point, respectively, $x_{m,avg}$ and $x_{o,avg}$ are the arithmetic mean values of the model data and observations, and σ_m and σ_o are the standard deviations for the model data and the observations. The bias and correlation can then easily be applied to retrieve both temporal and spatial statistics.

5.2 Data Analysis and Assimilation

Data analysis and data assimilation are statistical methods used for constraining a model with observations or finding an atmospheric state in model space from observations, by merging information from the two data sets based on characterised errors statistics from each source of information. The result of the merging is referred to as an analysis. Data assimilation refers to the process of merging observations with a dynamic model. The easiest way to do this is to sequentially perform model integration forward in time, followed by a data analysis, followed by the next time-integration, etc. This thesis uses data analysis to constrain the lateral boundary conditions of CTM by use of satellite-based lidar observations.

There are different statistical approaches that can be used to solve the problem of combining information from the a priori and the observations. One commonly used statistical method is called variational analysis, which seeks an analysis by minimising a cost-function to obtain a maximum-likelihood solution to the data analysis problem. This method is used in this thesis and is presented in Sect. 5.2.1.

Model estimates and observations are rarely derived/measured in the same physical space. In a CTM with a discretized 3D geographical domain, each grid box contains information about aerosols and their mass mixing ratios for each chemical component and size bin. The model can thus provide a vector $\mathbf{x} \in \mathbb{R}^n$ of the atmospheric aerosol state, where n is the number of model variables, or dimension of model space. Similarly, we have a set of m observations $\mathbf{y} \in \mathbb{R}^m$, with for example m_1 different observational parameters and m_2 different wavelengths ($m_1 \cdot m_2 = m$). In order to make observations and model results comparable to each other, we need a forward model, or an observation operator,

$\hat{H} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{x} \mapsto \hat{H}(\mathbf{x})$ that maps the model space variables into observation space. Section 5.2.2 will discuss the use of different observation operators in the context of finding an analysis to constrain model space aerosols.

5.2.1 Variational analysis - 3DVAR

The problem of finding the atmospheric state $\mathbf{x}$ based on observations $\mathbf{y}$ is an ill-posed inverse problem. In variational data analysis one takes a probabilistic approach to solving this problem. More specifically, one maximises the conditional probability that the atmosphere is in state $\mathbf{x}$ given observations $\mathbf{y}$, denoted by $P(\mathbf{x}|\mathbf{y})$. According to Bayes theorem

$$P(\mathbf{x}|\mathbf{y}) \propto P(\mathbf{x})P(\mathbf{y}|\mathbf{x}), \tag{5.3}$$

where $P(\mathbf{x})$ denotes the prior probability of the atmosphere being in state $\mathbf{x}$, and $P(\mathbf{y}|\mathbf{x})$ is the conditional probability for obtaining observations $\mathbf{y}$ given that the atmosphere is in state $\mathbf{x}$. Assuming Gaussian statistics, we have

$$P(\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \cdot \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\right) \tag{5.4}$$

$$P(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{1}{2}(\hat{H}(\mathbf{x})-\mathbf{y})^T \cdot \mathbf{R}^{-1} \cdot (\hat{H}(\mathbf{x})-\mathbf{y})\right) \tag{5.5}$$

where $\mathbf{x}_b$ denotes a background estimate of the atmospheric state given by a model forecast, and $\mathbf{B}$ and $\mathbf{R}$ denote the covariance matrices of the background estimate and the observations, respectively.

Observation errors come from measurement errors as well as forward model errors, as mentioned in Sect.-3.1.2. In practice when deriving the error covariance matrix $\mathbf{R}$, estimated satellite uncertainties in terms of standard deviation are used to form an error variance matrix. Model errors to derive the $\mathbf{B}$ error covariance matrix have to be estimated by some appropriate method. A common method applied in data analysis contexts, is the NMC method (Parrish and Derber 1992) which is based on correlating forecasts of different forecast lengths to derive an error covariance matrix. This method is also applied in the 3DVAR module, used in this thesis.

The conditional probability distribution $P(\mathbf{x}|\mathbf{y})$ can be summarised as in Eq.-5.6 and 5.7,

$$P(\mathbf{x}|\mathbf{y}) \propto \exp\left(-J(\mathbf{x})\right), \tag{5.6}$$

$$J = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \cdot \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}(\hat{H}(\mathbf{x})-\mathbf{y})^T \cdot \mathbf{R}^{-1} \cdot (\hat{H}(\mathbf{x})-\mathbf{y}) \tag{5.7}$$

where $J(\mathbf{x})$ is referred to as the cost function, since it can be interpreted as how "costly" it is for $\mathbf{x}$ to simultaneously deviate from the background estimate and the observations within the given error bounds. To find the most probable atmospheric aerosol state, one has to seek that state for which the probability distribution attains its maximum. This is done by seeking the minimum of $J(\mathbf{x})$ and computing the gradient $\nabla_x J$ in a descent algorithm to iteratively search for the minimum of $J(\mathbf{x})$.

The data analysis problem is usually ill-posed because the dimension of the model space is typically much larger than that of the observation space ($m << n$). This prompts the question which of the model variables can actually be constrained by the observations. The information content of the observations

can be studied by methods borrowed from retrieval theory (e.g. Rodgers, 2000). This helps us to understand to what extent the analysis depends on the a priori (background) estimate, which model variables are directly updated by the observations, and which model variables are merely indirectly updated by being statistically correlated to the directly updated variables.

5.2.2 Observation operator and aerosols

The observation operator maps the model space a priori to a model equivalent observation. In practice, when implementing the observation operator $\hat{H}(\mathbf{x})$ in a variational method, it is linearised with a first-order Taylor expansion as in Eq. 5.8,

$$\hat{H}(\mathbf{x}) = \hat{H}(\mathbf{x}_b) + \mathbf{H}(\mathbf{x}_b - \mathbf{x}) \tag{5.8}$$

where $\mathbf{H}$ is the Jacobian of $\hat{H}$ at $\mathbf{x} = \mathbf{x}_b$. This approximation can only be done if $\hat{H}$ is linear or mildly non-linear, which needs to be considered when formulating the forward model. The continuation of this section will focus on how to define an observation operator linking an aerosol concentration field given by a CTM to a model equivalent observation of aerosol optical properties for use in variational data analysis.

Aerosol optics modelling

Aerosol optical properties depend on three major properties,

- the aerosol size distribution
- the refractive index of the materials of which the aerosols are composed of
- the morphology of the particles.

Depending on the level of detailed description of an aerosol field given by a CTM, an aerosol optics model need to be adapted and invoke appropriate assumptions regarding the physical and chemical properties of aerosols. The chemical composition of particles is given by a CTM. If aerosol micro-physical processes are included in the model, then the size-distribution, the size-dependent chemical composition, and the mixing state of aerosols are also provided by the model output. From the chemical composition, an optics model can compute the refractive index of the particles. Refractive indices of different materials are commonly derived from laboratory measurements. The effective refractive indices of homogeneously mixed materials can be computed from those of pure materials by use of appropriate mixing rules (known as effective medium theories). The optics model employed in the CTM contains look-up tables of optical properties for various refractive indices and wavelengths. Morphology refers to the shape of a particle, to whether a particle consists of one or several chemical components, and how different components are mixed. Information about the mixing state is, as already stated, commonly given by a CTM. The aerosols are either externally or internally mixed, i.e., each particle consist of one chemical species or one particle can consist of many. However, if an aerosol is inhomogeneously internally mixed, assumptions regarding their internal mixing has to be made

by the optics model. Information about the shape of the particles is not provided by CTMs and needs to be regarded in an optics model. The simplest approach to simulate morphologies by CTMs and optics models, are to assume that all particles are externally mixed spheres. This assumption results in a linear relationship between the aerosol concentrations of the CTM and the optical properties (which is a requirement for variational data analysis/assimilation). However, these assumptions are known to introduce errors, especially for solid particles such as mineral dust and black carbon, which exists as both externally mixed particles with more complex structures as well as inhomogeneously internally mixed particles. This is especially problematic for black carbon since, it is known to absorb infrared radiation, both as freshly emitted (externally mixed) and aged (internally mixed). Aged BC particles can have a coating of a soluble compound, which can increase the absorption cross section (Kahnert et al. 2014, and references therein). Wherefore, simulating poorly derived optical properties for black carbon thence used in climate models results in erroneous climate forcings. This has prompted modellers to include more realistic descriptions of solid particles in aerosol optical models (Jacobson 2000).

In this thesis two different aerosol optical models are used, one simple model with only externally mixed homogeneous spheres, and another including internally mixed particles and a more realistic description of externally mixed black carbon as fractal aggregates and internally mixed as a core-grey-shell configuration (also seen in Fig. 2.2b).

6
Summary and Outlook

6.1 Summary of Paper A

Paper A focuses on a method of evaluating lateral boundary conditions (LBCs) in regional chemical transport models. The methodology is divided into two distinct parts. The first part is inspired by the work done by Henderson et al. (2014), in which LBCs are evaluated for a North American model domain by comparing global chemical transport model (CTM) data at the lateral boundaries with satellite retrievals from MOPITT and OMI for CO and ozone. We perform a similar evaluation, but over a European model domain and use global CTM data from the global version of the EMEP MSC-W model, as LBCs. In the second part of the evaluation, we apply these LBCs to the regional CTM MATCH (Multi-scale Atmospheric transport and CHemistry) model, developed by the Swedish Meteorological and Hydrological Institute (SMHI), and compare with independent data sets from the satellite sensor AIRS and ground based measurements from the GAW network. The regional model is also forced with dynamical and climatological boundary conditions derived from the global EMEP model. With this methodology we address questions that cannot be addressed fully by only evaluating the LBCs at the boundaries. More specifically, we address the following questions: i) Does the comparison between LBCs and satellite retrievals allow us to quantify the accuracy of the boundary values close to the surface?, ii) How much does the LBCs impact in-domain concentrations in the free troposphere? and iii) Which improvements can one attain by using dynamic as opposed to climatological LBCs?

The results show that, i) a direct comparison between LBCs and satellite retrievals can quantify a bias change close to the surface, ii) the LBCs have a large impact on the free troposphere concentrations, affecting the bias and the spatial correlation, and iii) the largest improvements from using dynamical LBCs, compared to climatological LBCs, are most apparent when using the sensitive metrics of AOT40 as well as accurately capturing the variability close the the western inflow boundary.

By contrast to satellite observations for trace gases, there are no retrieval products available for aerosol concentration and composition. Thus, if we want to use satellite observations to evaluate or improve lateral boundary conditions for aerosols, then we first need to solve an inverse problem. This requires some

preparations. We need to (i) develop an aerosol-optics forward model that maps aerosol concentrations to optical observables; and (ii) we need to understand how much information on aerosol concentrations and compositions we are able to obtain from satellite observations. Papers B and C provide such preparatory work, while paper D deals with lateral boundary conditions for aerosols.

6.2 Summary of Paper B

Paper B describes the implementation of a new aerosol optics model in the regional CTM MATCH. The new optics model assumes more realistic particles and mixing states, where both external and internal mixtures are simulated. Special account is also taken to black carbon (BC), where externally mixed BC is treated as fractal aggregates and inhomogeneous internal mixtures of BC are modelled with a core-grey-shell model. These two BC models have previously been shown to reproduce absorption and scattering by morphologically complex BC particles (Kahnert 2010b, Kahnert et al. 2013). The main question addressed in this paper, is whether or not such detailed description of particles, offered by this new optics model, has a significant impact on radiometric quantities and radiative forcings. The difference between this new optics model and a simple model, treating all particles as externally mixed spheres, is gauged with the well-understood effect of including or omitting aerosol dynamics in a CTM. Thus we calculate aerosol optical properties with four different model configurations, 1) the MATCH mass-transport model (i.e., with aerosol dynamics switched off) together and the simple optics model; 2) the MATCH mass-transport model and the new optics model; 3) MATCH including the aerosol dynamics module SALSA (MATCH-SALSA) and the simple optics model; and 4) the MATCH-SALSA and the new optics model.

Aerosol optical properties and radiative forcings (derived with the LibRadtran package (Kylling et al. 1998)) are compared between the four model set-ups. The comparisons show that the impact of assuming more realistic morphologies of model particles in the optics model is of the same order of magnitude as including aerosol dynamics in a chemical transport model. This is important, since many climate and remote sensing applications today employ simplistic models to simulate aerosol optical properties.

6.3 Summary of Paper C

Paper C investigates the information content of multi-wavelength lidar observation of aerosols. The main questions are (i) How many model variables can be controlled by any given set of observations? (ii) How does the information content of the observations change as we increase the number of wavelengths, or as we complement backscattering with extinction measurements? (iii) How does the information content depend on the observation errors? and (iv) Which aerosol variables in the model can actually be constrained by the limited information contained in the measurements?

Information content from different sets of multi-wavelength lidar observations with varying observation errors, is derived by utilising singular-value decomposition (SVD) of the normalised Jacobian of an observation operator, to

find the signal degrees of freedom N_s and Shannon entropy H. We analyse which aerosol model variables are most strongly related to the signal degrees of freedom in two ways. (i) Mathematically, we can obtain this information by inspecting the singular vectors. (ii) More practically, we perform a controlled observation system simulation experiment and test which model variables are most faithfully retrieved in the analysis process.

Both N_s and H are used to represent the information content, and both increase when the number of observations increases. However, both measures are strongly dependent upon the observation error σ_o. The larger the observation error, the lower the growth rate of N_s and H with the number of measurements. We find that the total aerosol mass mixing ratio (PM_{10}), as well as the sulfate mass mixing ratio, can be well constrained by assimilating backscattering observations at three wavelengths and extinction observations at two wavelengths.

A similar study was done by Kahnert (2018), where a SVD of the observation operator was applied to retrieve information about the signal- and noise related components in order to apply it to the same 3DVAR system. However, the significant difference is that the constraints are applied in the abstract phase space instead of the physical space. This leads to a fully Bayesian formulation and to the conclusion that a constrained analysis (compared to an unconstrained) would only be slightly improved if the background error statistics would be inaptly derived. Otherwise, the new implementation of the constraints leads to a significant reduction of CPU time, by order of one magnitude.

6.4 Summary of Paper D

The results from papers B and C have paved the way for the final paper in this thesis. Paper D aims to derive lateral boundary conditions for aerosols from satellite-based lidar observations, and evaluate them by implementing them in a CTM. Compared to Paper A, new LBCs are derived for aerosols by constraining a CTM background estimate with aerosol extinction observations (EXT) at 532 nm form CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) lidar measurement. The observation-derived LBCs are evaluated by running the model for a one month period and comparing in-domain concentrations to in situ ground observations of PM_{10} (Paper C showed that PM_{10} was the best aerosol model variable to constrain).

By analysing 67 profiles of CALIPSO EXT data for July 2008 with a 3DVAR analysis and a hemispheric version of MATCH-SALSA, new LBCs are created. The 3DVAR analysis includes the newly developed aerosol optical model in Paper B and the information content constraints in Paper C and Kahnert (2018). An indirect evaluation is done at the surface level of the model-domain by comparing two regional MATCH-SALSA runs with 51 EMEP measurement stations, measuring PM_{10}. One model run is driven by hemispheric MATCH-SALSA as LBCs, and one by the new LBCs. The comparison shows that the model bias gets reduced, from 26 % when using the hemispheric MATCH-SALSA as LBC to 6 % when using the analysed LBCs. This is a large improvement since only 67 profiles and a total number of 981 data points were analysed. The information analysis revealed 889 signal degrees of freedom, which entails that of 981 observation points, 889 model variables could be independently constrained. Keeping in mind that the lateral boundary consists of 8140 grid points with a total of 76

variables in each grid box.

6.5 Outlook

The main goal of this thesis was to confront regional CTM simulations with observations to better represent long-range transport. This was done by evaluating lateral boundary conditions for both trace gases and aerosols. A methodology was developed for evaluating LBCs both directly at the boundary as well as indirectly by evaluating CTM results derived from using different LBCs. This methodology was first applied to LBCs of trace gases; the aim was, thereafter, to apply it to LBCs of aerosols and to create LBCs derived from constraining a large scale model with aerosol extinction retrievals. However, this could only be possible if we had an aerosol optics model and knew how to use the information provided by the observed retrievals to constrain model results. After testing a newly developed aerosol optical model and applying a method for analysing the information content of aerosol extinction and backscattering measurements, we created new LBCs. These LBCs were further implemented and indirectly evaluated in a regional CTM. Unfortunately, the full methodology of a direct and indirect evaluation of the LBCs could not be fully adapted, due to time constraints. Thus, future work could involve the following:

- Analysing or assimilating a larger set of observations. The present fourth study only analyse level 3 monthly averaged extinction data for one month. If observations over a longer period of time were assimilated, it would be possible to conduct a more robust statistical analysis of the effect from the newly created LBCs.

- Assimilating level 2 data, which has a higher temporal resolution than level 3, could better capture dynamical features of the atmospheric aerosols.

- Using attenuated backscattering measurements, instead of extinction. The retrieved extinction data from CALIPSO is derived from attenuated backscattering assuming a lidar-ratio. Using attenuated backscattering would circumvent errors associated with that assumption. An observation operator of attenuated backscattering is under development.

- Increasing the information content by using a larger set of observation to be able to constrain more model variables. This could be achieved by combining different types of independent observations.

6.6 Contribution to papers

Paper A I set-up and ran the different model runs as well as post-processed the model and observation data for comparisons, supervised by Michael Kahnert, Lennart Robertsson and Abhay Devasthale at SMHI. Global EMEP MSC-W model data was provided by David Simpson. Further, I conducted the result analysis and writing assisted by Michael Kahnert and Abhay Devasthale.

Paper B I set-up and ran the MATCH model runs assisted by Michael Kahnert. Michael Kahnert ran the two optics models as well as the radiative transfer model to retrieve the radiative forcings. The results was post-processed by me and and visualised by both me an Michael. The writings was done by both me a Michael Kahnert.

Paper C M Kahnert has derived and implemented the information constraint in the 3DVAR module and analysed most results. I ran model runs for the different observation parameters and assisted in the result analysis.

Paper D I set-up and ran the hemispheric MATHC-SALSA with assistance from Manu Thomas at SMHI. Meteorological input from ECMWF was provided by Lennart Robertsson. I further retrieved the CALIPSO extinction data and processed it to be used in the 3DVAR data analysis and ran the 3DVAR program to create new lateral boundary conditions. Michael Kahnert ran the regional version of MATCH-SALSA. The model results was post-processed by me. The result analysis and writing of the paper was done together with Michael Kahnert.

Bibliography

C. Andersson, R. Bergström, C. Bennet, L. Robertson, M. Thomas, H. Korhonen, K. H. J. Lehtinen, and H. Kokkola. MATCH-SALSA — multi-scale atmospheric transport and CHemistry model coupled to the SALSA aerosol microphysics model — Part 1: Model description and evaluation. *Geosci. Model Dev.*, 8:171–189, 2015.

S. Pal Arya. *Introduction to Micrometeorology*. Academic Press, 2001.

F. Auclair, P. Marsaleix, and C. Estournel. Sigma coordinate pressure gradient errors: Evaluation and reduction by an inverse method. *Journal of Atmospheric and Oceanic Technology*, 17(10):1348–1367, 2000. doi: 10.1175/1520-0426(2000)017<1348:SCPGEE>2.0.CO;2. URL `https://doi.org/10.1175/1520-0426(2000)017<1348:SCPGEE>2.0.CO;2`.

J. Bak, X. Liu, J. H. Kim, K. Chance, and D. P. Haffner. Validation of omi total ozone retrievals from the sao ozone profile algorithm and three operational algorithms with brewer measurements. *Atmospheric Chemistry and Physics*, 15(2):667–683, 2015. doi: 10.5194/acp-15-667-2015. URL `https://www.atmos-chem-phys.net/15/667/2015/`.

T. Berg, C. Dye, J. E. Hanssen, J. Munthe, J.-P. Putaud, A. Reisell, J. Schaug, N. Schmidbauer, A. Semb, K. Torseth, H. Thelle Uggerud, K.-E. Yttri, and W. Aas. *EMEP manual for sampling and analysis*, 2014.

T. C. Bond and R. W. Bergstrom. Light absorption by carbonaceous particles: An investigative review. *Aerosol Sci. Technol.*, 40:27–67, 2006.

A. Bott. A positive definite advection scheme obtained by non-linear renormalization of advective fluxes. *Mon. Wea. Rev.*, 117:1006–1015, 1989.

M. T. Chahine, T. S. Pagano, H. H. Aumann, R. Atlas, C. Barnet, J. Blaisdell, L. Chen, M. Divakarla, E. J. Fetzer, M. Goldberg, C. Gautier, S. Granger, S. Hannon, F. W. Irion, R. Kakar, E. Kalnay, B. H. Lambrigtsen, S.Y. Lee, J. Le Marshall, McMillan W. W., L. McMillin, E. T. Olsen, H. Revercomb, P. Rosenkranz, W. L. Smith, D. Staelin, L. L. Strow, J. Susskind, D. Tobin, W. Wolf, and L. Zhou. Airs: Improving weather forecasting and providing new data on greenhouse gases. *B. Am. Meteorol. Soc.*, 87:911–926, 2006.

M. N. Deeter, H. M. Worden, D. P. Edwards, J. C. Gille, and A. E. Andrews. Evaluation of mopitt retrievals of lower-tropospheric carbon monoxide over the united states. *J. Geophys. Res.*, 117:D13306, 2012.

M. N. Deeter, S. Martínez-Alonso, D. P. Edwards, L. K. Emmons, J. C. Gille, H. M. Worden, C. Sweeney, J. V. Pittman, B. C. Daube, and S. C. Wofsy. The mopitt version 6 product: algorithm enhancements and validation. *Atmospheric Measurement Techniques*, 7(11):3623–3632, 2014. doi: 10.5194/amt-7-3623-2014. URL `https://www.atmos-meas-tech.net/7/3623/2014/`.

J. R. Drummond, J. Zou, F. Nichitiu, J. Kar, R. Deschambaut, and J. Hackett. A review of 9-year performance and operation of the mopitt instrument. *Advances in Space Research*, 45(6):760–774, 2010. doi: https://doi.org/10.1016/j.asr.2009.11.019.

B.J. Finlayson-Pitts and J.N. Pitts Jr. *Chemistry of the Upper and Lower Atmosphere*. Acamdemic Press, 2000.

G. L. Francis, M. N. Deeter, S. Martinez-Alonso, J. C. Gille, D. P. Edwards, D. Mao, H. M. Worden, and D. Ziskin. *MOPITT, Algorithm Theoretical Basis Document, Retrieval of Carbon Monoxide Profiles and Column Amounts from MOPITT Observed Radiances, (Level 1 to Level 2)*, 2017.

I.E. Galbally and M.G. Schultz. Global atmosphere watch - guidelines for continuous measurements of ozone in the troposphere. Technical Report 110, World Meteorological Organisation, 2013.

R.T. Gooby and Y. L. Yung. *Atmospheric Radiation: Theoretical Basis*. Oxford University Press Inc., 1995.

B. H. Henderson, F. Akhtar, H. O. T. Pye, S. L. Napelenok, and W. T. Hutzell. A database and tool for boundary conditions for regional air quality modeling: description and evaluation. *Geosci. Model Dev.*, 7:339–360, 2014.

M. Z. Jacobson. A physically-based treatment of elemental carbon optics: Implications for global direct forcing of aerosols. *Geophys. Res. Lett*, 27:217–220, 2000.

P. Jiménez, R. Parra, and J. M. Baldasano. Influence of initial and boundary conditions for ozone modeling in very complex terrains: A case study in the northeastern Iberian Peninsula. *Environ. Modell. Softw.*, 22:1294–1306, 2007.

M. Kacenelenbogen, M. A. Vaughan, J. Redemann, R. M. Hoff, R. R. Rogers, R. A. Ferrare, P. B. Russell, C. A. Hostetler, J. W. Hair, and B. N. Holben. An accuracy assessment of the caliop/calipso version 2/version 3 daytime aerosol extinction product based on a detailed multi-sensor, multi-platform case study. *Atmospheric Chemistry and Physics*, 11(8):3981–4000, 2011. doi: 10.5194/acp-11-3981-2011. URL `https://www.atmos-chem-phys.net/11/3981/2011/`.

M. Kahnert. Modelling the optical and radiative properties of freshly emitted light absorbing carbon within an atmospheric chemical transport model. *Atmos. Chem. Phys.*, 10:1403–1416, 2010a.

M. Kahnert. On the discrepancy between modelled and measured mass absorption cross sections of light absorbing carbon aerosols. *Aerosol Sci. Technol.*, 44:453–460, 2010b.

M. Kahnert. Information constraints in variational data assimilaiton. *QJRMS*, 2018. doi: 10.1002/qj.3347.

M Kahnert, T. Nousainen, and H. Lindqvist. Models for integrated and differential scattering optical properties of encapsulated light absorbing carbon aggregates. *Optics Express*, 21:7974–7993, 2013.

M. Kahnert, T. Nousiainen, and H. Lindqvist. Review: Model particles in atmospheric optics. *J. Quant. Spectrosc. Radiat. Transfer*, 146:41–58, 2014.

M Kampa and Castanas. E. Human health effects of air pollution. *Environmental Pollution*, 151(2):362 – 367, 2008. doi: https://doi.org/10.1016/j.envpol.2007.06.012.

J. Kar, M. A. Vaughan, K.-P. Lee, J. L. Tackett, M. A. Avery, A. Garnier, B. J. Getzewich, W. H. Hunt, D. Josset, Z. Liu, P. L. Lucker, B. Magill, A. H. Omar, J. Pelon, R. R. Rogers, T. D. Toth, C. R. Trepte, J.-P. Vernier, D. M. Winker, and S. A. Young. Calipso lidar calibration at 532 nm: version 4 nighttime algorithm. *Atmospheric Measurement Techniques*, 11(3):1459–1479, 2018. doi: 10.5194/amt-11-1459-2018. URL `https://www.atmos-meas-tech.net/11/1459/2018/`.

Z. Klimont, K. Kupiainen, C. Heyes, P. Purohit, J. Cofala, P. Rafaj, J. Borken-Kleefeld, and W. Schöpp. Global anthropogenic emissions of particulate matter including black carbon. *Atmospheric Chemistry and Physics*, 17 (14):8681–8723, 2017. doi: 10.5194/acp-17-8681-2017. URL `https://www.atmos-chem-phys.net/17/8681/2017/`.

H. Kokkola, H. Korhonen, K. E. J. Lehtinen, R. Makkonen, A. Asmi, S. Järvenoja, T. Anttila, A.-I. Partanen, M. Kulmala, H. Järvinen, A. Laaksonen, and V.-M. Kerminen. SALSA — a sectional aerosol module for large scale applications. *Atmos. Chem. Phys.*, 8:2469–2483, 2008.

M. Kroon, J. F. de Haan, J. P. Veefkind, L. Froidevaux, R. Wang, R. Kivi, and J. J. Hakkarainen. Validation of operational ozone profiles from the ozone monitoring instrument. *J. Geophys. Res.*, 116:D18305, 2011.

A. Kylling, A. F. Bais, M. Blumthaler, J. Schreder, C. S. Zerefos, and E. Kosmidis. The effect of aerosols on solar uv irradiances during the photochemical activity and solar ultraviolet radiation campaign. *J. Geophys. Res.*, 103: 26,051–26,060, 1998.

K.N. Liou. *An Introduction to Atmospheric Radiation*. Academic Press, 2002.

N. Papagiannopoulos, L. Mona, L. Alados-Arboledas, V. Amiridis, H. Baars, I. Binietoglou, D. Bortoli, G. D'Amico, A. Giunta, J. L. Guerrero-Rascado, A. Schwarz, S. Pereira, N. Spinelli, U. Wandinger, X. Wang, and G. Pappalardo. Calipso climatological products: evaluation and suggestions from earlinet. *Atmospheric Chemistry and Physics*, 16(4):2341–2357, 2016. doi: 10.5194/acp-16-2341-2016. URL `https://www.atmos-chem-phys.net/16/2341/2016/`.

D. F. Parrish and J. C. Derber. The national meteorological centre's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, 120:1747–1763, 1992.

A. C. Povey and R. G. Grainger. Known and unknown unknowns: uncertainty estimation in satellite remote sensing. *Atmospheric Measurement Techniques*, 8(11):4699–4718, 2015. doi: 10.5194/amt-8-4699-2015. URL `https://www.atmos-meas-tech.net/8/4699/2015/`.

L. Robertson, J. Langner, and M Engardt. An Eulerian limited-area atmospheric transport model. *J. Appl. Meteor.*, 38:190–210, 1999.

C.D. Rodgers. *Inverse Methods for Atmospheric Sounding*. World Scientific, 2000. doi: 10.1142/3171. URL `https://www.worldscientific.com/doi/abs/10.1142/3171`.

V.V. Roshchina and V.S. Roshchina. *Ozone and Plant Cell*. Springer Science+Business Media Dordrecht, 2003.

J.H. Seinfeld and S.N. Pandis. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. Wiley-Interscience, John Wiley and Sons, Inc., 2016.

M. Shiraiwa, K. Ueda, A. Pozzer, G. Lammel, C.J. Kampf, A. Fushimi, S. Enami, A.M. Arangio, J. Fröhlich-Nowoisky, Y. Fujitani, A. Furuyama, P.S.J. Lakey, K. Lelieveld, J.and Lucas, Y. Morino, U. Pöschl, S. Takahama, A. Takami, H. Tong, B. Weber, A. Yoshino, and K. Sato. Aerosol health effects from molecular to global scales. *Environmental Science & Technology*, 51(23):13545–13567, 2017. doi: 10.1021/acs.est.7b04417. URL `\url{https://doi.org/10.1021/acs.est.7b04417}`. PMID: 29111690.

D. Simpson, Y. Andersson-Sköld, M.E. Jenkin, EMEP Cooperative Programme for Monitoring, and Evaluation of the Long Range Transmission of Air Pollutants in Europe. *Updating the Chemical Scheme for the EMEP MSC-W Oxidant Model: Current Status*. EMEP-MSC-W Note 2-93. Norwegian Meteorological Institute, Meteorological Synthesizing Centre-West, 1993. URL `http://books.google.se/books?id=MexPywAACAAJ`.

D. Simpson, A. Benedictow, H. Berge, R. Bergström, L. D. Emberson, H. Fagerli, C. R. Flechard, G. D. Hayman, M. Gauss, J. E. Jonson, M. E. Jenkin, A. Nyíri, C. Richter, V. S. Semeena, S. Tsyro, J.-P. Tuovinen, Á. Valdebenito, and P. Wind. The EMEP MSC-W chemical transport model – technical description. *Atmos. Chem. Phys.*, 12(16):7825–7865, 2012.

J. L. Tackett, D. M. Winker, B. J. Getzewich, M. A. Vaughan, S. A. Young, and J. Kar. Calipso lidar level 3 aerosol profile product: version 3 algorithm design. *Atmospheric Measurement Techniques*, 11(7):4129–4152, 2018. doi: 10.5194/amt-11-4129-2018. URL `https://www.atmos-meas-tech.net/11/4129/2018/`.

Y. Tang, G. R. Carmichael, N. Thongboonchoo, T. Chai, L. W. Horowitz, R. B. Pierce, J. A. Al-Saadi, G. Pfister, J. M. Vukovich, M. A. Avery, G. W. Sachse, T. B. Ryerson, J. S. Holloway, E. L. Atlas, F. M. Flocke, R. J. Weber, L. G.

Huey, J. E. Dibb, D. G. Streets, and W. H. Brune. Influence of lateral and top boundary conditions on regional air quality prediction: A multiscale study coupling regional and global chemical transport models. *J. Geophys. Res.*, 112, 2007.

M. Tiedtke. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, 117(8):1779–1800, 1989. doi: 10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2. URL https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.

J. M. Wallace and P. V. Hobbs. *Atmospheric Sciences - An introductory survey.* Academic Press, Elsevier - International Geophysics Series, 2006.

R. Zellner. *Global Aspects of Atmospheric Chemistry.* Steinkopff and Springer, Darmstadt and New York, 1999.