Detailed analysis of MAX-DOAS measurements in Bremen: Spatial and temporal distribution of aerosols, formaldehyde and nitrogen dioxide

---

Tim Bösch

Dissertation

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
(Dr. rer. nat)

Insitut für Umweltphysik
Fachbereich Physik und Elektrotechnik
Universität Bremen
Oktober 2018

Bearbeitungszeitraum:	15. Mai 2014 bis November 2018
Erstgutachter:	Prof. Dr. John P. Burrows
Zweitgutachter:	Prof. Dr. Thomas Wagner
Betreuer:	Dr. Andreas Richter

ABSTRACT

In this thesis, spatial and temporal tropospheric inhomogeneities in the distribution of nitrogen dioxide (NO_2), formaldehyde ($HCHO$) and aerosols are investigated. The analysis was done on a three years dataset (2015 - 2017) of ground-based multi-axis differential optical absorption spectroscopy (MAX-DOAS) measurements in Bremen.

MAX-DOAS measurements were investigated in three different azimuthal viewing directions in order to analyse lateral changes in the distribution of NO_2 and $HCHO$ in Bremen. A clear seasonality was found and explained by anthropogenic and biogenic emissions for NO_2 and $HCHO$, respectively. While no significant azimuthal variability for $HCHO$ was found, NO_2 differs strongly for the westerly and southerly directions due to lateral inhomogeneities and a more frequent pointing towards the sun which has a strong impact on the results. In order to localize possible dominant emission sources of NO_2 within the area of Bremen, the onion peeling approach was successfully applied by usage of the three fitting windows in the ultra-violet (UV) and visible (vis) spectral range. Strong emitters could be identified having a large impact on average NO_2 results.

A major challenge for the analysis of trace gases in the troposphere is the usually insufficient knowledge of aerosols, which might have a large impact on spectroscopic measurements. The novel MAX-DOAS profiling algorithm BOREAS was developed and its accuracy is validated with the help of synthetic data as well as ancillary measurements of the CINDI-2 field campaign (Cabauw, the Netherlands, 2016). In contrast to other algorithms, BOREAS' aerosol information are retrieved by minimizing the difference of O_4 optical depths of measurements and forward modelling calculations. The resulting aerosol extinction coefficient profiles were used for the retrieval of vertical trace gas concentration profiles. In this thesis, several retrieval modes and various ways of improving the regularization between measurement and a priori constraints as well as the selection of proper a priori profiles by use of a priori pre-scaling were investigated.

The BOREAS algorithm was finally applied to the full MAX-DOAS dataset, and three years of aerosol and trace gas vertical profiles from the measurement location Bremen are presented and discussed with the help of in-situ as well as AERONET measurements. Seasonal, weekday and diurnal cycles for aerosols and NO_2 were found which could be attributed to near surface emissions mainly from traffic and power plants. The seasonal cycle of $HCHO$ is found to be dominated by biogenic emissions in summer, in addition to a smaller fraction of anthropogenic emissions in winter. While NO_2 and aerosols are mainly focussed in layers close to the surface, larger $HCHO$ concentrations could also be observed in the complete planetary boundary layer showing the need for the analysis of vertical concentration profiles of trace gases in the troposphere.

SCIENTIFIC PUBLICATIONS

Articles in peer-reviewed journals

- Frieß et al., 2018 (in preparation)
- Gratsea et al., 2019 (in preparation)
- Kreher et al., 2018 (in preparation)
- Merlaud et al., 2019 (in preparation)
- Peters et al., 2019 (in preparation)
- Tirpitz et al., 2018 (in preparation)
- **Bösch, T.**, Rozanov, V., Richter, A., Peters, E. Rozanov, A., Wittrock, F., Merlaud, A., Lampel, J., Schmitt, S., de Haij, M., Berkhout, S., Henzing, B., Aptiluey, A. den HOED, M., Vonk, J., Tiefengraber, M., Müller, M., and Burrows, J. P.: 'BOREAS - a New MAX-DOAS Profi le Retrieval Algorithm for Aerosols and Trace Gases'. In: Atmos. Meas. Tech. Discuss. pp. 1–39, doi: 10/gdqbdd, 2018.
- Meier, A. C., Schönhardt, A., **Bösch, T.**, Richter, A., Seyler, A., Ruhtz, T., Constantin, D.-E., Shaiganfar, R., Wagner, T., Merlaud, A., Van Roozendael, M., Belegante, L., Nicolae, D., Georgescu, L., and Burrows, J. P.: 'High-Resolution Airborne Imaging DOAS Measurements of NO2 above Bucharest during AROMAT'. In: Atmos. Meas. Tech., 10.5, pp. 1831–1857, doi : 10.5194/amt-10-1831-2017, 2017.
- Peters, E., Pinardi, G., Seyler, A., Richter, A., Wittrock, F., **Bösch, T.**, Van Roozendael, M., Hendrick, F., Drosoglou, T., Bais, A. F., Kanaya, Y., Zhao, X., Strong, K., Lampel, J., Volkamer, R., Koenig, T., Ortega, I., Puentedura, O., Navarro-Comas, M., Gómez, L., Yela González, M., Piters, A., Remmers, J., Wang, Y., Wagner, T., Wang, S., Saiz-Lopez, A., García-Nieto, D., Cuevas, C. A., Benavent, N., Querel, R., Johnston, P., Postylyakov, O., Borovski, A., Elokhov, A., Bruchkouski, I., Liu, H., Liu, C., Hong, Q., Rivera, C., Grutter, M., Stremme, W., Khokhar, M. F., Khayyam, J., and Burrows, J. P.: 'Investigating differences in DOAS retrieval codes using MAD-CAT campaign data'. In: Atmos. Meas. Tech., 10, pp. 955-978, doi: 10.5194/amt-10-955-2017, 2017.
- Wagner, T., Beirle, S., Benavent, N., **Bösch, T.**, Chan, K. L., Donner, S., Dörner, S., Fayt, C., Frieß, U., García-Nieto, D., Gielen, C., González-Bartolome, D., Gomez, L., Hendrick, F., Henzing, B., Jin, J. L., Lampel, J., Ma, J., Mies, K., Navarro, M., Peters, E., Pinardi, G., Puentedura, O., Puķīte, J., Remmers, J., Richter, A., Saiz-Lopez, A., Shaiganfar, R., Sihler, H., Van Roozendael, M., Wang, Y., Yela, M.: 'Is a Scaling Factor Required to Obtain Closure between Measured and Modelled Atmospheric O_4 Absorptions?'. In: Atmos. Meas. Tech. Discuss. pp. 1–85, doi: 10/gfb74z, 2018.

Selected contribution to conferences

- **Bösch, T.**, Peters, E., Richter, A., Wittrock, F., Rozanov, V., Burrows, J. P.: 'Introduction into IUP Bremen's new MAX-DOAS profile retrieval algorithm BOREAS', 8th International DOAS Workshop, Yokohama, Japan, September 2017.
- **Bösch, T.**, Meier, A., Schönhardt, A., Peters, E., Richter, A., Ruhtz, T., Burrows, J.P.: 'Airborne measurements of different trace gases during the AROMAT-2 campaign with an Avantes spectrometer', EGU General Assembly, Vienna, Austria, April 2016.
- **Bösch, T.**, Peters, E., Hilboll, A., Richter, A., Burrows, J. P.: 'About the information content of the DOAS polynomial', 7th International DOAS Workshop, Brussels, Belgium, July 2015.

Contents

CHAPTER

I

INTRODUCTION

Air pollution results from the release of substances into the air which might be harmful for living organisms or Earth's environment. While the exposure to pollutants might have severe effects to human beings localized in the contaminated area, emissions also deteriorate living conditions for all life on a global scale.

The most famous example of the release of air pollutants having an adverse effect on Earth's atmosphere is the extensive use of chlorofluorocarbons (CFCs) in refrigerants, propellants and solvents in the last century. It was proposed by MOLINA and ROWLAND [1974] that the emissions of CFCs would result in a strong decline of the stratospheric ozone layer which prevents high energetic ultra-violet radiation from entering Earth's lower atmosphere and causing harm to humans. Since the early 1980s, an annual appearance of the ozone hole has been detected (e.g. SOLOMON, 1988). Earlier, it was shown by CRUTZEN [1970] that nitrous oxide (N_2O) released in the troposphere is also able to destroy ozone in a catalytic cycle, together with nitrogen dioxide (NO_2). Their research was one of the reasons for signing a ban of the usage of CFCs within the *Montreal protocol* in September 1987.

This example shows the urgent need for an adequate knowledge of air pollution and its effect on Earth's climate and human health. The world health organization (WHO) recently reported that 7 million premature deaths might be linked to air pollution[1] highlighting again the importance of research in this field.

On a local scale, air pollution due to emissions of facilities and vehicles using combustion processes is frequently discussed by media and politicians at present, owing to the strong impact of pollutants on the respiratory system of humans, the acidification of soil and its impact on plant growth. Not only because of the *Diesel scandal*[2] but also due to the transgression of thresholds for NO_2 in several cities of Germany, the awareness of air pollution being a problem for day-to-day life has increased. The Federal Environment Agency in Germany (Umweltbundesamt) has set the annual NO_2 threshold to the limit of $40\,\mu g/m^3$ which has been exceeded (sometimes by more than $50\,\%$) for more than the

[1] WHO article - 7 million premature deaths annually linked to air pollution. `http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/`, date: 14.09.2018

[2] For cars of the Volkswagen concern (among others nowadays), emission factors were wrongly stated and the common emission measurements were falsified by activating a special software which reduces emissions when doing laboratory tests. See e.g. `https://en.wikipedia.org/wiki/Volkswagen_emissions_scandal`, date: 14.09.2018.

half of all traffic-close air quality measurement sites in Germany, since the year 2010[3]. However, further measurements and research is needed in order to evaluate the temporal and spatial variability of high level concentrations to prevent harm to living beings and the environment.

Different techniques have been used for the measurement of trace gases and aerosols which can be separated into in-situ and remote sensing techniques. Differential Optical Absorption Spectroscopy (DOAS), as an example of remote sensing, is a frequently used and well-established measurement method for evaluating trace gas and aerosol informations. Early instruments applying what we now call the DOAS method, were first operated by BREWER et al. [1973] and NOXON [1975] who made zenith-sky measurements of stratospheric NO_2. PERNER et al. [1976] and PLATT and PERNER [1980] used artificial light sources in long-path DOAS measurements for the analysis of near-surface trace gas concentrations.

The zenith-sky DOAS technique was applied and further improved within several studies in the late 1980th and the 1990th (e.g. POMMEREAU and GOUTAIL, 1988; EISINGER et al., 1997; WAGNER et al., 1998; RICHTER et al., 1999; ROSCOE et al., 1999). In addition to these ground-based instruments, satellite measurements were getting more and more important at the end of the last century and with the early 2000th (e.g. BURROWS et al., 1995; BOVENSMANN et al., 1999; BURROWS et al., 1999; CALLIES et al., 2000). However, even though the spectral and vertical resolution of space-born instrument has been improved within the last two decades (e.g. VEEFKIND et al., 2012), near-surface concentration values or tropospheric vertical profiles are still not accurately retrievable showing the need for ground-based measurements and analyses.

MAX-DOAS instruments have been used for more than 15 years in order to analyse the chemical composition of Earth's troposphere (HÖNNINGER and PLATT, 2002; BOBROWSKI et al., 2003; LESER et al., 2003; WITTROCK et al., 2004; WAGNER et al., 2004; van ROOZENDAEL et al., 2004). Their advantages compared to satellite instruments are the low costs, the high sensitivity for the lowermost layers of the troposphere as well as the high temporal and spatial resolution. The retrieval of vertical profiles of aerosol extinctions and trace gas concentrations from MAX-DOAS observations is an under-constrained inverse problem which has to be solved with the help of radiative transfer models (RTM) and profiling algorithms. Several authors have successfully calculated vertical profiles in the past but the uncertainties on the retrieval results are still high (WITTROCK et al., 2004; HENDRICK et al., 2004; FRIESS et al., 2006; WITTROCK, 2006; CLÉMER et al., 2010; WANG et al., 2017).

In this thesis, Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) measurements are used for the quantification of trace gas and aerosol amounts in the city

[3] Umweltbundesamt - https://www.umweltbundesamt.de/themen/luft/luftschadstoffe/stickstoffoxide, date: 14.09.2018.

of Bremen. For this purpose, a novel algorithm for the calculation of aerosol and trace gas vertical profiles was developed and is presented and validated in order to decrease existing uncertainties in the determination of trace gas and aerosol concentrations in the lower troposphere. Together with the analysis of integrated concentration values with various tools and methods, the profiling results are used for a discussion of the spatial and temporal distribution of NO_2 in the area of Bremen. In addition, DOAS retrievals in the UV spectral range enable the discussion of formaldehyde ($HCHO$) which is another hazardous trace gas of interest whose uncertainties are even higher than those for NO_2 (WAGNER et al., 2011; PINARDI et al., 2013; VLEMMIX et al., 2015; GRATSEA et al., 2016). Therefore, HCHO integrated values as well as vertical profiles are further discussed within this work in order to increase the knowledge about its temporal and spatial distribution.

Outline of this thesis

The second chapter of this thesis focusses on relevant background information on atmospheric physics and chemistry (Ch. II). The solar system as well as the Earth's atmospheric composition are introduced, followed by a description of the basic interaction phenomena between electromagnetic radiation and the atmosphere. Fundamental information on radiative transfer and the DOAS approach are given as well as an introduction into the chemistry of important trace gases and the properties of aerosols.
Within the third chapter (Ch. III), the measurement location Bremen and its climatology are initially introduced. Subsequently, the chapter focusses on the instrumentation, its calibration, as well as a preliminary discussion of slant column densities. The main part of this chapter describes the conversion to vertical column densities and their discussion with respect to spatial and temporal inhomogeneities in the distribution of tropospheric NO_2 and $HCHO$. This analysis shows the need for a more detailed consideration of results retrieved within different fitting windows. Therefore, the onion peeling approach is applied in order to extract possible signals of individual emitters from MAX-DOAS measurements and to localize these emitters in the area of Bremen.
The fourth chapter introduces IUP Bremen's novel MAX-DOAS profiling algorithm BOREAS. The performance and accuracy of BOREAS is shown with the help of a large synthetic aerosol and NO_2 dataset created with the radiative transfer model SCIATRAN. This synthetic analysis is followed by a validation study of profiling results retrieved from data of the CINDI-2 campaign, which was held in Cabauw (the Netherlands), in 2016. The chapter is completed by the discussion of a large dataset of retrieval results from three years of MAX-DOAS measurements in Bremen.
Finally, the most interesting results are summarized together with a presentation of the main conclusions drawn in this study, in addition to a listing of recommended future work (Ch. V).

CHAPTER

II

INTRODUCTION TO ATMOSPHERIC PHYSICS AND CHEMISTRY

Atmospheric remote sensing comprises the analysis of individual constituents of Earth's atmosphere by using natural or artificial interaction phenomena from a more or less remote observation location. These constituents often undergo temporal, vertical and lateral changes and require an accurate analysis because of their impact on Earth's climate and human health.
This chapter introduces the sun as the major radiation source, the interaction of its radiation with the atmosphere and provides detailed information on the atmosphere as well as its components from a physical and chemical point of view. The last subsections of this chapter introduce the theoretical background used within this thesis.

2.1 The solar system, the sun and its impact on Earth's atmosphere

The Sun, as the central star of our solar system, formed 4.6 billion years ago and supplies all bodies within the heliosphere[1] with different forms of energy. This solar energy output can be categorized into four phenomena: solar wind, solar flares, coronal mass ejections and solar radiation (see BOTHMER and DAGLIS [2007] for further details).

The solar wind is a particle stream which has its source in Sun's corona and consists mainly of protons, electrons, and alpha particles. Although it consists of free charge carriers, it can be considered as quasi neutral. Since the *Debye length*[2] is small, in contrast to the distance of the charge carriers to each other, the solar wind can be considered as a plasma. A plasma obeys *Alfvén's Theorem*[3] and, therefore, leads to a movement of magnetic field lines when the plasma moves itself. Because of that, the solar wind leads to a varying deformation of Earth's magnetic field which may even cause field lines to reconnect to a *Magnetotail* behind the Earth, if the pressure is very high (compare Fig. 2.1). This reconnection can be understood as one of the possible ways for charged

[1] The heliosphere describes the volume in space which is affected by the sun.
[2] The length for which the potential of a charge carrier decreases by 1/e.
[3] See DAVIDSON [2002] for further details.

particles to reach the upper atmosphere. In general, Earth's magnetosphere provides sufficient protection for life on Earth due to the fact that energetic particles[5] tend to move along the field lines rather than penetrating it. Nevertheless, energetic particles are able to break through this magnetic shield when their kinetic energy is high enough, which might lead to an ionization and excitation of constituents in the upper atmosphere. This phenomenon is called *Aurora* and is more frequent in higher latitudes because of the perpendicular alignment of magnetic field lines and the surface near the poles.

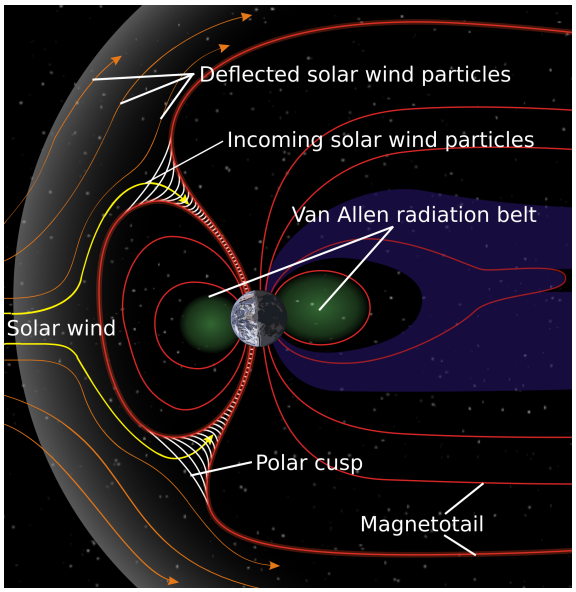


Fig. 2.1: *Schematic representation of the solar wind and its impact on Earth's magnetosphere.*[4]

While the solar wind can be considered as a more or less slow and steady particle stream, which is emitted in all directions, **solar flares and coronal mass ejections** (CME) are sudden eruptions of solar energetic particles and radiation which have a certain propagation direction. Flares and CME's do not necessarily occur together and have different source layers within the Sun. Flares arise from the *Chromosphere*[6] and burst through the corona into space, whereas CME's have their origin in Sun's corona. The cause of the eruption of CME's and flares is strongly connected with the reconnection of magnetic fields in bipolar regions of the *photosphere*, but all mechanisms which lead to these phenomena are still a topic of current research.

Solar radiation is the most important energy output form as it is the reason of Earth's habitability, due to its heating effect on the atmosphere and its additional interaction potential with the Biosphere, known as *Photosynthesis*. This radiation takes the form of electromagnetic waves and is mainly in the visible frequency domain, as can be seen in Figure 2.2. The electromagnetic spectrum of Sun's radiation (Fig. 2.2) has the shape of a black body with a temperature around 5762 K (ASCHWANDEN, 2006). Its source is in the photosphere, which is the brightest of all layers of our central star. The solar radiation

[4] Source: Picture of wikipedia user Medium69 was slightly changed by myself. URL: `https://en.wikipedia.org/wiki/File:Structure_of_the_magnetosphere-en.svg`, Date of download: 14.04.2017

[5] E.g. in the form of ionizing radiation which has deleterious effects on human beings.

[6] The sun is divided into several components and layers. From the core to the outermost part, these layers are known as: core, radiation zone, convective zone, photosphere, chromosphere and corona.

at Earth's surface depends strongly on the geometry, weather condition and abundancy of absorbers in the atmosphere (see also later sections of this chapter for more details).

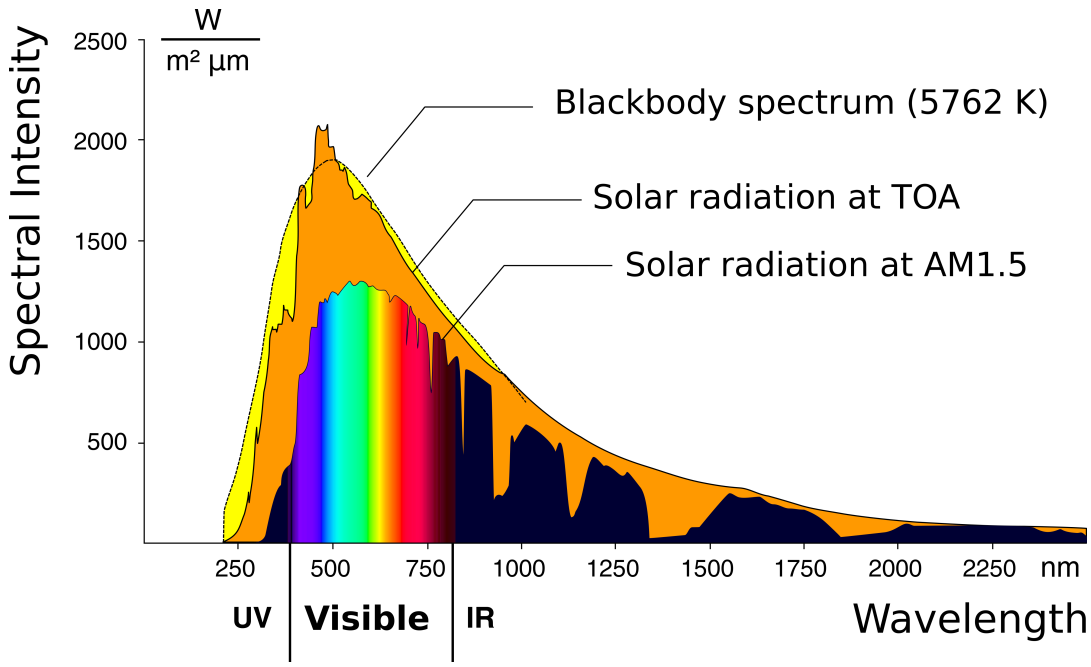


Fig. 2.2: *Comparison plot of the solar irradiance at top of the atmosphere (orange), the solar radiation at an air mass of 1.5 (blue/multi color) and the spectrum of a black body at* $5762\,\mathrm{K}$ *(yellow).*[7]

In addition to the Biosphere[8] also technologies as e.g. power generators and satellites in an orbit around the Earth are strongly influenced by Sun's energy forms as they might be damaged by large energy events like CME's or flares. Table 2.1 gives an overview of the energy amount emitted by the Sun for the individual phenomena as well as the time they need to reach the Earth.

	Speed $(\mathrm{km/s})$	**Time to Earth**	**Total power** (W)
solar wind	$\leq 450 - 800$	$\sim 3\,\mathrm{days}$ to $5\,\mathrm{days}$	4×10^{20}
large CME	$200 - 3000$	$\sim 1\,\mathrm{days}$ to $5\,\mathrm{days}$	1×10^{23}
large flare	3×10^{8}	$\sim 8.3\,\mathrm{min}$	1×10^{23}
solar radiation	3×10^{8}	$\sim 8.3\,\mathrm{min}$	4×10^{26}

Tab. 2.1: *Speed, time to reach Earth and the total power output for different forms of energy of the Sun. From* BOTHMER and DAGLIS [2007].

[7] Source: Picture of wikipedia user Baba66 was slightly changed by myself. URL: `https://de.wikipedia.org/w/index.php?title=Datei:Sonne_Strahlungsintensitaet.svg&lang=de`, Date of download: 13.04.2017

[8] *The biosphere is the global ecological system integrating all living beings and their relationships, including their interaction with the elements of the lithosphere, geosphere, hydrosphere, and atmosphere.* Definition from wikipedia: `https://en.wikipedia.org/wiki/Biosphere`, Date: 20.06.2018

Whereas there is no warning time for radiation events and flares due to their propagation speed which is the speed of light, there is a sufficiently long advance warning time for CME's and enhanced solar wind events. This is important because geomagnetic storms caused by solar wind shock waves lead to strong geomagnetic fluctuations which, in turn, induce geomagnetic currents in electrical conductors which might destroy them. For that reason, it is crucial for modern society to constantly observe the Sun from satellites in Earth's orbit via remote sensing (e.g. Solar Dynamics Observatory, SDO) or from Lagrange-Point L1 via in-situ measurements (e.g. Deep Space Climate Observatory, DSCOVR).

To quantify the amount of radiation from an emitter like the sun, the terminology *radiant flux* describes the overall emitted energy per time (or power, $\sim 3.9 \times 10^{26}\,\mathrm{W}$, ASCHWANDEN, 2006). This energy is emitted in all directions in space and goes through spherical shells in different distances to the emitter. In general, the radiant flux Φ_e through an emitting surface A_e is called flux density $E_e = \Phi_e / A_e$. This flux density through the surface of the Sun with its radius r_s is called *irradiance* $E_s = \Phi_e / (4\pi r_s^2)$. When referring to the flux through an unit area in the distance of the Earth ($1\,\mathrm{AU}$[9]), this quantity is called *solar constant* and it is on average $1368\,\mathrm{W\,m^{-2}}$ (see e.g. HOLLOWAY and WAYNE, 2015). Due to the large distance of Sun and Earth, a perpendicular throughput can be assumed. If a fraction of this irradiance is emitted by a reference plane $E_s = \mathrm{d}\Phi_e / \mathrm{d}A$ to a solid angle $\mathrm{d}\Omega$, this term is called *radiance* $L = \mathrm{d}^2\Phi_e / (\cos(\vartheta)\mathrm{d}A\mathrm{d}\Omega)$ with the angle ϑ between propagation direction and the normal vector of the reference plane. In short, radiance is the irradiance through a certain solid angle.

2.2 The structure of Earth's atmosphere

Earth's atmosphere underwent different steps from the beginning primordial state to the current composition. After Earth formed through gravitational agglomeration processes to a *Protoplanet*, the first stable atmosphere was formed by degasing of volcanoes. It consisted mainly of water vapor $\mathrm{H_2O}$, carbon dioxide $\mathrm{CO_2}$ and hydrogen sulphide $\mathrm{H_2S}$. The high water vapor content was lost during the cooling of the planet which led to a phase of constant rain. This phase ended in the formation of the first oceans and seas which were in turn a habitat for the first bacteria. These lifeforms led to an increase in methane $\mathrm{CH_4}$ and nitrogen $\mathrm{N_2}$ which were in addition to hydrogen sulphur the basic ingredient for the formation of Earth's Biosphere. The oxygen $\mathrm{O_2}$ content increased through the time as the number of micro-organisms increased, which photosynthesised carbon dioxide $\mathrm{CO_2}$ to oxygen $\mathrm{O_2}$ (further details in SEINFELD and PANDIS, 2006).

[9] One astronomical unit ($1\,\mathrm{AU}$) is per definition $149.6\,\mathrm{Mio\,km}$, which is the mean distance of Sun and Earth.

2.2.1 The chemical composition of the atmosphere

The composition of the atmosphere as we know it today consists primarily of three main gases: nitrogen N_2, oxygen O_2 and argon Ar (see Fig. 2.3). All other gases are much less frequent and are called *trace gases*. They comprise no more than 0.1 percent of the total amount of gases in the atmosphere. However, their impact on Earth's climate and the *greenhouse effect* is crucial.

Trace gases can be divided into constant and highly variable components. Water vapor H_2O, ozone O_3 and nitrogen dioxide NO_2 can be considered as variable, because they depend strongly on weather conditions and annual variations[10](and anthropogenic sources for NO_2). On the other hand, the noble gases (Ar, Ne, He) are chemical inert, which means, that they are evenly distributed in the atmosphere. H_2O[11], CH_4 and CO_2 are among the trace gases with the most drastic impact on the greenhouse effect (see section 2.2.3), which makes them an important subject of current research. Another important greenhouse gas got dubious fame in the last century, because of its rapid decrease in the atmosphere above the poles. Ozone O_3 is a trace gas with average abundance but its impact on Earth's energy budget is large due to its shielding effect in the ultraviolet spectral range[12]. Details on trace gases, their sources, sinks and reaction chains can be found in Section 2.6.

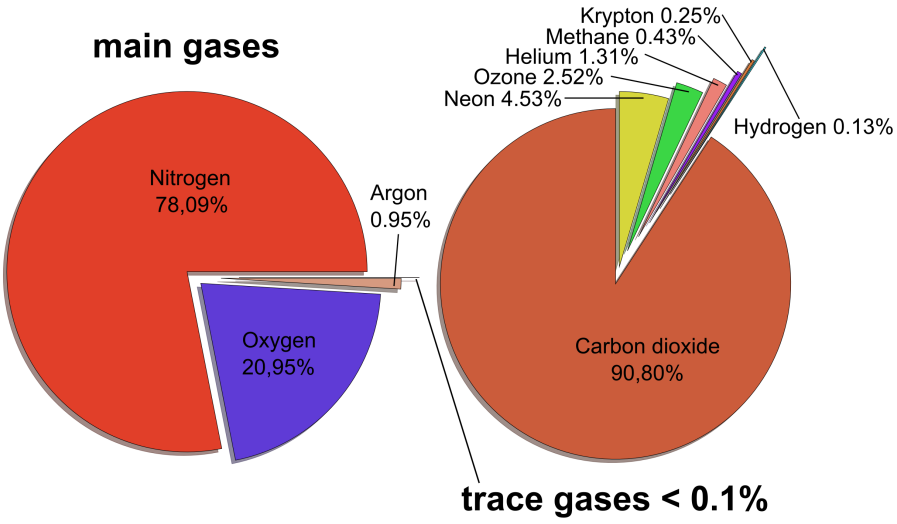


Fig. 2.3: Left: *Main gases of Earth's atmosphere.* ***Right:*** *Composition of trace gases.* H_2O *was neglected due to its high variability but it is on average 0.4%. Values were extracted from* RAITH and BAUER [2001].

[10] E.g. seasonal vegetation cycle due to plant growth in spring and its alongside reduction of CO_2.

[11] Water vapour makes a small fraction of the Earth's atmosphere (0 % to 3 %) but its concentration varies on lateral and temporal scales.

[12] A schematic representation of spectral ranges and classifications can be found in the appendix in Fig. AF.2.

2.2.2 The physical layering of the atmosphere

The atmosphere can be split up into several layers according to Earth's vertical temperature profile (see Fig 2.4). The extrema of the temperature curve describe the boundaries of each layer. From surface to space, the layers are known as *Troposphere*, *Stratosphere*, *Mesosphere* and *Thermosphere*. The outermost layer of Earth's atmosphere is called *Exosphere* (not shown in Fig 2.4). It starts at the *Thermopause* at a height of around $700\,\mathrm{km}$ and reaches up to $10\,000\,\mathrm{km}$. Since the density is low, Exosphere is neither important for meteorology nor other scientific phenomena which might interfere with life on Earth. Because of its low density, there is nearly no friction for satellites which makes it a preferred working environment for Earth observing satellites.

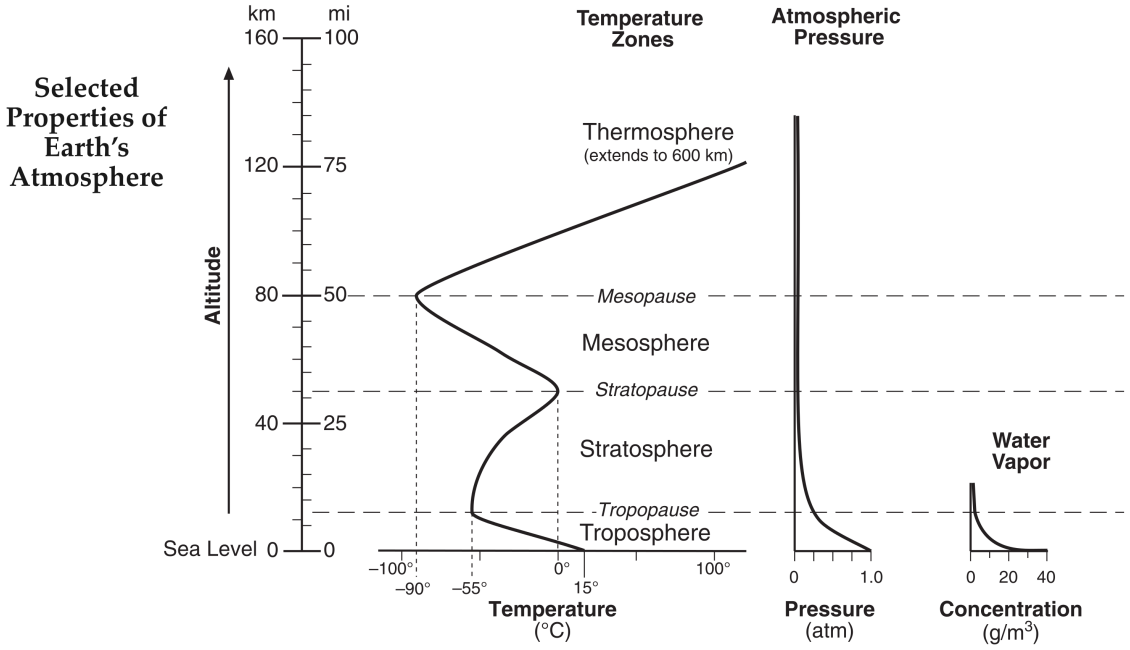


Fig. 2.4: *Schematic representation of temperature, pressure and water vapor vertical profiles and layering of Earth's atmosphere.*[13]

The four layers depicted in Figure 2.4 are rather important for Earth's Biosphere, because they include all phenomena which affect life.

The **troposphere** is the lowermost layer, beginning at the surface and extending up to an average height of $12\,\mathrm{km}$. The temperature decreases within this layer nearly linearly, with an inversion point at the *Tropopause*, which defines the boundary to the stratosphere. The reason for this temperature profile is the heating of the surface by sunlight and the adiabatic cooling with decreasing pressure and altitude (barometric formula).The troposphere contains almost the entire water vapour content of the atmosphere and its pressure

[13] From the New York State Education Department. Reference Tables for Physical Setting/EARTH SCIENCE. Available from `http://www.p12.nysed.gov/assessment/reftable/earthscience-rt/esrt2011-engr.pdf`, Date of download: 15.04.2017

decreases exponentially, indicating that the largest fraction of atmospheres mass is in the lowermost layer. Within Earth's troposphere there is another significant layer which is called *planetary boundary layer* (PBL). This sublayer is defined as the part which is directly influenced by the surface and it varies in its thickness from a few hundred meters up to several km. Here, due to convection, there is a strong mixing of all components and changes in temperature, moisture etc. might happen fast. Furthermore, the boundary layer contains the majority of anthropogenically produced trace gases and aerosols, because their sources are mostly located on Earth's surface. In highly polluted regions (e.g. industrial, cities), the PBL can be seen with the naked eye, because of the high amount of aerosols concentrated within the boundary layer (see Chapter 2.7). The second part of the troposphere lies above the PBL and is called *free troposphere*. This layer includes anthropogenic pollution because of air planes and uprising air by low pressure systems and thunderstorms. Rising air has a natural boundary at the tropopause, because the increasing temperature in the stratosphere prevents air from rising further.
The **stratosphere** extends from the tropopause up to an altitude at around $55\,\mathrm{km}$. The temperature increases steadily with altitude as well as the ozone volume mixing ratio does up to O_3's maximum amount. The stratospheric regime of high ozone concentration is called *ozone layer* and its source is oxygen which is photo dissociated by ultraviolet radiation from the Sun ($O_2 + \mathrm{hv} \longrightarrow 2\,O$), and connects with another oxygen molecule to form ozone ($O + O_2 + M \longrightarrow O_3 + M$). The ozone molecule is dissociated as well but will form another ozone molecule immediately. The absorption of ultraviolet light in turn creates heat, which is the reason for an increasing temperature with altitude (see section 2.6). Because of this positive temperature gradient, the stratosphere is quite stable and no convection is occurring.
The stratosphere's upper limitation point is called *stratopause*, which separates it from the **mesosphere**. Within this layer, the temperature decreases again up to the *mesopause* which is the coldest place on Earth with temperatures down to $-100\,^{\circ}\mathrm{C}$ (LATIF, 2009) at a height around $85\,\mathrm{km}$. Within the mesosphere, certain waves propagate (e.g. gravity waves, planetary waves) and *noctilucent clouds* can be observed. Above the mesosphere, the next layer is called **thermosphere**, which ends at a height of around $500\,\mathrm{km}$. Here, energetic solar radiation in the ultraviolet spectral range photodissociates molecules which leads to free charge carriers in the form of ions and electrons[14]. Within the mesosphere, temperature increases again due to absorption processes up to $2500\,\mathrm{K}$.

2.2.3 The greenhouse effect

The terminology *greenhouse effect* is often used in the media to describe the warming of the atmosphere in the same breath as pollution. This indicates, that the greenhouse effect

[14] Which creates the *ionosphere*, a layer defined by its ionized constituents rather than the temperature.

is only caused by human and that it has an overall negative effect on Earth's climate. But to be more accurate, it must be divided into the natural and the anthropogenic greenhouse effect. Whereas the latter one can be considered as an adverse alteration of the climatic conditions, the natural greenhouse effect is the reason why life on Earth (as we know it) is even possible.

In section 2.1 it was stated, that the solar constant is $S = 1368\,\mathrm{W\,m^{-2}}$ at a distance of $1\,\mathrm{AU}$. If we consider the Earth as a disk with radius r_E, it receives the power of $S\pi\mathrm{r}_E^2$. Since our planet is not a perfect absorber, it reflects a certain fraction A (albedo) of the incoming radiation back to space but absorbs the radiant flux $S(1-A)\pi\mathrm{r}_E^2$. With an average albedo of $A = 0.3$, this radiation is in a global mean $239\,\mathrm{W\,m^{-2}}$ (cf. Fig. 2.5), with the assumption that the Sun radiated area ($\pi\mathrm{r}_E^2$) is four times smaller than Earth's spherical surface area ($4\pi\mathrm{r}_E^2$). In other words, there is a steady radiative energy input by the sun which would yield to a more and more increasing temperature without an

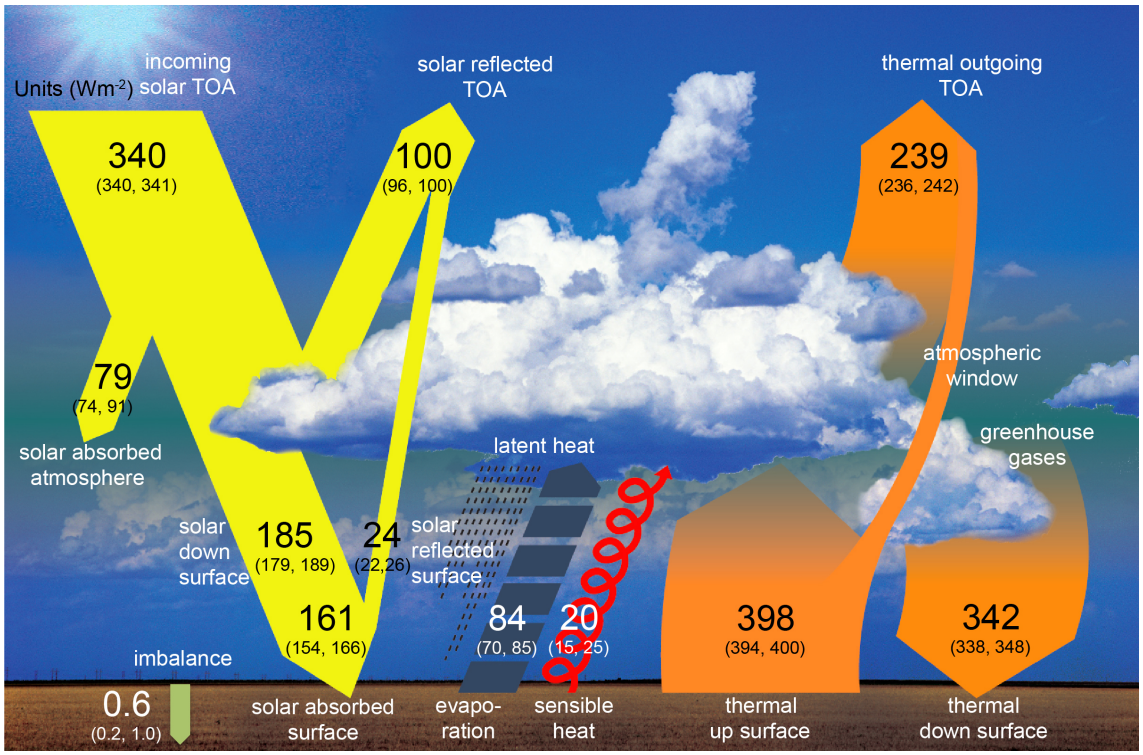


Fig. 2.5: *Energy budget of Earth's atmosphere. Global and annual average of total solar irradiance:* $1360\,\mathrm{W\,m^{-2}}/4 = 340\,\mathrm{W\,m^{-2}} \mathrel{\widehat{=}} 100\%$ *of incoming radiation. From* IPCC [2013].

equalizing energy output by the Earth. This output can be understood by considering *Stefan-Boltzmann's law*. The formula describes the energy output of a black body radiator which depends only on the temperature T to the power of four, multiplied by the *Stefan-Boltzmann constant* σ ($\sigma = 5.670367 \cdot 10^{-8}\mathrm{W\,m^{-2}\,K^{-4}}$) $E = \sigma T^4$. In equilibrium, the part of the solar radiation, which is not absorbed by the Earth, must therefore equal the

outgoing temperature depending radiation by Stefan-Boltzmann's law $(1-A)S/4 = \sigma T_e^4$. T_e is called *equilibrium temperature* and has the value $T_e = 255\,\mathrm{K}$. The average temperature on Earth' surface is, however, around $288\,\mathrm{K}$, which leads to a difference of $33\,\mathrm{K}$. The phenomenon of surface temperature being higher than the equilibrium temperature is commonly known as natural greenhouse effect.
For a better understanding of the incoming and outgoing radiation, *Planck's law* describes the emitted radiation of a black body at a given temperature T,

$$B_\lambda(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/(\lambda k_B T)} - 1} \tag{2.1}$$

with *Planck's constant* h, the speed of light c, the wavelength λ and the *Boltzmann's constant* $k_B = 1.38 \cdot 10^{-23}\mathrm{J/K}$. The Sun's incoming radiation has a maximum in the visible frequency range at a temperature of around $T = 5900\,\mathrm{K}$ (compare Fig. 2.2). Planck's law states, that at the equilibrium temperature of Earth's atmosphere, the emitted radiation is in the infra-red spectral range at around $10\,\mu\mathrm{m}$ $(288\,\mathrm{K})$.
Figure 2.5 shows, that there are even more processes involved[15]. For example, the incoming radiation is reflected at clouds or at regions with a high surface reflectance (e.g. snow covered regions) as well. The explanation of the natural greenhouse effect can be found by comparing the radiation emitted by Earth's surface $(398\,\mathrm{W\,m^{-2}})$ and the fraction of this which leaves the atmosphere $(239\,\mathrm{W\,m^{-2}})$. The difference of this thermal radiation is absorbed by the atmosphere and can be understood as the underlying cause of the natural greenhouse effect. This absorption can be explained by a certain number of atmospheric compounds. The most important one is water vapour $\mathrm{H_2O}$, which is responsible for $\sim 66\,\%$ of the absorption of thermal radiation, followed by carbon dioxide $\mathrm{CO_2}$ $(\sim 30\,\%)$ and other trace gases $(\leq 4\,\%$ e.g. methane $\mathrm{CH_4}$, nitrous oxide $\mathrm{N_2O}$, ozone $\mathrm{O_3}$; from LATIF, 2009).

In contrast to the natural greenhouse effect, the *anthropogenic greenhouse effect* is defined as the change of trace gas concentrations (e.g $\mathrm{CO_2}$ or $\mathrm{CH_4}$) or the emission of new trace gases (e.g. chlorofluorocarbon, CFC) by humans, which lead to a change in Earth's energy budget. The best known example for anthropogenic pollution of the atmosphere and the in turn change in trace gas concentration based on this pollution was the tremendous use of CFC's in the previous century. Their emission has lead to a decrease of the stratospheric ozone concentration which increased the transmission of solar radiation in the ultraviolet spectral region at the poles. Furthermore, since CFC are also greenhouse gases with absorption bands in the infra-red, outgoing thermal radiation would have been trapped in the atmosphere. The CFC increase would have warmed the planet in a drastic

[15] For a detailed discussion on Earth's energy budget, the reader is encouraged to read more in relevant literature e.g. RAITH and BAUER [2001], LATIF [2009] and HOLLOWAY and WAYNE [2015].

way when CFC use would not have been restricted by the *Montreal protocol* in the year 1987[16]. To quantify the impact of changes in trace gas concentrations on Earth's climate, *radiative forcing* (RF) is defined as the change of the energy budget between incoming radiation absorbed by the Earth's climate system and the emitted infra-red radiation. Table 2.2 shows RF values of the most important atmospheric compounds and effects. The RF numbers in this table describe the change between the pre-industrial era at the year 1750 and today (2011). Here, red colour represents positive RF and green values show negative RF. It can be seen that there is not only a positive forcing but also effects shielding the atmosphere from radiation and leading to a decreasing average temperature. Since a couple of these effects balance each other and forcing values of some species are known with a higher accuracy than others, the investigation of possible trends in Earth's energy budget is complex. Nevertheless, the total net anthropogenic radiative forcing is positive, which means that the global average temperature will continue to rise until positive and negative forcing effects balance each other. As commonly known, this can only be done by a decrease in anthropogenic emissions of greenhouse gases (or by questionable geoengineering solutions).
Although this thesis is not directly linked to the greenhouse effect and its implication for Earth's climate, the investigated trace gases and atmospheric effects (see also Sec. 2.6) play an important role in Earth's energy budget.

	RF Terms	Comp.	RF values ($\mathrm{W\,m^{-2}}$)	loc
Anthropogenic	Long-lived greenhouse gases	CO_2	1.68 [1.33 to 2.03]	VH
		CH_4	0.97 [0.74 to 1.20]	H
		N_2O	0.17 [0.13 to 0.21]	VH
	Ozone (strat. + trop.)	O_3	0.30 [-0.15 to 0.65]	H
	Short-lived gases	CO	0.23 [0.16 to 0.30]	M
		NO_x	-0.15 [-0.34 to 0.03]	M
	Total Aerosol direct effect		-0.27 [-0.77 to 0.23]	H
	Total Aerosol Cloud albedo eff.		-0.55 [-1.33 to -0.06]	L
	Surface albedo		-0.15 [-0.25 to -0.05]	M
Natural	Solar Irradiance		0.12 [0.06 to 0.30]	M
	Total net anthrop. RF		2.29 [1.13 to 3.33]	H

Tab. 2.2: *Global averaged radiative forcing estimations for the main compounds and effects for 2011 relative to 1750. The level of confidence (loc) is divided into VH: very high, H: high, M: medium and L: low). Table reproduced from IPCC,* STOCKER [2014].

[16] The connection between O_3 depletion and the greenhouse effect is even more complicated as stated here since the radiative forcing due to less O_3 is negative but the net forcing including CFC's is

2.3 Interaction mechanisms of radiation and Earth's atmosphere

In the previous section, the focus was on the source of radiation and on the physical aspects of Earth's atmosphere, with a besides mention of interaction mechanisms like scattering and absorption. In this section, radiation itself and the way how it interacts with molecules in the atmosphere is on focus. This part is rather a brief summary of the most important contexts than a detailed discussion. The reader is referred to the relevant literature (e.g. Haken and Wolf [2002], Haken and Wolf [2006], Demtröder [2006], Demtröder [2010]) for more information on the topic.

2.3.1 What is electromagnetic radiation?

Even though light has always been an omnipresent phenomenon, the concept of light and electromagnetic radiation described as an electromagnetic wave was developed in the last two centuries only. In the beginning 18th century, *William Herschel* and *Johann Wilhelm Ritter* discovered, that there is measurable radiation with wavelength shorter and longer than those of the visible range ($\sim$1800). However, it took more than six decades (1864) until *James Clark Maxwell* found mathematical expressions which explain light as fluctuating electric and magnetic fields travelling with the speed of light. From that moment, all phenomena of classical electromagnetism could be explained. Furthermore, it was possible to understand the *interference* of light in the Double-slit experiment by *Thomas Young* and further effects (like *coherence*) with the wave nature of electromagnetic radiation.
On the other hand, in the late 18th century, the photoelectric effect could not be fully explained by *Alexandre Edmond Becquerel* and *Heinrich Hertz* (1886) with Maxwell's equations. It needed two further decades until *Max Planck* introduced the idea of light as discrete quantities of energy and *Albert Einstein*, who defined this effect as light which propagates in the form of particles, the so called *photons* (1905). Even though Einstein's theory was able to extend the understanding of light, it was not able to explain all of its aspects. This issue lead to the concept of *wave-particle duality*, which states that electromagnetic radiation can be described as particles and waves to fulfil the requirements of the wavelike and the particle-like character of radiation.
In the following subsections, the concept of radiation as electromagnetic wave and as particle is shortly introduced.

2.3.1.1 Radiation as electromagnetic waves

Electromagnetic waves can be explained with Maxwell's equations:

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{2.2}$$

$$\nabla \cdot \vec{E} = \frac{\eta_f}{\epsilon} \tag{2.3}$$

$$\nabla \times \vec{B} = \mu \left(\vec{j} + \epsilon \frac{\partial \vec{E}}{\partial t} \right) \tag{2.4}$$

$$\nabla \cdot \vec{B} = 0 \tag{2.5}$$

$\vec{E}$ and $\vec{B}$ are the electric and magnetic fields, respectively. η_f is the total charge density and $\vec{j}$ is the current density. Furthermore, μ and ϵ are the magnetic permeability and the permittivity, which are the product of the relative permeability μ_r (permittivity ϵ_r) and the magnetic constant μ_0 (electric constant ϵ_0).

$$\mu = \mu_0 \mu_r \qquad \text{and} \qquad \epsilon = \epsilon_0 \epsilon_r$$

When considering Sun's radiation, the light propagates through space, which is mostly an uncharged region ($\eta_f = 0$) without currents ($\vec{j} = 0$). The Maxwell equations then transform into vacuum equations. By taking the curl ($\nabla\times$) of equations 2.2 and 2.4, the resulting formulations are wave equations for the magnetic and electric fields.

$$\Delta \vec{E} - \epsilon \mu \ddot{\vec{E}} = 0 \tag{2.6}$$

$$\Delta \vec{B} - \epsilon \mu \ddot{\vec{B}} = 0 \tag{2.7}$$

A comparison to the wave equation for u ($\ddot{u} = c_1^2 \Delta u$) leads to a relationship between the speed of light in free space c and the *refractive index* n defined as the ratio of c and the velocity v with which light propagates through a medium.

$$c_1 = \frac{1}{\sqrt{\mu_r \mu_0 \epsilon_r \epsilon_0}} = \frac{c}{\sqrt{\mu_r \epsilon_r}} = \frac{c}{n} \quad \text{with} \quad c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}, \quad n = \sqrt{\mu_r \epsilon_r} \tag{2.8}$$

A possible solution of the wave equations 2.6 and 2.7 can be found as plane periodic waves.

$$\vec{F} = \vec{F_0} \cdot e^{i(\vec{k}\vec{r} - \omega t)} \qquad \text{with} \qquad F = E, B \tag{2.9}$$

These solutions are satisfying the Maxwell equations, which shows, that wave vector $\vec{k}$ is perpendicular to both field vectors $\vec{E_0}$ and $\vec{B_0}$ (see Fig. 2.6). In other words, the energy transfer in direction of $\vec{k}$ is perpendicular to the oscillations, which characterize the moving wave.

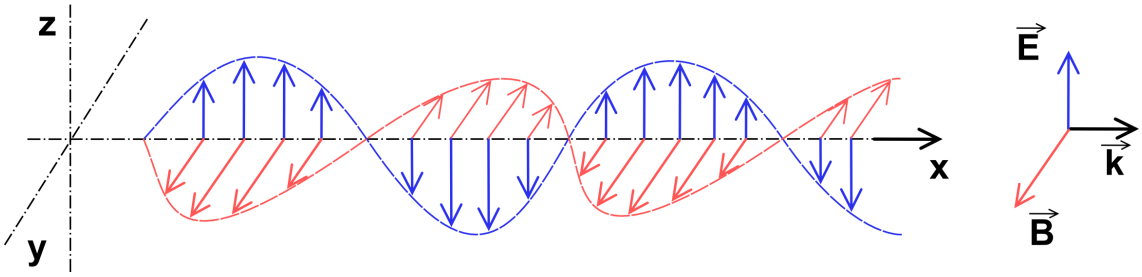


Fig. 2.6: *Schematic representation of an electromagnetic wave with its field vectors perpendicular to the propagation direction.*[17]

The energy density of an electromagnetic wave is the sum of the energy densities of its individual components.

$$w_{em} = w_{el} + w_m = \frac{1}{2}\epsilon_0 E^2 + \frac{1}{2\mu_0} B^2 \tag{2.10}$$

Since the electric and magnetic field components are perpendicular to each other, the field amplitudes for a plane parallel electromagnetic wave in vacuum can be connected with the speed of light $|B| = c^{-1}|E|$. With this relation, the energy density can be transformed to $w_{em} = \epsilon_0 E^2$. The *Poynting-vector* is defined as the vector which describes the direction of energy propagation.

$$\vec{S} = \epsilon_0 c^2 \vec{E} \times \vec{B} \tag{2.11}$$

The intensity of electromagnetic waves is now found as the energy per time, which is transported to an area or the absolute value of the Poynting vector. In terms of the energy density, the intensity is

$$I_{em} = c \cdot w_{em} = c \cdot \epsilon_0 E^2. \tag{2.12}$$

2.3.1.2 Radiation as particles

In addition to the wave nature of electromagnetic radiation, light can also be explained by discrete energy quantities (photons) with the energy $E_{ph} = h\nu$. The energy of these photons depends only on Planck's constant $h \approx 6.626 \times 10^{-34}\,\mathrm{J\,s}$ and the frequency ν. The energy density of a beam of photons with a photon density n is $w_{ph} = n \cdot h\nu$.
The concept of photons is substantial for atmospheric sciences, because phenomena like emission and absorption of light by atmospheric species can be explained easily (see section 2.3.2). Furthermore, the measurement of electromagnetic radiation in the visible

[17] Source: Picture of wikipedia user SuperManu. URL: `https://de.wikipedia.org/wiki/Datei:Onde_electromagnetique.svg`, Date of download: 07.05.2017

spectral range requires optical instead of electrical engineering (with electronic devices e.g super-heterodyne receivers for the latter case). Whereas the first engineering subject enables the measurement of radiation in its particle nature, the latter one is confined to electromagnetic waves. Between both areas lies the *terahertz gap*, which is the spectral region ($0.1\,\mathrm{THz}$ to $10\,\mathrm{THz}$) for which there is no suitable measurement technology developed yet. In other words, the frequency range with wavelengths higher than $\sim 0.1\,\mathrm{THz}$ can not be measured by using the wavelike nature of radiation (e.g. with antennas). Here, the particle representation of electromagnetic waves is used.
Measuring devices for photons (e.g. photo diodes) utilize effects like the *photoelectric effect* to measure the intensity

$$I_{ph} = c \cdot w_{ph} = c \cdot n \cdot h\nu \tag{2.13}$$

of the incoming radiation. The number of photons per seconds which reach the detector is proportional to the number of electrons emitted as free charge carriers, which can be measured by electrical circuits. Table 2.3 highlights quantities of both, the wavelike and the particle nature of electromagnetic radiation.

	wave	**particle**
energy (J)	$dW = wdV = \epsilon_0 E^2 Acdt$	$E_{ph} = h\nu$
energy density ($\mathrm{J/m^3}$)	$w_{em} = \epsilon_0 E^2$	$w_{ph} = n \cdot h\nu$
intensity ($\mathrm{W/m^2}$)	$I_{em} = c \cdot \epsilon_0 E^2$	$I_{ph} = c \cdot n \cdot h\nu$

Tab. 2.3: *Short comparison of important values of electromagnetic radiation as waves and particles.*

2.3.1.3 Polarisation of radiation

Equation 2.9 explains how the electromagnetic field oscillates in one single plane. The field vector is always pointing in the same direction, a case which is referred to as *linear polarisation*. Possible solutions of the wave equations 2.6 and 2.7 can also be formulated by using an additional phase parameter ϕ. In this case, the field vectors can be rewritten as: $\vec{F} = \vec{F_0} \cdot \exp i(\vec{kr} - \omega t + \phi)$. Generally, phase shifts with $\phi = 90°$ are referred to as *circular polarisation* and $\phi \neq 90°$ and $\phi \neq 0°$ means *elliptical polarisation*.

The full state of polarization is fully described with the help of *Stokes vector*.

$$\vec{I} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} \tag{2.14}$$

The first element represents the scalar intensity I, which was used before (see Sec. 2.3.1.1). The other three elements Q, U and V are used to characterise the *degree of polarisation* (dop), the preferred orientation and the kind of polarisation (compare PETTY [2006] for an introduction). The degree (dop) is defined as

$$\Pi = \frac{\sqrt{Q^2 + U^2 + V^2}}{I}. \tag{2.15}$$

The elements Q and U represent the linear polarisation states, whereas V describes the circular polarisation (see examples in Tab. 2.4). Additional degrees of linear and circular polarisation are defined as the ratio of the matching vector elements and the intensity I ($\sqrt{Q^2 + U^2}/I$ and V/I).

Expression	(I, Q, U, V)
Horizontal polarisation	$(1, 1, 0, 0)$
Linear polarisation at $+45°$	$(1, 0, 1, 0)$
Right circular polarisation	$(1, 0, 0, 1)$
Unpolarized	$(1, 0, 0, 0)$

Tab. 2.4: *Examples of elements of Stokes vector and their expression.*

2.3.2 Interaction phenomena of radiation with atmospheric compounds

It was already noted that radiation undergoes different types of interaction processes with species in the atmosphere. So far, terminologies like absorption, emission and scattering were used without further clarifications. The subsequent sections provide explanations for these effects and are a short introduction into interaction phenomena between light and molecules. For further details, the reader is referred to literature like DEMTRÖDER [2010], LIOU [2002], BURROWS et al. [2011], and PLATT and STUTZ [2008].

2.3.2.1 Absorption and Emission

In section 2.1 it was stated, that Sun's electromagnetic spectrum (Fig AF.2) is comparable to a blackbody. However, a sun observing spectroscopic instrument would show, that the solar spectrum is not a continuous curve of electromagnetic radiation. In contrast to a blackbody radiator, the irradiance spectrum includes some dark parts which can be allocated to certain wavelengths. Furthermore, the number of dark parts in the radiance spectrum, measured on Earth's surface, is even larger than in the spectrum measured in space. The missing parts in the solar spectrum are called *Fraunhofer lines* and *telluric lines* and can be explained by absorption of radiation by different elements in Sun's photosphere and Earth's atmosphere, respectively.
In addition to absorption, there are phenomena in nature which lead to an increasing intensity at certain wavelengths. A popular example are *auroras*[18], which can be explained as an excitation of molecules by solar particles in the atmosphere and the subsequent emission of photons at certain wavelengths.

Energy levels and transitions

In order to understand absorption and emission, the *Bohr* model of atoms is helpful. It states, that an atom consists of a positively charged core (nucleus), which contains all the mass (protons and neutrons). Electrons circle around this core on certain orbits without losing energy. In Figure 2.7, a schematic representation is shown, which explains that ab-

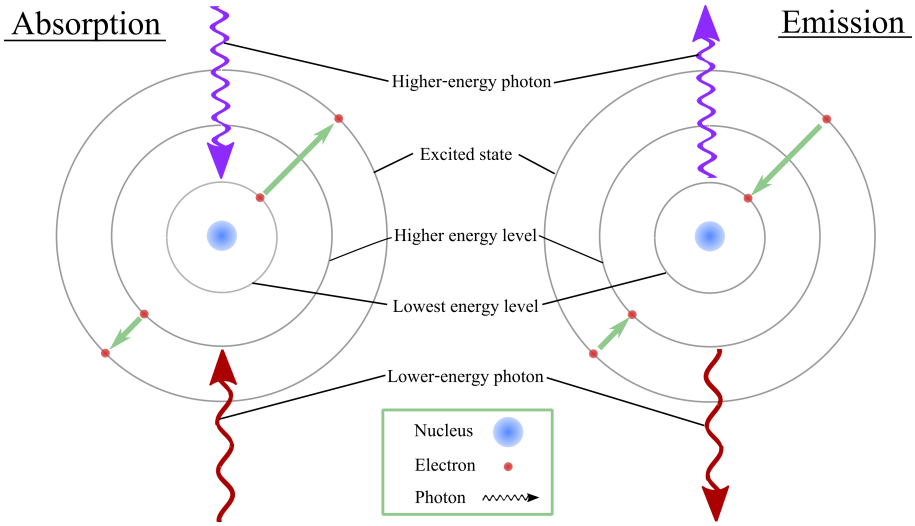


Fig. 2.7: *Illustration of absorption and emission of different energetic photons in the simple representation of Bohr's model.*

[18] Auroras are luminous effects which can be observed in the sky, mostly in higher latitudes. The colour of auroras is determined by the excited atom or molecule. E.g. excited atomic oxygen emits at $630\,\mathrm{nm}$ which produces a reddish glow.

sorption of photons is equivalent to an excitation of the electron to a higher energy level. On the other hand, the *spontaneous emission* of photons will lead to a lower energy level as the energy difference between both states is emitted as radiation.
Bohr's atomic theory is a simple attempt to explain the discrete energy transfer when dealing with the particle nature of electromagnetic waves. Unfortunately, it is not able to explain energy state values for atoms with more than one electron. Furthermore, quantum mechanic considerations make clear, that there can not be fixed orbitals and known energy levels, because *Heisenberg's uncertainty principle* states that location and velocity/momentum (and therefore energy) cannot be measured precisely at the same time ($\Delta x \cdot \Delta p \sim h$).

The *atomic orbital model* is a more suitable description because it satisfies the uncertainty principle by considering orbitals as a probability of presence $|\Psi(r)|^2$ (with $\Psi(r)$ as wave function) for electrons. The beforehand mentioned energy levels E are then the Eigenvalues of the *Hamiltonian* H. The time-independent *Schrödinger* equation

$$H\Psi(\vec{r}) = E\Psi(\vec{r}) \tag{2.16}$$

can be solved in the case of the hydrogen atom analytically in spherical coordinates by using a product of angular factors $\Theta(\theta)$ and $\Phi(\phi)$, and a radial factor $R(r)$. The overall wave function is then $\Psi(r, \theta, \phi) = R(r)\Theta(\theta)\Phi(\phi)$. The solution of the angular part consists of spherical harmonics $Y_l^m(\theta, \phi)$ and the full solution is described by:

$$\Psi(r, \theta, \phi) = R_{n,l}(r)Y_l^m(\theta, \phi) \tag{2.17}$$

While the principal quantum number n indicates in which electronic shell the electron is located (energy of photon), l and m denote the *angular quantum number* and the *magnetic quantum number* respectively. For the example of the hydrogen atom, l describes the angular momentum of an electron in its shell defined by n with the z-component of angular momentum defined by the quantum number m. The energy itself depends only on the principal quantum number n in the simple case of the hydrogen atom.
Schrödinger's equation can be solved analytically only for one-electron-systems but there are approximations (e.g. *Born-Oppenheimer*) for multi-particle systems as well.

Atoms have the unique characteristic, that the movement of electrons is the only part which needs to be considered for energy aspects because of the inertia of the heavy nucleus. The energy levels of atoms are therefore fully described by the radial distance between both objects. A molecule is a more complex system which might undergo different kinds of movement. As a simple example, the hydrogen molecule consists of two hydrogen atoms with a covalent bond, meaning that both atoms share an electron. As the

distance core-electron is important for energy considerations, the internuclear distance between both atoms is crucial as well.
Furthermore, the molecule itself is free to rotate and vibrate which leads to an expansion of the energy term which describes the system. The degrees of freedom (dof) changes from three (translational movement in all directions of space) for a single atom to $3N = f_{trans} + f_{vib} + f_{rot}$ for a molecule with N atoms. These motions must be added as additional terms in Schrödinger's equation which results in a new set of quantum numbers.

The force between two charged particles or atoms is called *Coulomb-force* and it decreases with the squared distance. When considering a molecule's potential energy for covalent bonding, the force between both atoms is strong, when they are close to each other. When the distance increases, the force decreases until the Coulomb-force is exceeded (this effect is known as *dissociation*). In this case, the energy levels for both atoms can be considered as that for single atoms. The potential which describes this behaviour is called *Morse potential* $V(r)$ and it is shown in Figure 2.8.

$$V(r) = D_E(1 - e^{-a(r-r_0)})^2 \tag{2.18}$$

Here, D_E is the dissociation energy which needs to be added to the system for separating both atoms. The constant a is a value which defines the width of the potential. An

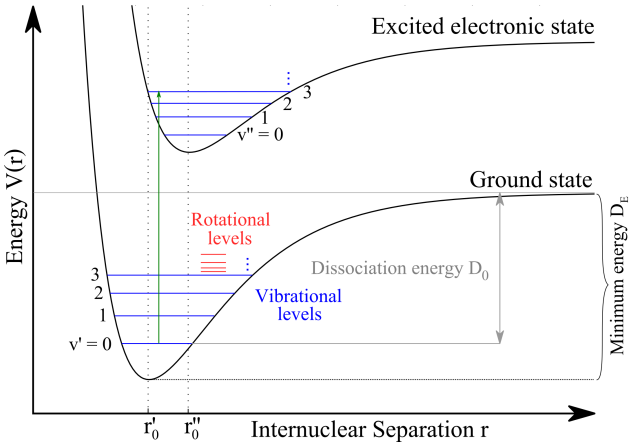


Fig. 2.8: *Illustration of the Franck-Condon principle. The black curves indicate the ground state and an excited electronic state of a diatomic molecule. The vibrational levels are displayed in blue, the rotational levels in red. The green arrow represents a transition (absorption) from the ground state to the 3rd vibrational level of the excited electronic state.*

excitation of electrons means, that the electron density between both nuclei decreases. As a consequence of the larger distance between electron and nuclei, the internuclear distance increases as well (from r'_o to r''_o). The potential flattens and the Dissociation energy D_E is smaller. The electronic transition from the ground state to an excited state is caused by *absorption* of a photon (green arrow in Fig. 2.8). The reversed process is producing the *emission* of a photon. In contrast to atoms, molecules have two further excitation schemes which are based on the additional degrees of freedom. They are called vibrational and rotational transition.

- **Rotational transition:** These transitions occur in the sub-mm or microwave spectral range for small energies between $10^{-3}\,\mathrm{eV}$ to $10^{-2}\,\mathrm{eV}$. The rotational energy levels depend on the moment of inertia I of the molecule, as well as the total angular momentum $\vec{J}$:
  $E_{rot} = \frac{\hbar^2}{2I} \cdot J(J+1)$
  The energy levels depend on J and are therefore not equally spaced:
  $\Delta E_{rot} = E_J - E_{J-1} = \frac{\hbar^2 J}{I}$

- **Vibrational transition:** Pure vibrational transitions occur in the infra-red spectral region. The order of energy is around $0.1\,\mathrm{eV}$. Vibrational energy levels can be considered, in an approximation, as those of an harmonic oscillation:
  $E_{vib} = (\nu + 0.5)\hbar\omega$
  With the vibrational quantum number ν as a natural number. In this approximation, the energy levels are equidistant. However, to be precise, the levels decrease with increasing ν.

- **Electronic transition:** In the near infra-red, visible and ultra violet spectral region electronic transitions dominate. The order of energy is around $1\,\mathrm{eV}$. A change in electronic energy might be accompanied by vibrational and rotational transition characteristics as well.

An excitation is often connected to an electronic transition, whose exciting photon had enough energy to *push* the system to an excited state. The vibrational energy level for such a transition does not depend only on the energy:
The *Franck-Condon principle* states, that the distance between the nuclei can not change during an electronic transition, because the time span for this process is much shorter ($\sim 10^{-15}\,\mathrm{s}$) than that for a possible vibration ($\sim 10^{-13}\,\mathrm{s}$). The final vibrational energy level is then the most probable one with the initial nuclei separation r'. This means from a quantum mechanic perspective, that the vibrational wave functions of the ground state and the excited state have the largest overlap of all other vibrational energy levels.

Line formation and line broadening

So far, absorption and transmission was treated as a discrete amount of energy which was given (taken) from (to) the discussed system. The Franck-Condon principle explains, that there might be a possible energy range which ends in the same, most probable vibrational state. Furthermore, *Heisenberg's energy-time uncertainty principle* ($\Delta E \cdot \Delta t \geq \hbar/2$) states that a transition is always connected with a certain energy range due to the finite lifetime $\tau = \Delta t$ of ground and excited energy level.
As a consequence, transitions have a *natural line broadening* which can be observed even

with the highest spectral resolution of the measuring instrument. The spectral intensity $I_{nat}(\omega)$ of this broadening can be approximated as that of an harmonic oscillations to a *Lorentz profile*

$$I_{nat}(\omega) = I_0 \frac{\delta\omega_{nat}/2\pi}{(\omega - \omega_0)^2 + (\delta\omega_{nat}/2)^2}, \tag{2.19}$$

with the natural linewidth $\delta\omega_{nat}$ as full width at half maximum (FWHM). This width depends not on the energy difference between ground and excited state but on the lifetime τ of the excited state.
Another broadening phenomenon is called *Doppler broadening* and it is based on the Brownian motion of molecules in a gas. The mid-frequency of a transition underlies the *Doppler effect* and it changes with the velocity of the particle. The lineshape is Gaussian with the Doppler width $\delta\omega_{dop}$.

$$I_{nat}(\omega) = I_0 \cdot e^{-\delta\omega_{dop}(\omega-\omega_0)^2/2} \tag{2.20}$$

Figure 2.9 shows in the top right corner a schematic comparison of the two profile shapes. In addition, *pressure broadening* (or *collision broadening*) is a Lorentz-shaped line broadening which is caused by collisions of molecules within gases. These collisions reduce the lifetime of the excitation. The number of collisions, and therefore the width of the line, depend on the gas pressure and on the average velocity of the molecules inside the gas.
For the overall line width, one has to consider that the Doppler broadening is not only working on the initial monochromatic intensity but on the Lorentz-shaped natural and pressure broadend line. The result will be a convolution of Lorentz and Gauss curves which results in a *Voigt profile* (see Fig. 2.9). As can be seen from the figure, all broade-

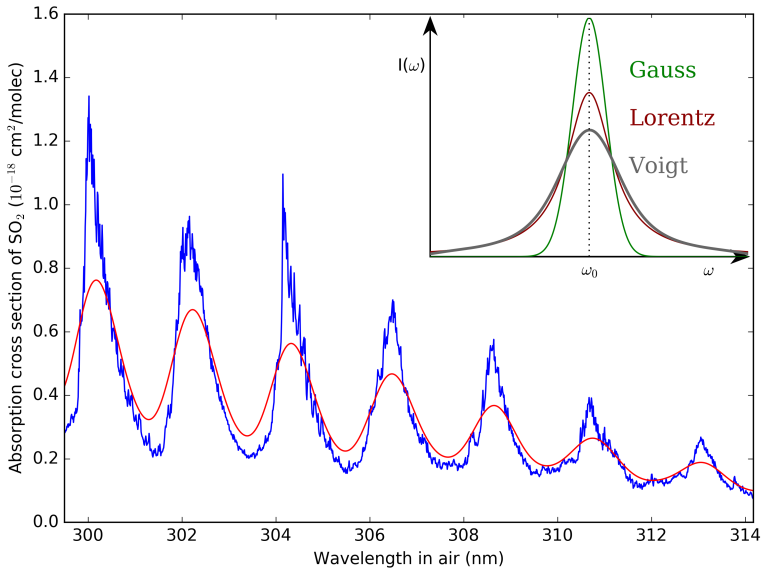


Fig. 2.9: *Blue curve: High resolution absorption cross section of sulfur dioxide* SO_2 *(*VANDAELE *et al., 1994) and the with a gaussian (FWHM* $= 0.8\,\mathrm{nm}$*) convoluted cross section (red). The top right plot shows a schematic represenation of different line broadening shapes.*

ning mechanisms decrease the peak intensity of the transition. Furthermore, the number of wavelengths which contribute to one transition increases as well. A comparison with a measured absorption cross section[19] shows further consequences. The blue curve depicts, for one electronic transition, different vibrational transitions. Not fully resolved rotational transitions lead to noise like features on top of each peak. Cross sections like the depicted one are used in the DOAS analysis (see section 2.5) to retrieve differential slant column densities (dSCD). As another smearing effect, the finite width of the entrance slit of a spectrometer leads to a further broadening of each line. The measured intensity spectrum $I_{meas}(\lambda)$ is therefore a convolution of the instruments slit function $SF(\lambda)$ and the true intensity.

$$I_{meas}(\lambda) = (I_{true} * SF)(\lambda) \equiv \int_{-\infty}^{\infty} I_{true}(\lambda - \lambda')SF(\lambda')d\lambda' \tag{2.21}$$

An example of a convoluted cross section is shown in red. Only the smoothed vibrational transitions remain.

2.3.2.2 Scattering

In contrast to interaction phenomena which are accompanied by a change in energy (absorption, emission) only, *scattering* is defined primarily as a change in direction of light caused by another object. Within the atmosphere, these objects are atoms, molecules and smaller particles which scatter incoming radiation. Supplementary to absorption and emission, measured intensities differ a lot by assuming additional scattering processes, as the light might have been scattered into or out of the incident light beam to a detector. In addition to the characteristic of changing directions, it can be associated with a change in energy as well. Therefore, scattering can be separated in an *elastic* and an *inelastic* part.
Many phenomena in the atmosphere can be explained by scattering processes such as the blue colour of the sky, the white colour of clouds and rainbows.

Elastic scattering

The easiest possible scattering case is the elastic scattering of particles whose size (diameter d) is sufficiently smaller than the wavelength (or $d \ll 1$). It is called *Rayleigh scattering* (RAYLEIGH, 1899). This interaction process is elastic because it does not involve a change in energy of the scattered photon.
When the particle is much smaller than the wavelength, each part of it is influenced by the radiation simultaneously. Considering light as electromagnetic radiation (Eq. 2.9),

[19] An absorption cross section is measured in a laboratory by using Lambert-Beer's law with the help of a light source on a gas filled cell.

this means, that the incoming external oscillating electric field $\vec{E}_{ext}$ polarizes the particle by separating the positive and negative charge carriers. The particle becomes an electric dipole with the induced dipole moment $\vec{p}$.

$$\vec{p} = \alpha \vec{E}_{ext} exp(i\omega t) \tag{2.22}$$

Here, α is the polarizability, which describes the ability of a physical system to form dipoles. $\vec{p}$ depends on the frequency $\omega = 2\pi\nu$ and on the size and composition of the particle as the polarizability does. Generally, α is a tensor because the polarizability is anisotropic. For simplicity, the isotropic case is considered here and α is a scalar.
As the dipole is oscillating, it produces in turn electromagnetic radiation of the same wavelength, which is the (outgoing) elastic scattered radiation.
With the above explanation, a reason for elastic scattering was found but there is a need to know the direction of scattering because measured intensity might be enhanced or attenuated. In general, for the description of the angular distribution of scattered light a so called *phase function* $P(\Theta)$ is used. This function can be understood as a probability density for scattering into a certain direction with the angle Θ between incident and outgoing photon. It is defined as normalized to 1:

$$\frac{1}{4\pi} \int_0^{\pi} \int_0^{2\pi} P(\Theta) \sin\Theta d\phi d\Theta = 1 \tag{2.23}$$

and is in case of Rayleigh scattering

$$P_{Ray}(\Theta) = \frac{3}{4} \left(1 + \cos^2\Theta\right). \tag{2.24}$$

This equation is valid under the assumption of small particles ($d \ll 1$) with spherical shape and it satisfies the normalization condition 2.23. In Figure 2.10 the Rayleigh phase function for isotropic scattering as well as unpolarized and polarized scattered light is depicted. As can be seen, there is a greater probability for scattering in forward or backward direction than perpendicular to the incident light beam in case of polarized radiation. Equation 2.24 describes the probability of elastic scattering in a certain direction, when a photon hits a small and spherical particle without considering what the probability is that a photon hits this particle. The definition of a cross section is of use here as it directly links probability with characterizing parameters. When N spherical objects are arranged in an area A with each having a cross area of $\sigma_{geom} = \pi r^2$, then the probability P that an incident target beam hits one of these particles is:

$$P = \frac{\sigma_{geom} N}{A} \Leftrightarrow \sigma_{geom} = \frac{PN}{A} \tag{2.25}$$

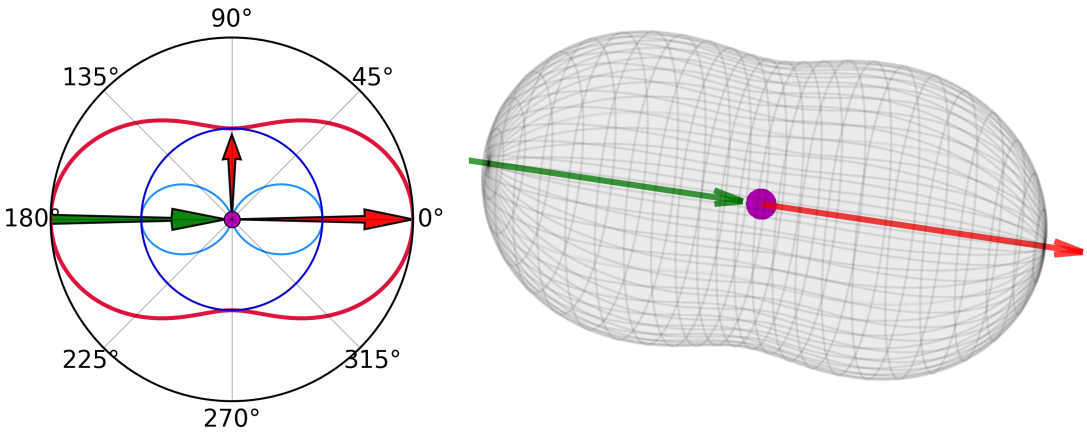


Fig. 2.10: *Schematic representation of the Rayleigh phase function. Green arrows indicate the incident photon, red possible outgoing photons and the magenta dot represents the scattering particle.* ***Left:*** *Polar diagram of non polarised incident light (red), isotropic light (dark blue) and vertical polarised light (light blue).* ***Right:*** *3D plot of Rayleigh's phase function.*

Here, the assumption of particles having a certain cross area is referred to as geometrical cross section σ_{geom}. A cross section has usually the unit cm^2/molecule (the exception is O_4, with cm^5/molecules^2 as it is linked to the O_2 concentration, GREENBLATT et al., 1990). When dealing with spherical objects, the *size ratio* x is another useful quantity because it directly links wavelength λ and particle radius r:

$$x \equiv \frac{2\pi r}{\lambda} \tag{2.26}$$

In Figure 2.11, the relationship between wavelength and particle size for different size ratios is depicted. One important finding is that different particle types might scatter in up to three scattering regimes, depending on the energy of the incoming photon. In the visible spectral range, *Rayleigh* and *Mie* scattering are the most important regimes besides geometric optics for large particle where *Snell's* law ($n_1 sin(\theta_1) = n_2 sin(\theta_2)$) is applied.

For the scattering of light, the before mentioned geometrical approximation is not valid because σ also depends on the energy (and therefore wavelength) and on optical properties of the particle and the surrounding medium. The latter one is introduced as the relative index of refraction $n = N_2/N_1$ with the complex refractive indices N_1 and N_2, as the particles and mediums optical properties respectively. The cross section σ_{Ray} for Rayleigh scattering is now defined as

$$\sigma_{Ray} = \frac{128\pi^5 r^6}{3\lambda^4} \left| \frac{n^2 - 1}{n^2 + 2} \right|^2 \propto \frac{r^6}{\lambda^4}, \tag{2.27}$$

with the important proportionality of $\sigma_{Ray} \propto r^6/\lambda^4$ (RAYLEIGH, 1899). This is an important finding because it explains several phenomena in the atmosphere. The blue sky and

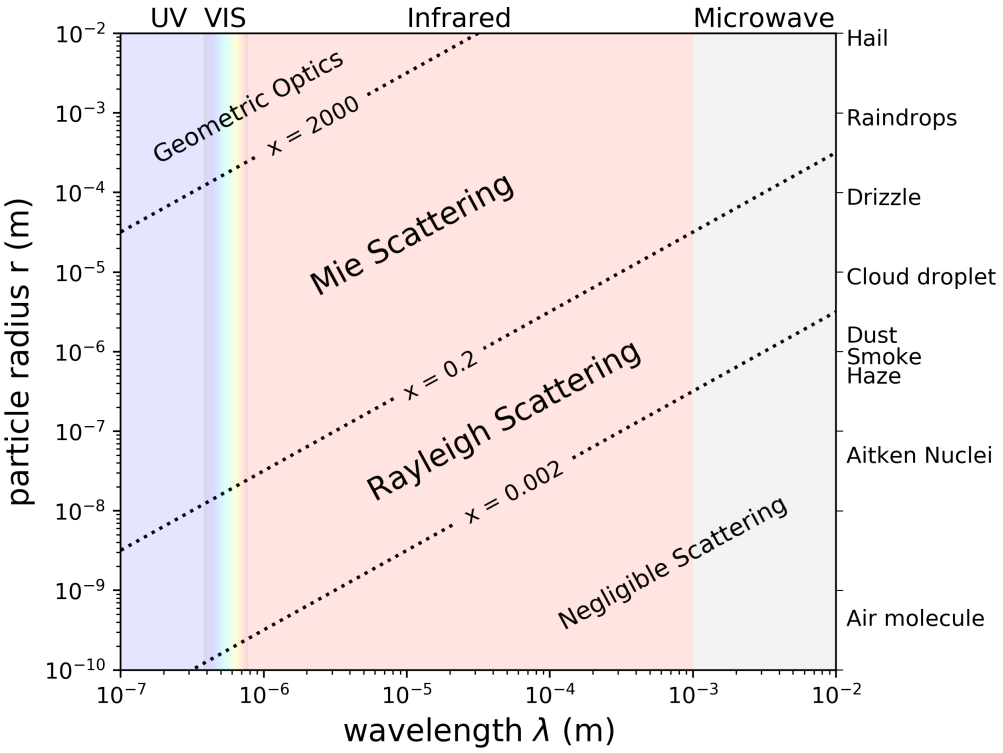


Fig. 2.11: *Wavelength versus particle radius for different size ratio scattering regimes and particle types. Reproduced after* PETTY [2006].

the reddish colour during the sunset are prominent examples for the impact of Rayleigh scattering.

So far, the explanation has focused on the Rayleigh scattering regime for particles smaller than the incident wave's wavelength. For larger spherical scattering objects, Mie theory can be used to describe the wavelength and size dependency in the size distribution domain with $x \gtrsim 1$. In general, Mie theory means a rather complicated series expansion of coefficients by use of Maxwell's equations (see section 2.3.1.1). Doing this, the scattering cross section is

$$\sigma_{Mie} = \frac{2\pi r^2}{x^2} \sum_{j=1}^{\infty} (2j+1)(|a_j|^2 + |b_j|^2) \tag{2.28}$$

with the Mie scattering coefficients a_j and b_j, representing the electric and magnetic multipoles. The leading term ($j = 1$) describes the Rayleigh scattering so that Equation 2.27 could have been found with the help of Mie theory. For larger particles, more terms have to be considered and the wavelength dependency is

$$\sigma_{Mie} \propto \lambda^{-\kappa} \qquad \text{with} \quad \kappa = 0...2. \tag{2.29}$$

Note, that σ_{Mie} is less wavelength dependent than σ_{Ray}. Another important finding can be seen by comparing Figure 2.12 with 2.13. Larger particles tend to scatter light primarily into the forward direction. This is the reason for effects like the increase in brightness of a car's dirty wind shield when driving into the direction of the sun.

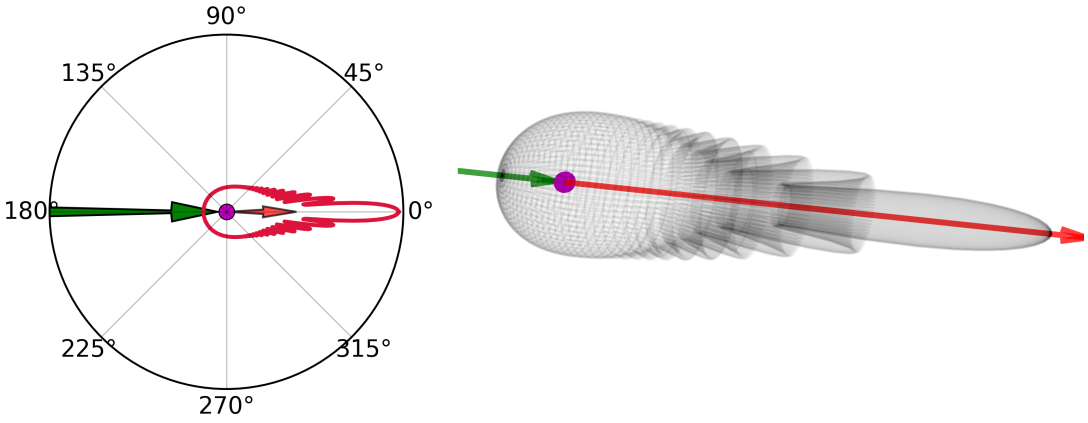


Fig. 2.12: *Schematic representation of the Mie phase function. Green arrows indicate the incident photon, red a possible outgoing photon and the magenta dot represents the scattering particle.* ***Left:*** *Polar diagram of a particle with $d = 3\,\mu\text{m}$ and $n = 1.0 - 0.5i$ (red).* ***Right:*** *3D plot of Mie's phase function.*

In addition to the strong forward peak, several side peaks can be found which make the phase function much more complex than P_{Ray}. This circumstance prevents the usage of Mie phase functions for simple considerations due to the enormous computational effort. To solve this problems, empirically found analytic phase functions are frequently used which include less parameters for the description of several particles. In Figure 2.13 the aerosol phase function after HENYEY and GREENSTEIN [1941] is depicted. The phase function is defined as

$$P_{HG}(\Theta) = \frac{1 - g^2}{(1 - 2g\cos(\Theta) + g^2)^{1.5}} \tag{2.30}$$

with asymmetry factor g

$$g = 0.5 \int_0^{2\pi} P(\Theta)\cos(\Theta)\mathrm{d}\cos(\Theta). \tag{2.31}$$

which is a measure for the ratio between forward and backward scattering. In Figure 2.10, two cigar-like probability shapes for the asymmetry factors 0.3 (blue) and 0.6 (red) are shown. There is a strong tendency for forward scattering rather than backward when g is higher. Isotropic scattering means $g = 0$ (green).

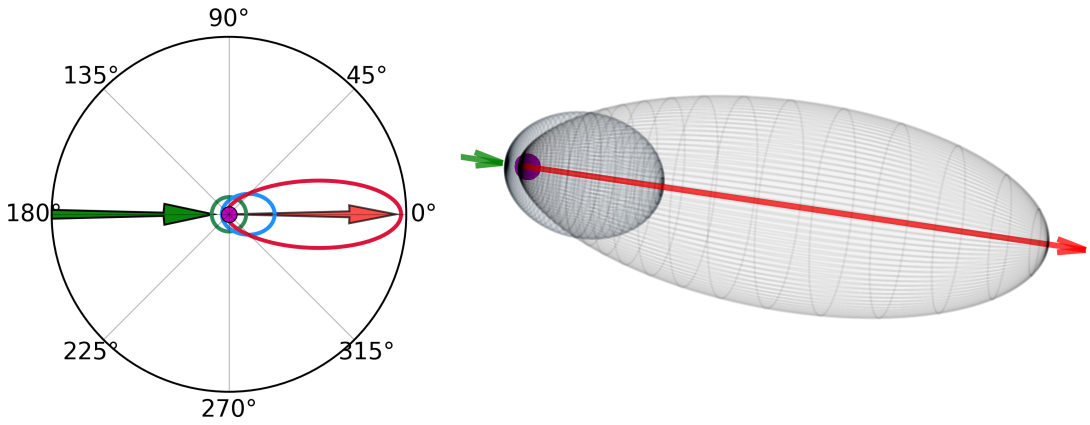


Fig. 2.13: *Schematic representation of the Henyey-Greenstein phase function. Green arrows indicate the incident photon, red a possible outgoing photon and the magenta dot represents the scattering particle.* ***Left:*** *Polar diagram of the phase function with $g = 0$ (green), $g = 0.3$ (blue) and $g = 0.6$ (red).* ***Right:*** *3D plot of Henyey-Greenstein's phase function for $g = 0.$ (small cigar) and $g = 0.6.$ (big cigar)*

Unfortunately, P_{HG} is not capable of reproducing measured aerosol phase functions well into the forward direction and fails completely for the backward peak (see also discussion in Sec. 3.7.1). This is the reason why often two Henyey-Greenstein phase functions are used for forward and backward scattering via $P_{2HG}(\Theta, \beta g_1, g_2) = \beta P_{HG_1}(\Theta, g_1) + (1 - \beta) P_{HG_2}(\Theta, g_2)$ respectively (see e.g. TOUBLANC [1996]).

Inelastic scattering

In addition to elastic processes explained in the previous section, there are also scattering processes which are accompanied with a change in energy. One important example for inelastic scattering is called *Raman scattering*. In contrast to elastic scattering, the emitted photon might have an energy higher ($h\nu^+$) or lower ($h\nu^-$) than the initially incident photon's energy $h\nu_0$. To be more precise, the energy of the emitted photon is the difference between an excited state and a relaxed state of the molecule ($h\Delta\nu = h(\nu_0 - \nu^{\pm})$). The new state might be a vibrational or rotational state of the atom's or molecule's ground state. The first case is called vibrational raman scattering (VRS) and it is much weaker than the rotational effect (RRS) in the atmosphere (see PLATT and STUTZ [2008]). Since there are several quantized energy states possible, the matching cross section shows many lines at different energies, the so called *Stokes* lines. If the energy of the emitted photon is higher than the initial energy, the associated lines are called *Anti-Stokes* lines. The latter effect is weaker because the atom or molecule must already have been in an excited state for emitting a photon with higher energy than the initial one.
As a consequence of inelastic scattering in the atmosphere, Fraunhofer lines were found to be not as deep as expected by GRAINGER and RING [1962]. Nowadays, this filling-in is

referred to as *Ring effect*. BRINKMANN [1968] and KATTAWAR et al. [1981] showed, that this effect is due to wavelength shifts by RRS or VRS. Since there are more photons at wavelengths close to a Fraunhofer line than within this line, it is more likely that the energy is changed in a way that the line is filled-in rather than filled-out. This effect can also be found for strong absorbers like ozone. The Ring effect is an important phenomenon which is accounted for with an additional pseudo cross section within the DOAS fit (see Sec. 2.5).

Multiple scattering

Imagine an incident photon at a certain wavelength measured by an instrument. How should we know where it came from? Was it transmitted through the atmosphere or scattered one metre away from the instrument? Since there does not exist a quantity for historical interaction processes of each photon, it is often assumed, that every photon had only one interaction before it was measured. That this is not true can be observed by taking a look at clouds. The lack in blue colour indicates, that Rayleigh scattering of air molecules was somehow obscured. This is due to *multiple scattering* within optical thick layers like clouds leading to a strongly increased amount of Mie scattering. A photon might be scattered a hundred times within clouds before it leaves into the direction of the instrument where it was measured. Multiple scattering is an important phenomenon not only in clouds, but also in aerosol layers with weakly absorbing aerosols. Retrieval artefacts are possible when multiple scattering leads to stronger signals of trace gases when e.g. an aerosol enhances multiple scattering and therefore the probability for scattering at a trace gas molecule. Because of the importance for radiative transfer, multiple scatter processes have to be considered when dealing with radiative transfer models like SCIATRAN (see Sec. 2.4.1).

2.4 Radiative transfer in Earth's atmosphere

In the previous sections, the sun was introduced as the source of light on Earth. Interaction phenomena of sunlight in Earth's atmosphere were explained but how these processes can be accounted for within analytical expressions is the content of this section. Radiative transfer through the atmosphere is a rather complicated matter because several species, different light paths, vertical (and horizontal) inhomogeneities within a spherical atmosphere create the need for sophisticated equations and solution schemes. Here, a short introduction is given but a detailed discussion is out of scope for this thesis. The reader is referred to literature like CHANDRASEKHAR [1960], LIOU [2002], PETTY [2006] or JACOBSON [2005] for more information about this topic.[20]

[20] This section is based on PETTY [2006].

For a simple consideration, a homogeneous atmosphere with one absorber only can be assumed. An initial intensity I_0 will be attenuated exponentially on its way through the absorber so that the final intensity after travelling the distance s is

$$I = I_0 \cdot exp(-\beta_a s), \tag{2.32}$$

where β_a is called *absorption coefficient* with the unit (1/m). This equation is valid when the attenuation of intensity is due to absorption only. In section 2.3 it was explained, that a photon might also be scattered out of the light path. To account for this, the sum of absorption and scattering coefficients is defined as extinction coefficient

$$\beta_e = \beta_a + \beta_s. \tag{2.33}$$

Another important quantity is the *single scattering albedo* ω (SSA), defined as

$$\omega = \frac{\beta_s}{\beta_e} = \frac{\beta_s}{\beta_a + \beta_s}. \tag{2.34}$$

For a more precise description of light passing the atmosphere than Equation 2.32, we introduce an infinitesimal layer of thickness ds and the change in intensity within this layer as

$$dI_{ext} = -I(s)\beta_e(s)ds. \tag{2.35}$$

Note that the intensity depends on the wavelength λ but this is omitted here for simplicity. An integration between s_1 and s_2 leads to the *Beer-Lambert's law*

$$I(s_2) = I(s_1)exp\left(-\int_{s_1}^{s_2} \beta_e(s)ds\right). \tag{2.36}$$

Another crucial quantity is the so called *optical depth* τ

$$\tau(s_1, s_2) \equiv -\int_{s_1}^{s_2} \beta_e(s)ds. \tag{2.37}$$

Now, all important relationships are derived for a more accurate description of radiative transfer. Consider again an infinitesimally small layer with thickness ds in the atmosphere. During the passage of radiation through this layer, light might be absorbed, scattered or emitted by atoms or molecules. Whereas absorption and emission are straightforward, scattering needs to be accounted for in two different ways. Light might be scattered into the thin layer from adjacent layers increasing the intensity. Scattering can also lead to a decrease in radiation, when photons are scattered in a direction $\mathbf{\Omega}'$ which is not the

direction of propagation $\mathbf{\Omega}$[21]. The overall change in intensity dI is the sum of the changes due to extinction dI_{ext} (Eq. 2.35), emission dI_{emi} and scattering dI_{sca}[22].

$$dI = dI_{ext} + dI_{emi} + dI_{sca} \tag{2.38}$$

$$= -\beta_e I ds + \beta_a B(T) ds + \frac{\beta_s}{4\pi} \int_{4\pi} P(\mathbf{\Omega}, \mathbf{\Omega}') I(\mathbf{\Omega}') d\Omega' ds. \tag{2.39}$$

The emission term dI_{emi} is here defined to equalize an absorbing atmosphere only, by emitting the totally absorbed quantity $\beta_a ds$ as a black body obeying Planck's law B (Eq. 2.1) in the infra-red spectral region[23]. In contrast to the change due to extinction dI_{ext}, the last term describes scattering of photons from all directions $\mathbf{\Omega}'$ into ds which leads to an increase in intensity. $I(\mathbf{\Omega}')$ is multiplied by the phase function $P(\mathbf{\Omega}, \mathbf{\Omega}')$ and integrated over solid angles 4π.
Dividing both sides of Equation 2.39 by $d\tau = -\beta_e ds$ and introducing the single scattering albedo ω (SSA), Equation 2.39 can be reformulated as:

$$\frac{dI(\mathbf{\Omega})}{d\tau} = I(\mathbf{\Omega}) - (1-\omega)B - \frac{\omega}{4\pi} \int_{4\pi} P(\mathbf{\Omega}, \mathbf{\Omega}') I(\mathbf{\Omega}') d\Omega' \tag{2.40}$$

This is a general and often used formulation for the radiative transfer through the atmosphere but further remarks should be noted. In the equation above, the intensity is a scalar quantity but polarization would introduce a vector defined as Equation 2.14 with a 4 x 4 phase matrix $\mathbf{P}(\mathbf{\Omega}, \mathbf{\Omega}')$. In addition, for many problems it is assumed that a plane parallel atmosphere is sufficient for the description of all important processes. Here, an explanation of how to solve the before mentioned problems is clearly out of scope, but this short illustration of radiative transfer enables the introduction of SCIATRAN in the next section.

2.4.1 The radiative transfer model SCIATRAN

IUP Bremen's in-house radiative transfer model SCIATRAN is a frequently used, well proven radiative transfer model (RTM) which is capable of calculating various atmospheric parameters (ROZANOV et al., 2014). It provides products in a full spherical atmosphere including polarisation, multiple scattering and rotational and vibrational raman scattering. SCIATRAN covers the spectral region between 175.44 nm and 44 μm but not all products are available throughout the full spectral range. Some of its final products are intensities (Stokes vectors), weighting functions, air mass factors (AMF), box air mass factors (BAMF) and inversion retrieval products (e.g. trace gas/aerosol profiles).

[21] The solid angle vector $\mathbf{\Omega}$ summarizes a set of azimuthal ϕ and polar angles θ and is a representation of specific directions in space. An infinitesimal solid angle in spherical coordinates is defined as $d\Omega = \sin\theta d\theta d\phi$.
[22] The dependency of the changes of intensity dI on s is suppressed from now on.
[23] This term can be neglected in the UV/VIS spectral region.

SCIATRAN solves Equation 2.40 by applying boundary conditions as constraints which enable the numerical solution of the radiative transfer. The upper boundary condition utilizes the assumption of a monodirectional unpolarized light beam[24]. The lower boundary condition is assumed to obey a bidirectional reflection surface[25]. The radiative transfer equation is solved employing the discrete-ordinates method and a source function integration technique. The reader is referred to ROZANOV et al. [2014] for further details.
In this study, SCIATRAN is used for the calculation of differential slant column densities (Ch. III), air mass factors (Ch. III), box air mass factors and aerosol profiles (Ch. IV).

2.5 The DOAS approach

Differential optical absorption spectroscopy (DOAS) is a frequently used and well-known technique for the analysis of trace gases from measurements of direct or scattered light. First DOAS applications reach back to NOXON [1975] and PLATT and PERNER [1980] but essential improvements on measurement and analysis were made in the years between the early 80s and late 90s (e.g. SOLOMON et al. [1987], RICHTER [1997]). Whereas the technique was first used for ground-based measurements, satellite platforms became more and more important with early publications by BOVENSMANN et al. [1999] and BURROWS et al. [1999].
Since the technique is generally based on Lambert-Beer's law (see eq. 2.36), it is applicable to active or passive measurement of artificial, solar, or lunar light. For introductory literature, the reader is referred to PLATT and STUTZ [2008] or BURROWS et al. [2011]. As the general DOAS equation (Eq. 2.48) is a simplification of more accurate analytic formulations, the reader is also encouraged to read MARQUARD et al. [2000], ROZANOV and ROZANOV [2010] or PUĶĪTE et al. [2010] for a discussion of limitations of DOAS.

2.5.1 Derivation of the DOAS equation[26]

In this section, a derivation of the DOAS equation is explained from a more general starting point with a discussion on all approximations which are needed for the derivation. Note that this discussion is similar to the derivation of the minimization problem of the aerosol retrieval in Section 4.2.1.
To make things easier, consider the functional Taylor series expansion of the logarithm of

[24] With the cosine of solar zenith angle μ_0 and the azimuth angle ϕ_0 the upper boundary condition is $I(\mu,\phi) = F_0\delta(\mu-\mu_0)\delta(\phi-\phi_0)I$. F_0 is a constant factor here.
[25] The lower boundary condition is then: $I(\tau_0,\mathbf{\Omega}) = 1/\pi \int d\phi' \int d\mu' \mu R(\mathbf{\Omega},\mathbf{\Omega}')I(\tau_0,\Omega') + \epsilon B(T)I$ with the surface emissivity ϵ and a bi-directional reflectance function $R(\mathbf{\Omega},\mathbf{\Omega}')$.
[26] This section is based on ROZANOV and ROZANOV [2010].

the sun normalized intensity $\ln I$[27] around a linearisation point $\overline{\beta}_a(z)$ for an atmosphere with one absorber assuming that its absorption coefficient profile is $\beta_a(z)$:

$$\ln I(\lambda, \beta_a(z)) = \ln \overline{I}(\lambda, \overline{\beta}_a(z)) + \int_0^H \frac{\delta \ln I(\lambda, \beta_a(z))}{\delta \beta_a(z)} \delta\beta_a(z) dz + \epsilon_{lin}(\lambda) \tag{2.41}$$

This equation can be understood as follows. The first term describes the intensity at a linearisation point which is the quantity which would have been measured if no variations of intensity due to changes in the absorber were present. The second term describes the linear contribution to variation of the logarithm of the sun normalized intensity caused by the variation of the absorber concentration from the surface up to the altitude H. $\delta\beta_a(z)$ is the perturbed vertical absorption coefficient profile which is defined as the difference between absorption coefficient profile $\beta_a(z)$ and the profile at the linearisation point $\overline{\beta}_a(z)$ ($\delta\beta_a(z) = \beta_a(z) - \overline{\beta}_a(z)$). The first multiplier of the integrand is the functional derivative of $\ln I(\lambda, \beta_a(z))$ with respect to absorption coefficient $\beta_a(z)$ calculated at the linearisation point $\overline{\beta}_a(z)$.

$$W(\lambda, z) = \left. \frac{\delta \ln I(\lambda, \beta_a(z))}{\delta\, \beta_a(z)} \right|_{\overline{\beta}_a} \tag{2.42}$$

The quantity $W(\lambda, z)$ is also called *weighting function*. In one sentence, the second term describes the variation of intensity $\ln I(\lambda, \beta_a(z))$ caused by a change in the absorber's absorption coefficient profile $\beta_a(z)$ through all layers within the atmosphere. Since the relationship between intensity and absorber does not necessarily has to be a linear one, $\epsilon_{lin}(\lambda)$ describes the error due to the assumption of linearity. Equation 2.41 is a rather complex formulation but further simplifications can be made. Consider a non-absorbing atmosphere by replacing $\delta\beta_a(z) = 0 - \overline{\beta}_a(z)$ and substituting Equation 2.42 into 2.41, we have

$$\ln I(\lambda, 0) = \ln \overline{I}(\lambda, \overline{\beta}_a(z)) - \int_0^H W(\lambda, z) \overline{\beta}_a(z) dz + \epsilon_{lin}(\lambda). \tag{2.43}$$

Taking further into account that $\beta_a(z) = \sigma(\lambda)\, n(z)$[28] (see also Eq. 2.25), the following equation provides a linear relationship between the logarithm of the sun normalized intensity and an absorbers number density profile $n(z)$:

$$\ln I(\lambda, n(z)) = \ln I(\lambda, 0) + \int_0^H W(\lambda, z) \sigma(\lambda) n(z) dz + \epsilon_{lin}(\lambda) \tag{2.44}$$

[27] As a simplification, the normalization with a solar spectrum I_0 ($\ln I/I_0$) is suppressed.

[28] Here, the absorption cross section $\sigma(\lambda)$ is assumed to be height independent. Furthermore, it can be seen, that $\beta_a(z)$ depends on the wavelength. The dependency will be suppressed in this section.

The linearisation error can usually be neglected as it is small around the linearisation point when no strong absorption species are present and the scattering processes are sufficiently weak. Moreover, it was already explained that atom or molecule absorption happens on a high frequency scale, whereas Rayleigh and Mie scattering can be considered as smoothly varying with wavelength (see Sec. 2.3.2.1 and 2.3.2.2). Under these assumptions, $\ln I(\lambda, 0))$ can be approximated by a low order polynomial $\sum_{j=1}^{m} = a_j \lambda^j$ with degree m. The final Taylor series expansion is given by

$$\ln I(\lambda, n(z)) = \int_0^H W(\lambda, z)\sigma(\lambda)n(z)dz + \sum_{j=1}^{m} a_j \lambda^j. \tag{2.45}$$

Here, the general definition of the so called *slant column density* $S(\lambda)$[29] is introduced as follows

$$S(\lambda, n(z)) = -\int_0^H W(\lambda, z)n(z)dz, \tag{2.46}$$

which leads to

$$\ln I(\lambda, n(z)) = \ln I(\lambda, 0) - S(\lambda, n(z))\sigma(\lambda) + \sum_{j=1}^{m} a_j \lambda^j. \tag{2.47}$$

Note, that by applying an exponent to the above equation it would yield in a form of Lambert-Beer's law (see Eq.2.36).
So far, the whole consideration was done for one absorber only. Fortunately, the expansion to several absorbers is a simple task. Since further absorber would introduce additional absorption cross sections σ_i, a summation over all $S_i(\lambda, n_i(z))\sigma_i(\lambda)$ in Equation 2.47 satisfies the need for the description of the species contained in a typical measurement.

$$\ln\left(\frac{I(\lambda)}{I_0(\lambda)}\right) = -\sum_{i=1}^{l} S_i(\lambda)\sigma_i(\lambda) + \sum_{j=1}^{m} a_j \lambda^j \tag{2.48}$$

Here, the solar spectrum I_0 was explicitly written as it will be referred to within the next subsection. This equation is valid for all trace gas absorbers l and it is usually referred to as *DOAS equation*.

[29] In general, the integration along a vertical axis is referred to as vertical column density rather than slant column density. Since no specific geometry was introduced so far, I will use the terminology slant column density.

2.5.2 Limitations and enhancements of the DOAS approach

On the way to Equation 2.48, several assumptions and approximations were made and are summarized below:

- Local linearity of $\ln I$ was assumed, regarding the variation of absorbers and scattering and reflection processes.
- Absorber cross sections have to include high frequency variations.
- Rayleigh and Mie scattering have to vary slowly with wavelength, so that a polynomial can be used as replacement.
- Surface reflectance has to show broadband features which can be replaced by the polynomial as well.
- Absorber cross sections are independent of altitude.
- Normal instruments have a certain field of view (FOV), so that the DOAS equation needs to be integrated over the full FOV for an accurate description.
- Geometry aspects are not considered so far. Scattering phase functions and slant light paths need further assumptions.
- $I_0(\lambda)$ was considered as scattering term only but might be contaminated by absorption features as well.

Below, quantities will be defined which are also important for the analysis and discussion within this thesis and which explain how one can easily cope with some of the above mentioned bullet points. Other assumptions have a smaller impact on the presented results and can be neglected.

Practical considerations

Equation 2.48 was derived with the assumption, that $I_0(\lambda)$ was measured without absorption. This is easily possible when dealing with satellite measurements, because a solar irradiance spectrum can be used as initial intensity. For ground-based measurements, a measurement of $I_0(\lambda)$ without consideration of absorbing species is not possible. A better understanding can be achieved by looking at Figure 2.14 showing a typical Multi Axis DOAS (MAX-DOAS) measurement geometry.

Every intensity spectrum, measured with a ground based instrument, shows Fraunhofer lines as well as absorption structures of species in the light path s through Earth's atmosphere. From a practical viewpoint, $I_0(\lambda)$ (also called reference spectrum) is measured in a geometry which minimizes the length of s so that absorption can be considered as relatively small. This is often done by using a spectrum in zenith direction $I_{0_{zen}}(\lambda)$, which still

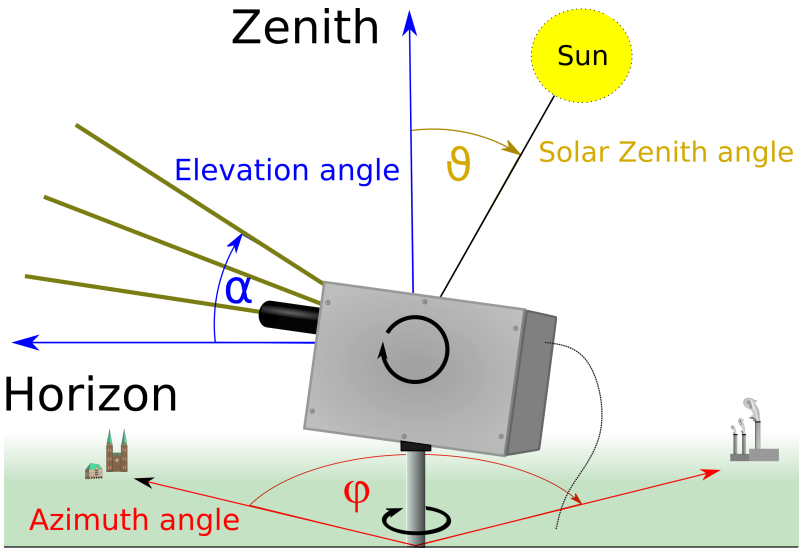


Fig. 2.14: *Schematic representation of a typical MAX-DOAS measurement geometry (*BÖSCH et al., 2018*).*

contains a certain amount of trace gas absorption which has to be considered. To do so, S is redefined as *differential slant column density* ΔS:

$$\Delta S(\lambda, \mathbf{\Omega}) = S(\lambda, \mathbf{\Omega}) - S(\lambda, \mathbf{\Omega_0}) = \int_{s_1}^{s_2} n(s)ds - \int_{s_1'}^{s_2'} n(s)ds' \tag{2.49}$$

$$= -\int_0^H [W(\lambda, \mathbf{\Omega}, z) - W(\lambda, \mathbf{\Omega_0}, z)]n(z)dz \tag{2.50}$$

Here, $\mathbf{\Omega_0}$ is a geometry vector which describes a set of variables representing the zenith (or reference) direction. $\mathbf{\Omega}$ summarizes a certain viewing geometry. As Figure 2.14 demonstrates, a geometry vector can be described by an elevation angle α (also called line of sight, LOS), the solar zenith angle ϑ (SZA) and an azimuth angle φ. Note, that for the description of a scattering angle for e.g. RTM calculations, a simple telescope viewing azimuth angle $\varphi = \varphi_{vaa}$ (VAA) is not enough. The relative azimuth angle φ_{raa} (RAA) is therefore defined as the difference between instruments azimuth viewing angle and the solar azimuth angle φ_{saa} measured clockwise from the south:

$$\varphi_{raa} = \varphi_{vaa} - \varphi_{saa} \tag{2.51}$$

Similar to Equation 2.50, the optical thickness τ (see also Eq.2.37) can be formulated as *differential slant optical thickness*:

$$\Delta\tau(\lambda, \mathbf{\Omega}) = \tau(\lambda, \mathbf{\Omega}) - \tau(\lambda, \mathbf{\Omega_0}) = \sigma(\lambda)\Delta S(\lambda, \mathbf{\Omega}) \tag{2.52}$$

Effective light path, Air mass factor and vertical columns

As a summary of the considerations above: A MAX-DOAS measurement is done by measuring an intensity spectrum $I(\lambda, \mathbf{\Omega})$ and a reference spectrum $I_0(\lambda, \mathbf{\Omega}_0)$ and by applying the DOAS fit (cf. Sec. 3.4) to the logarithm of the ratio of both spectra to retrieve a differential slant column density $\Delta S(\lambda, \mathbf{\Omega})$. Unfortunately, ΔS can not be used to make a statement about the trace gas amount in a certain location or direction. Since Equation 2.48 includes only one integration, the matching light path should be considered as an *effective light path* rather than a specific one. Neither the actual length H (or L for slant light paths) is known, nor the height depending trace gas concentration $n(z)$ (or light path depending $n(s)$).
The *vertical column density* V is defined as the vertically integrated absorber concentration from surface to top of the atmosphere (TOA):

$$V = \int_0^{TOA} n(z)dz \tag{2.53}$$

Since TOA is more or less known (or defined), vertical columns are a useful quantity for deriving wavelength independent integrated number concentrations along the vertical dimension. The quantity relating S to V is the so called *air mass factor* M (or AMF, see e.g. SOLOMON et al., 1987):

$$M(\lambda, \mathbf{\Omega}) = \frac{S(\lambda, \mathbf{\Omega})}{V} \tag{2.54}$$

AMFs are usually calculated with RTM's and need further assumptions on the specific trace gas and aerosol profiles, geometry, surface spectral reflectance and wavelength. For the conversion of differential slant column densities into vertical column densities, the specific reference AMF is needed so that $\Delta S/\Delta M = (S - S_0)/(M - M_0) = V$. The difference between initial and reference AMF is similar to ΔS called *differential air mass factor* ΔM (dAMF). Without RTM calculations, AMF conversion factors can be used by applying a geometric approximation (see HÖNNINGER et al., 2004). Under the assumption that the trace gas of interest is located below the last scattering point, the geometric AMF can be found as follows:

$$M_{geom}(\alpha) = \frac{S(\alpha)}{V} = \frac{1}{\sin(\alpha)}, \tag{2.55}$$

with elevation angle α. Another important concept is the so called *box air mass factor* $\mathscr{M}_z$ (BAMF, also block AMF) which can be understood as the altitude depending air mass

factor for layer z:

$$\mathscr{M}_z(\lambda, \mathbf{\Omega}) = \frac{S_z(\lambda, \mathbf{\Omega})}{V_z} \tag{2.56}$$

A similar definition as above is found for differential box air mass factors $\Delta\mathscr{M}_z$ (dBAMF) introducing the sensitivity of the partial differential slant column density ΔS_z to the partial vertical column density V_z in a specific layer z. Again, SCIATRAN is used for the calculation of dBAMF within this thesis.
The relation between M and $\mathscr{M}_z$ is found by summing up the individual partial slant column densities $S_z(\lambda, \mathbf{\Omega})$ and dividing by the total vertical column V (WAGNER et al., 2007):

$$M(\lambda, \mathbf{\Omega}) = \frac{1}{V} \sum_{z=0}^{TOA} \mathscr{M}_z V_z = \frac{1}{V} \sum_{z=0}^{TOA} S_z(\lambda, \mathbf{\Omega}) \tag{2.57}$$

By combining the definition of ΔS (Eq. 2.46), V (Eq. 2.53) with Equation 2.57, a link between weighting functions and dBAMF can be found:

$$\begin{aligned} \Delta S(\lambda, \mathbf{\Omega}) &= -\int_0^H [W(\lambda, \mathbf{\Omega}, z) - W(\lambda, \mathbf{\Omega_0}, z)] n(z) dz = \\ V \Delta M(\lambda, \mathbf{\Omega}) &= \sum_{z=0}^{TOA} \left[(\mathscr{M}_z(\lambda, \mathbf{\Omega}) - \mathscr{M}_z(\lambda, \mathbf{\Omega_0})) \int_{z_i}^{z_{i+1}} n_z(z') dz' \right] \end{aligned} \tag{2.58}$$

When infinitesimally small differences between layer z_i and z_{i+1} are considered, the sum in the second line is transformed into an integration and $\Delta\mathscr{M}_z(\lambda, \mathbf{\Omega}, z) = -\Delta W(\lambda, \mathbf{\Omega}, z)$. The definitions of weighting function and block air mass factor are important for the discussions in this thesis as they provide the sensitivity of measurements to absorber concentrations which can be used for the retrieval of vertical profiles (see Chapter IV).

2.6 Introduction into atmospheric chemistry

The knowledge of chemical processes in the atmosphere is crucial for the interpretation of trace gas measurements. The gaseous composition of Earth's atmosphere was shortly explained in Section 2.2.1. However, this explanation was focussed on a natural mean atmospheric structure. Deviations in the temporal and spatial distribution of individual components from this mean are common, meaning that either transport processes have happened or that sources and sinks have changed the specific amount of atmospheric constituents. In this thesis, only ground-based measurements are discussed so that the following introduction focusses on the tropospheric chemistry of important trace gases

only. This section highlights the most important sinks and sources as well as the chemical reactions of nitrogen dioxide (NO_2) and formaldehyde ($HCHO$). See e.g. JACOB [1999]; SEINFELD and PANDIS [2006]; HOLLOWAY and WAYNE [2015] for further details on this topic.

Air pollution

Air pollution can be defined as *a situation in which substances that result from anthropogenic activities are present at concentrations sufficiently high above their normal ambient levels to produce a measurable effect on humans, animals, vegetation, or materials* (def. by SEINFELD and PANDIS, 2006). Additionally, it should be noted that *natural air pollution* exists (e.g. emissions due to volcanic eruptions, forest fires) and might have a large impact on Earth's climate and the tropospheric trace gas composition.
In both cases, emitted substances can be separated to *primary* and *secondary* pollutants. The first case classifies pollutants which are directly emitted into the atmosphere; no further conversion was needed. Secondary pollutants describe substances which were created by reaction or interaction with other atmospheric compounds or electromagnetic radiation. Examples for primary air pollutants are nitrogen oxide (NO), carbon monoxide (CO) or particulate materials (e.g. soot, ash). Secondary pollutants are e.g. nitrogen dioxide (NO_2), ground level ozone (O_3) or atmospheric acids (such as nitric acid, HNO_3; or sulphuric acid, H_2SO_4). Note that some pollutants can be classified as both, primary and secondary, as they are emitted directly or can be formed from other substances (e.g. NO_2, SO_2).
Sources and sinks of pollutants come in various ways. As examples for natural sources, volcanic eruptions (mostly sulphur dioxide, SO_2; carbon dioxide, CO_2; and aerosols), digestion by animals (e.g. methane emission by cows, CH_4) or vegetation (volatile organic compounds, VOC[30]) can be considered as strong sources. On the other hand, anthropogenic emissions can be attributed for example to combustion processes (traffic and industrial facilities emitting e.g. NO, CO, SO), waste deposition (CH_4) or fertilization of farmlands (HNO_3). In general, the exact emission source is not easy to find from remote sensing data as variable winds, temporally changing emission rates and various sources, might mix up signals.
Sinks of pollutants can be classified into dry or wet deposition as well as chemical sinks. The latter one is the most important sink as any chemical change leads to final products and certain reactants meaning that it can be understood as source of the products while it is also a sink for the reactants. As an example, hydroxyl radicals (OH) act as a sink for methane (CH_4) because of the reaction to methyl radicals (CH_3) and water (H_2O):

[30] VOSs, such as $HCHO$ or benzene (C_6H_6), are organic compounds which evaporate easily at room temperature and are therefore considered as volatile.

$$CH_4 + OH \longrightarrow CH_3 + H_2O. \tag{2.59}$$

This reaction instantaneously forms the peroxy radical (CH_3O_2) by

$$CH_3 + O_2 \xrightarrow{M} CH_3O_2. \tag{2.60}$$

On the other hand, dry deposition means an interaction with the Earth's surface or with substances which have underwent an uptake on it. A simple example is gravitational sedimentation of aerosols. Also microbiological sinks are possible when e.g. CO is transported or emitted close to a surface with a high microbiological activity.
Wet deposition means the removal of water-soluble substances due to precipitation, clouds or fog. A popular example is the formation of acid rain. For example, when nitrogen dioxide (NO_2) reacts with water to form nitric acid (HNO_3):

$$3NO_2 + H_2O \longrightarrow 2HNO_3 + NO. \tag{2.61}$$

Since this thesis mainly discusses nitrogen dioxide (NO_2) and formaldehyde ($HCHO$), the next paragraphs focus on sources, sinks and the most important chemical reactions for understanding the relevant processes of these species.

Nitrogen dioxide

Nitrogen oxides $NO_x = NO + NO_2$ summarize nitric oxide (NO) and nitrogen dioxide (NO_2) and are frequent substances in the atmosphere (cf. Fig. 2.3). While NO is colourless, NO_2 has a brown-reddish tone and a characteristic smell. The tropospheric emission amounts are dominated by fossil fuel combustion and biomass burning (cf. Tab. 2.5). Smaller sources as soil emission, lightning or the NH_3 oxidation account only with percentages smaller than $13.3\,\%$. Due to the possible hazardous effects on human beings of the toxic nitrogen dioxide, nitrogen oxides are routinely measured by air quality in-situ networks in many cities in

trop. NO_x **source**	**Tg N yr**$^{-1}$	**%**
Fossil fuel combustion	21	46.1
Biomass burning	12	26.3
Soils	6	13.2
Lightning	3	6.6
NH_3 oxidation	3	6.6
Aircraft	0.5	1.1
Transport from stratosphere	0.1	0.2

Tab. 2.5: *Global tropospheric sources of Nitrogen oxides. Data taken from* JACOB [1999].

Germany (cf. BLUES network, 3.2). The lifetime of NO and NO_2 in the troposphere is a few seconds and ∼1 day, respectively. Because of these rather short lifetimes, nitrogen oxide abundances are usually largest close to the emission sources in the troposphere but NO_x sometimes can also be transported within reservoir species over long distances.
Within combustion processes, molecular oxygen is thermolyzed and subsequently reacts with molecular nitrogen (N_2) to form NO:

$$O_2 \overset{\text{heat}}{\longleftrightarrow} 2O \tag{2.62}$$

$$O + N_2 \longleftrightarrow NO + N \tag{2.63}$$

$$N + O_2 \longleftrightarrow NO + O \tag{2.64}$$

When the specific temperatures under which the above reactions occur are higher, the equilibria are shifted to the right increasing the amount of NO. The same mechanisms apply for lightning where the temperature is extremely high enabling an efficient production of NO.
After the release of NO, NO_2 is produced from reactions with ozone (O_3). Atomic and molecular oxygen reacts to O_3 with a third body species M, which removes excess energy:

$$O(^3P) + O_2 \xrightarrow{M} O_3 \tag{2.65}$$

M is usually molecular oxygen or nitrogen. This mechanism is valid for both, the troposphere and the stratosphere, and it is the main reason for Earth's ozone layer. $O(^3P)$ is initially produced by photolysis with higher frequency UV photons (or by quenching of excited oxygen atoms). Within the troposphere, cycling between NO and NO_2 occurs on a time scale of minutes (see e.g. JACOB, 1999) by a null cycle[31]:

$$O_3 + NO \longrightarrow NO_2 + O_2 \tag{2.66}$$

$$NO_2 + h\nu \xrightarrow{\lambda \lesssim 420\,\text{nm}} NO + O(^3P) \tag{2.67}$$

Here, $O(^3P)$ reacts via Equation 2.65 and the cycle starts again. The three equations above are known as *Leighton relationship* (LEIGHTON, 1961) and they create an equilibrium between NO, NO_2 and O_3 in the absence of other species. Since this cycle does also depend on photolysis, the equilibrium differs as the amount of sunlight varies. Another possibility for the creation of NO_2 involves peroxy radicals (HO_2 or CH_3O_2). Then, Equation 2.66

[31] A null cycle is a catalytic cycle that simply inter-converts chemical species without leading to net production or removal of any component. Definition from wikipedia, `https://en.wikipedia.org/wiki/Null_cycle`, date: 18.09.2018.

can be replaced by:

$$HO_2 + NO \longrightarrow NO_2 + OH \qquad \text{or} \tag{2.68}$$

$$CH_3O_2 + NO \longrightarrow NO_2 + CH_3O \tag{2.69}$$

The final products are NO_2 and either hydroxyl radicals (OH) or alkoxy radicals (CH_3O). As can be seen, these steps do not involve the removal of O_3 so that more and more ozone aggregates when NO_2 is photolysed again (via Eq. 2.67). This process is known as *summer smog* (also *photochemical smog*) formation.
The discussion of NO_x sinks differs depending on the time. The basic day-time removal of NO_x is the creation of nitric acid (HNO_3) via reaction of NO_2 with OH:

$$OH + NO_2 \xrightarrow{M} HNO_3 \tag{2.70}$$

During night-time, the reaction with O_3 and nitrate (NO_3) leads finally to dinitrogen pentoxide (N_2O_5)

$$NO_2 + O_3 \longrightarrow NO_3 + O_2 \tag{2.71}$$

$$NO_3 + NO_2 \xrightarrow{M} N_2O_5, \tag{2.72}$$

which may then react to HNO_3:

$$N_2O_5 + H_2O \longrightarrow 2\,HNO_3 \tag{2.73}$$

Due to the high solubility in water, HNO_3 can efficiently be removed from the atmosphere by wet-deposition (and contributes to *acid rain*). On the other hand, remaining amounts of NO_3 are quickly photolyzed during the day which starts again the catalytic cycle of nitrogen oxides.
Generally, the diurnal cycle of NO_2 is also dominated by anthropogenic emissions. Especially in the early morning hours and in the afternoon, high NO_2 values can be expected due to the increasing amount of commuter traffic emissions. Even though high near-surface NO_2 concentrations are still an issue nowadays, the reader should note that a decreasing trend in NO_2 over Europe in the last decades was observed (HILBOLL et al., 2013).
Additionally to the tropospheric diurnal cycle, a stratospheric cycle exists with an increasing trend of NO_2 throughout the day and a slower decrease in the night.

Formaldehyde

Formaldehyde ($HCHO$) is a colour-less and flammable gas with an irritating smell. It is the simplest aldehyde meaning that the short-chained structure making it highly water-

soluble. Because of this, HCHO can be washed out from the atmosphere by precipitation. The main sources for HCHO are the oxidation of methane (CH_4) and isoprene (C_5H_8) followed by non-methane volatile organic compounds (NMVOC) to a smaller amount. Direct emissions (biomass burning, fossil fuel combustion) can be considered as minor emission sources (cf. Tab. 2.6). Due to the relatively short lifetime of a few hours (SEINFELD and PANDIS, 2006), HCHO is not expected to be present in larger amounts farer away from possible emission sources. However, the oxidation of methane produces a more or less stable global background HCHO concentration. Enhancements from this background level are due to isoprene emission of plants or anthropogenic pollution close to larger forest or industrial areas, respectively.

trop. HCHO source	**Tg yr^{-1}**	**%**
CH_4 oxidation	960	60
C_5H_8 oxidation	498	30
Anthrop. NMVOC oxidation	116	7
Pyrogenic NMVOC oxidation	50	3
Direct emissions	< 17	< 1

Tab. 2.6: *Sources of formaldehyde. Data taken from* STAVRAKOU et al. [2009].

The oxidation of CH_4 is with OH including the intermediate products, methyl radical (CH_3), methoxy radical (CH_3O), methyl peroxy radical (CH_3O_2) and methylhydroperoxide (CH_3OOH):

$$CH_4 + OH \longrightarrow CH_3 + H_2O \tag{2.74}$$

$$CH_3 + O_2 \xrightarrow{M} CH_3O_2 \tag{2.75}$$

$$CH_3O_2 + NO \longrightarrow CH_3O + NO_2 \tag{2.76}$$

$$CH_3O_2 + HO_2 \longrightarrow CH_3OOH + O_2. \tag{2.77}$$

In the absence of NO, the latter equation is dominant and CH_3OOH is additionally decomposed to CH_3O and HCHO:

$$CH_3OOH + OH \longrightarrow HCHO + OH + H_2O \tag{2.78}$$

$$CH_3OOH + OH \longrightarrow CH_3O_2 + H_2O \tag{2.79}$$

$$CH_3OOH + h\nu \longrightarrow CH_3O + OH. \tag{2.80}$$

Finally, CH_3O from the Equations 2.76 and 2.80 directly reacts with O_2 to form HCHO (in addition to Eq. 2.78):

$$CH_3O + O_2 \longrightarrow HCHO + HO_2. \tag{2.81}$$

HCHO is destroyed either by photolysis or by reaction with OH (or by wet deposition, see

e.g. MELLER and MOORTGAT, 2000 for further details):

$$\mathrm{HCHO} + h\nu \xrightarrow{\lambda<365\,\mathrm{nm}} \mathrm{CO} + \mathrm{H_2} \tag{2.82}$$

$$\mathrm{HCHO} + h\nu \xrightarrow{\lambda<340\,\mathrm{nm}} \mathrm{HCO} + \mathrm{H} \tag{2.83}$$

$$\mathrm{HCHO} + \mathrm{OH} \longrightarrow \mathrm{CHO} + \mathrm{H_2O}. \tag{2.84}$$

Due to this sunlight dependence, there is a clear diurnal cycle with wet and dry deposition as dominant losses during the night. In total, photolysis accounts for $63\,\%$ of all losses, OH reactions for $29\,\%$, followed by wet and dry deposition, $6\,\%$ and $2\,\%$, respectively (data from unpublished TM4 model studies, personal communication, M. Vrekoussis, IUP Bremen).

Combustion processes

Due to the importance of combustion processes within this thesis, a brief introduction on the most important chemical reactions is given below.

Combustion processes usually involve combustible materials and an oxidizer to form an exothermic reaction in order to release heat. For anthropogenic combustion processes, this heat is then usually converted to kinetic energy (e.g. within engines). To be more precise, anthropogenic combustion refers to processes in which hydrocarbons react with oxygen to release heat and further products as e.g. carbon dioxide and water. Reactions as

$$\mathrm{C}_x\mathrm{H}_y + \left(x + \frac{y}{4}\right)\mathrm{O_2} \longrightarrow \frac{y}{2}\mathrm{H_2O} + x\mathrm{CO_2} \tag{2.85}$$

frequently happen during combustion processes (see e.g. FLAGAN and SEINFELD, 1988) and one of the simplest examples is the burning of methane (e.g. as included in biogas):

$$\mathrm{CH_4} + 2\mathrm{O_2} \longrightarrow 2\mathrm{H_2O} + \mathrm{CO_2}. \tag{2.86}$$

Within car engines, the combustion process is usually not clean meaning that additional substances are added to gasoline or diesel in order to increase the efficiency. However, these substances lead to further emission quantities such as CO, NO, $\mathrm{NO_2}$ or $\mathrm{SO_2}$. The clean burning of gasoline consists of reactions of hydrocarbons with 4 to 12 carbon atoms per molecule. For example

$$2\mathrm{C_6H_{14}} + 19\mathrm{O_2} \longrightarrow 14\mathrm{H_2O} + 12\mathrm{CO_2} \tag{2.87}$$

for gasoline or with e.g. $\mathrm{C_{12}H_{23}}$ for diesel. In addition to these hydrocarbons, alkanes,

alkenes and cycloalkanes are added to the fuel mixture in order to optimize the burning. Furthermore, since the oxygen for these combustion processes comes usually from ambient air, additional nitrogen molecules and other constituents (cf. Sec. 2.2.1) enter the burning process. This leads to the problem that there is strong pollutant formation during combustion within engines (or coal, biogas burning power plants). Two major mechanisms for the creation of NO_x within diesel engines are the so called *Fenimore* and *Zeldovich* mechanisms. The first case, also known as *prompt* NO_x, is formed within the cylinder at the front of the flame. CH radicals react with nitrogen molecules to form cynhydric acid (HCN):

$$CH + N_2 \longleftrightarrow HCN + N. \tag{2.88}$$

After the dissociation to N, NO can be created with the reaction of OH or O. This mechanisms does only lead to a smaller amount of NO compared to *Zeldovich*. Here, molecular nitrogen reacts with O, O_2 or OH under lower temperature conditions (but sufficient heat is still necessary for the creation of atomic O):

$$O + N_2 \longleftrightarrow NO + N \tag{2.89}$$

$$O_2 + N \longleftrightarrow NO + O \tag{2.90}$$

$$OH + N \longleftrightarrow NO + H \tag{2.91}$$

For both mechanisms, the equilibrium is shifted towards NO when the temperature is higher. Conversion to NO_2 might happen with Equation 2.66.
One problem arises when considering the efficiency of an engine. Since the burning is fast enough so that it can be considered as adiabatic, the total process can be assumed to be a *Carnot-cycle*[32]. Then the efficiency η_C can be found with the lowest T_l and highest temperature T_h within this cycle:

$$\eta_C = \frac{T_h - T_l}{T_l} \tag{2.92}$$

In other words, the higher the maximum temperature T_h the better the efficiency. On the other hand, when the temperature within the burning process is high, then the formation of pollutants increases. This indicates the difficulties of the development of efficient engines together with a more and more complicated after-treatment of exhaust gases in order to decrease the release of pollutants. A more detailed discussion is out of the scope of this thesis but the reader is referred to FLAGAN and SEINFELD [1988] for further details on air pollution due to combustion processes.

[32] *The Carnot cycle is a theoretical thermodynamic cycle which provides an upper limit on the efficiency that any classical thermodynamic engine can achieve during the conversion of heat into work*. Definition from wikipedia, `https://en.wikipedia.org/wiki/Carnot_cycle#Efficiency_of_real_heat_engines`, date: 20.09.2018.

2.7 Aerosols

So far, the terminology *aerosol* was used without giving a detailed explanation about their properties, impact on radiative transfer and the DOAS analysis. However, important quantities as the single scattering albedo ω, Henyey-Greenstein phase function ($P_{HG}(\Theta)$), and the aerosol size ratio x were already introduced (see Sec. 2.3.2.2).

In general, aerosols are defined as the suspension of fine particles (or particulate matter, PM) or liquid droplets in air or another gas (see e.g. BURROWS et al., 2011). Although the terminology aerosol describes the gas and the suspended matter together, in practice usually only the particles or droplets are considered in the description of aerosols.

Even though aerosols can be classified with the help of different properties (e.g. shape, mass, chemical composition, aerodynamic features...) size stands out because it can directly be linked with e.g. the source of particles, the impact on human health or Earth's radiative budget. The size (diameter) of suspended particles ranges from a few nanometres for aerosols with the origin of e.g. combustion processes up to $0.1\,\mathrm{mm}$ for e.g. dust and pollens. Due to this large range, several classifications are commonly used. As a rough distinction, aerosols are separated into fine and coarse particles with the separation limit of $2.5\,\mu\mathrm{m}$ (FINLAYSON-PITTS and PITTS JR, 1999). Here, the quantities PM_{10} and $PM_{2.5}$ are commonly used as the description of aerosols with effective diameters smaller than $10\,\mu\mathrm{m}$ and $2.5\,\mu\mathrm{m}$, respectively.

To be more accurate, the fine mode itself can be further separated into ultra fine particles ($\leq 0.01\,\mu\mathrm{m}$), *Aitken nuclei* ($0.01\,\mu\mathrm{m}$ to $0.1\,\mu\mathrm{m}$), and an accumulation range ($0.1\,\mu\mathrm{m}$ to $2\,\mu\mathrm{m}$). Since particles do not necessarily have to be spherical, their size or diameter should rather be understood as an effective value instead of a specific description.

Furthermore, in the real atmosphere two or more different aerosol types might be present at a specific location so that one value such as the diameter is not enough for a full description. *Size distributions* are often used for a more accurate representation which are then described by a log-normal distribution of the aerosol number concentration N_a for n different modes.

$$\frac{dN_a(r)}{dr} = \sum_{i=1}^{n} \frac{N_{a_i}}{r\sqrt{2\pi}\ln\sigma_i} \exp\left(-\frac{\ln^2(r/r_i)}{2\ln^2\sigma_i}\right) \tag{2.93}$$

Here, σ_i represents the standard deviation and r_i the geometric mean radius of the specific mode n. In many cases, bimodal size distributions ($n = 2$) are enough as there are often only two dominant aerosol types present (e.g. sea-salt and combustion related aerosols in coastal urban regions).

In addition to these microphysical aerosol properties, also optical properties are frequently used and are equally important as they are needed in the context of RTM calculations or DOAS analyses.
In the previous sections, the single scattering albedo ω was introduced as the ratio of scattering coefficient to extinction coefficient (Eq. 2.34). In many cases, this number does not vary strongly for a certain measurement location so that a constant aerosol type can usually be assumed. On the other hand, episodic aerosol events like Sahara dust transport to Europe or forest fire plumes might lead to changes in aerosol type over time, creating the need for dynamic adaptations to the prevalent situations.
The *aerosol optical depth* $\tau_(a)$ (AOD) is the altitude depending extinction coefficient $\beta_e(z)$ integrated over the altitude:

$$\tau_a = \int_0^{TOA} \beta_e(z)dz \tag{2.94}$$

In general, this quantity is wavelength dependent with the *Angström Exponent* α as a measure for the wavelength dependency which is frequently used to distinguish between aerosol types (see. e.g. TOLEDANO et al., 2007).

$$\frac{\tau_a(\lambda)}{\tau_a(\lambda_0)} = \left(\frac{\lambda}{\lambda_0}\right)^{-\alpha} \tag{2.95}$$

Here, two different AOD's $\tau_a(\lambda_0)$ and $\tau_a(\lambda)$ are measured at the wavelengths λ_0 and λ, respectively.
In this thesis, AOD, $\beta_e(z)$, ω and $P_{HG}(\Theta)$ (Eq. 2.30) are used to parametrize aerosols. They have an impact on MAX-DOAS measurements as they change the light path and therefore the differential slant column densities. How much they affect the measurement depends on the aerosol type, their location relative to the dominant trace gas concentration and the general geometry. For example, aerosols with a high SSA and g close to the surface might reduce the measured trace gas signal when measuring at low elevation angles because the effective light path is reduced due to the additional scattering events. In addition, absorbing aerosols decrease the signal due to a general loss of photons by absorption. For discussion of aerosols impact on O_4 dSCD's, the reader is referred to WAGNER et al. [2004]; FRIESS et al. [2006].

CHAPTER

III

MAX-DOAS MEASUREMENTS IN BREMEN

In this chapter, results of MAX-DOAS measurements taken on the roof of the institute of environmental physics (IUP) in Bremen are presented. It starts with a short introduction into the weather conditions in Bremen, the measurement location itself including nearby emission sources, followed by a description of the instrument, the data calibration and preparation.

3.1 Measurement location and weather conditions in Bremen

The city of Bremen (53.1°N, 8.8°E) has ~ 565.000 inhabitants and lies in the northern part of Germany with a distance of approximately $60\,\mathrm{km}$ to the North sea. The area around the city is dominated by flat terrain with a typical mid-latitudinal flora and fauna.

Weather conditions

The wind comes mainly from south-westerly directions (compare Fig. 3.1), with weak to medium wind speeds (low seasonal variations can be seen in Fig.AF.1, in the appendix). Due to these wind directions, urban aerosol types should be prevalent at the measuring site but also sea salt events are possible when oceanic air masses are transported towards the city. It is expected, that long range transport from the Ruhr area or the Netherlands is rare, so that the main trace gas signal is due to close emission sources of traffic, industry and agriculture. No larger city is located in a distance less than 100 km but the smaller city of Oldenburg (40 km to the west) might introduce further pollution events.

The sky is on average 316 days per year partly covered with clouds (coverage of 4/8 or higher[2]) and on 217 days per year fully or nearly fully covered (7/8 or higher). This shows the need for an analysis of longer time series for statistical reasons. In general, cloudy or rainy days have to be discussed separately since the analysis is more complicated e.g.

[1] Source: FTP Server - DWD [2018] - `ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany/climate/hourly/`, Date of download: 14.03.2018.

[2] For the cloud coverage classification, the sky is divided into 8 segments. If up to 3 segments show clouds, the sky is referred to as cloud free (0/8) or partly clouded (3/8). Between 4/8 and 6/8, the sky is cloudy. 7/8 or higher means nearly or full cloud coverage.

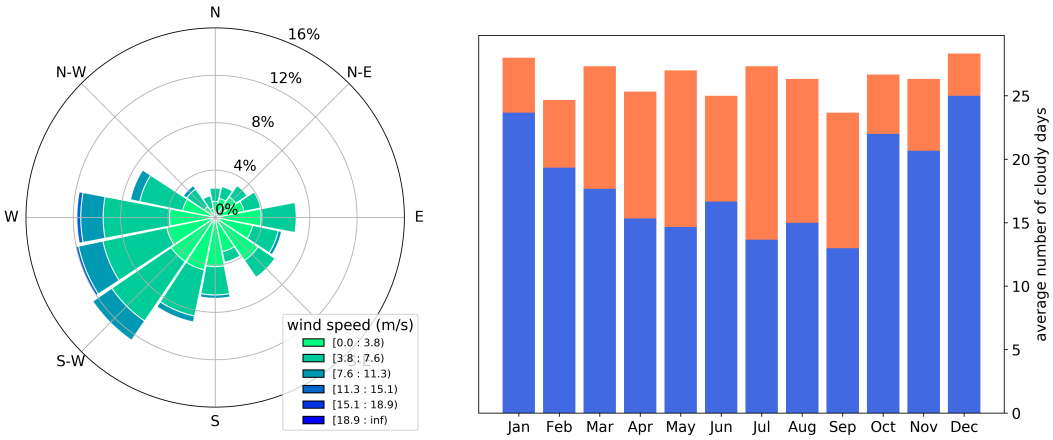


Fig. 3.1: Left: *Windrose of hourly wind data measured at Bremen airport with wind speed colour coded and percentages of all days as pie radius for the certain directions.* ***Right:*** *Average number of cloudy days per month. Blue bars indicate nearly or fully covered days (coverage: $\geq 7/8$), orange bars depict partially cloudy days (coverage: $\geq 4/8$). Both graphs include the years 2015 - 2017*[1].

due to changes in light paths, possible convection processes or wet deposition. Therefore, cloudy days are discarded within the analysis of this thesis which decreases the amount of data drastically.

Measurement location

The MAX-DOAS site in Bremen is part of the Bremian DOAS Network for Atmospheric Measurements (BREDOM) and provides data since the early 1990th. Starting with a zenith-only instrument, the telescope is nowadays able to measure in all hemispherical directions with a visible and a UV spectrometer (see Sec. 3.2 for further details) for a spectral coverage of the most important trace gas absorption bands.
Here, only three years of data from 2015 to 2017 are analysed (see also data availability in Fig. 3.5) because within these years, the instrument pointed always into the same three azimuthal directions. Although the MAX-DOAS instrument in Bremen has measured for many years, frequent changes in the set-up, viewing directions and spectral ranges prevent detailed and longer data evaluation under similar conditions. In Figure 3.2, the pointing directions of the instrument as well as the locations of power plants, wind measurement site and in situ[3] stations are shown. In general, the pollution over the city of Bremen is expected to be dominated by traffic emissions which should be measurable with data from direction s even though this is not exactly where the city centre is located. The prevalent wind direction should lead to enhanced pollution signal in the measurements. On the other hand, depending on the wind direction, the emission plume of the *SWB heat & po-*

[3] Data provided by Bremen's air quality network BLUES (Bremer Luftüberwachungssystem), BLUES [2017].

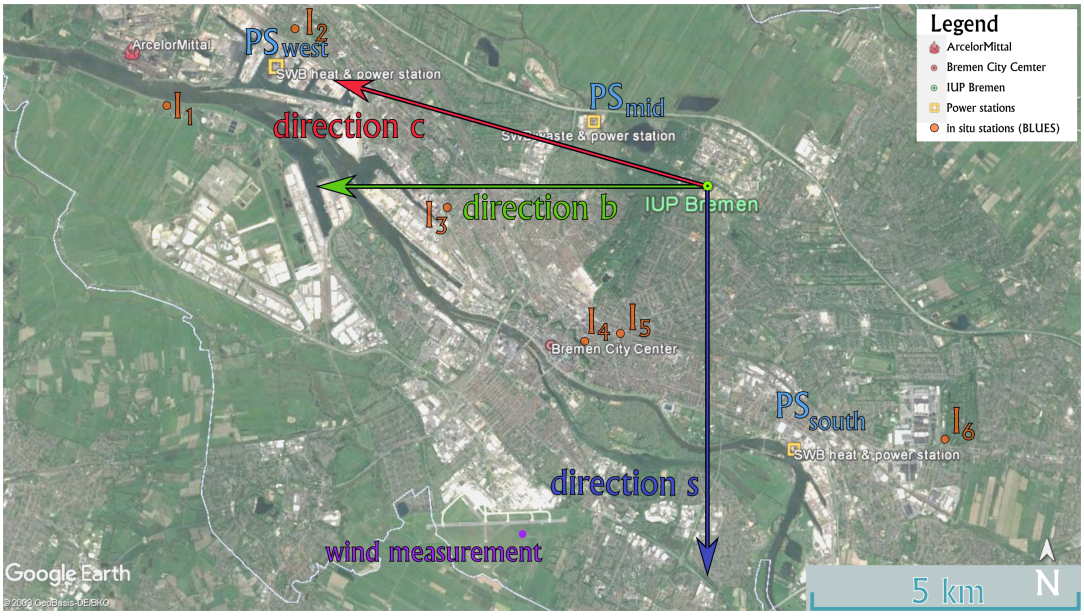


Fig. 3.2: *Map of Bremen with viewing directions of the MAX-DOAS telescope depicted as coloured arrows. Three power stations (PS) are represented as yellow rectangles. Orange dots show the location of different in situ stations (I_i). The purple dot shows the measurement location of climatology data at Bremen airport.*

wer station (PS_{south}) as well as emissions of air planes starting or landing at the regional airport of Bremen should increase the trace gas signal additionally. Direction b and c are arranged to measure emissions of two further power stations (PS_{west} and PS_{mid}) and the steel processing factory *ArcelorMittal Bremen GmbH*. The latter one is the largest emitter for nitrogen dioxide in the area of Bremen. The following table summarizes the average NO_2 emissions for the shown power plants and factories for the year 2015. All industrial complexes emit further species like e.g. carbon dioxide (CO_2) or nitrous oxide (N_2O) but since this thesis focuses on nitrogen dioxide, other compounds emitted are not discussed.

Name	**Abbrev.**	**NO_x/NO_2 (t/a)**	**SO_x/SO_2 (t/a)**
SWB heat & power, Hastedt	PS_{south}	576, 669	272, 297
SWB disposal, MHKW Bremen	PS_{mid}	640, 515	-
SWB heat & power, Hafen	PS_{west}	1070, 1070	961, 917
SWB disposal, MKK Bremen	PS_{west}	265, 261	-
ArcelorMittal/ Walzwerk	Arcelor	353, 396	-
ArcelorMittal	Mittal	1720, 2270	3350, 2940

Tab. 3.1: *Yearly average emission values of nitrogen and sulphur oxides represented as the ratios of nitrogen/sulphur oxides and nitrogen/sulphur dioxide for the largest emitters in Bremen for the year 2015, 2016 (in t per year). Data provided by the European Pollutant Release and Transfer Register,* E-PRTR [2018].

Furthermore, map 3.2 includes in situ stations of Bremen's air quality network (BLUES) as orange dots. Here, this data is used to validate results by comparing time series and daily and seasonal variations of nitrogen dioxide. The in situ stations provide data for different species and climatological datasets and are listed either as background or as traffic monitoring sites. The depicted in situ stations are described in Table 3.2. All stations measured throughout the full time period (2015 - 2017). Note, that I_1, I_2 and I_3 are located close to the industrial area around the river *Weser* for monitoring emissions from PS_{west} and ArcelorMittal. Together with measurements of direction c, this data will be used to quantify the emission sources (Sec. 3.7) with results which should match the data of the European Pollutant Release and Transfer Register (E-PRTR) listed in Table 3.1. Note, that although I_1 and I_4 provide additional climatology data, only measurements from the DWD (Deutscher Wetterdienst) site at the airport are used.

Name	Symbol	type	Species & data
Hasenbüren	I_1	BK	O_3, SO_2, NO, NO_2, PM10, PM2.5, cli.
Oslebshausen	I_2	BK	SO_2, NO, NO_2, NO_x, CO, PM10
Nordstraße	I_3	TFC	NO, NO_2, NO_x, PM10
Bremen Mitte	I_4	BK	O_3, SO_2, NO, NO_2, CO, PM10, cli.
Dobben	I_5	TFC	NO, NO_2, CO, PM10, Benzol
Bremen Ost	I_6	BK	O_3, SO_2, NO, NO_2, PM2.5

Tab. 3.2: *List of in situ stations and measured species of the Bremen air quality network (BLUES, 2017) within the city of Bremen. BK: background station, TFC: traffic monitoring site, cli: climatology data (pressure, temperature, wind speed and direction & rel. humidity).*

3.2 Instrumentation

Figure 2.14 shows a schematic representation of the measurement geometry of a MAX-DOAS instrument in Bremen. The pan-tilt head is able to point the telescope into any azimuthal direction φ with elevation angles α varying from horizon to zenith. In Section 2.5.2, measurement geometry describing angles like RAA (φ_{raa}, Eq. 2.51), SZA (ϑ) and VAA (φ_{vaa}) were already defined. Note that in this thesis, the viewing azimuth angle is always measured as clockwise with respect to the north.

The telescope itself is mounted on the roof of the IUP building in an altitude of $\sim 26\,\mathrm{m}$ above the surface, with the possibility of pointing into $\sim 270°$ azimuthal viewing directions. Only between the northern ($0°$) and western ($\sim 90°$) directions, parts of the building obscure the telescope's line of sight. The telescope unit can be seen in Figure 3.3 and is explained in detail in PETERS [2013].

Briefly: the light enters the telescope through a fused silica window (A) within a specific opening angle (also field of view, FOV) determined by a lens focussing the light into a light fibre mounted on an optical bench (B). The total field of view is $\sim 1°$ for the used telescope.

Additional elements are a video camera (E), which takes pictures in the pointing direction, a drying agent (F), which reduces the humidity within the telescope and calibration installations (C, D). During night, calibration measurements are routinely done by using a HgCd line lamp (D). To do this, the telescope points to the floor, the shutter falls down and opens the light path from the line lamp to the lens by reflecting the light with a Teflon coating on the shutter. The telescope unit is connected via four different cables. Two cables provide communication and power supply to the pan-tilt head and the line lamp. The third one is a USB cable for the power supply and video transmission of the camera. The last and most

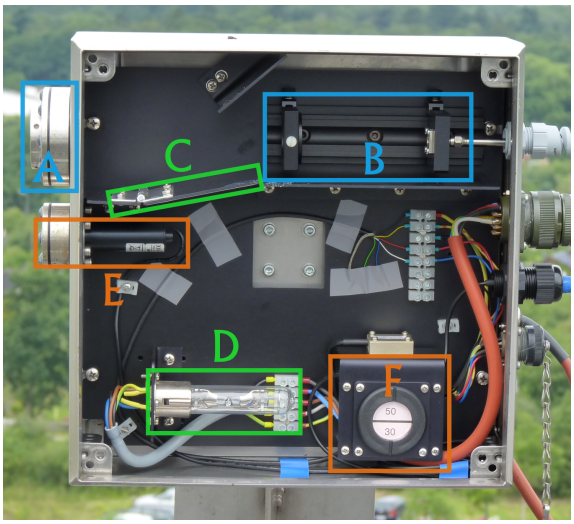


Fig. 3.3: *The MAX-DOAS telescope unit with elements of the optical path in blue, calibration components in green and additional elements in orange. A: Fused silica light entrance window, B: Optical bank (left: focusing lens, right: fibre mount and end of glass fibre), C: gravity driven Teflon shutter, D: HgCd calibration lamp, E: Fused silica window including video camera, F: drying agent.*

important cable attached to the telescope is the quartz light fibre. It consists of 76 individual light fibres which are split up in the laboratory to feed two spectrometers, one for the UV ($305\,\mathrm{nm}$ to $390\,\mathrm{nm}$) and one for the visible spectral range ($406\,\mathrm{nm}$ to $579\,\mathrm{nm}$). The two spectrometers (both Acton500) are build in the *Czerny-Turner* design and are temperature stabilised at $\sim 35^\circ$ C to prevent wavelength depending temperature shifts. This design means, that light which enters the spectrometers is reflected via a collimating mirror on a grating which diffracts the light on a focusing mirror which then focuses the spectrally separated light onto a CCD (*Charge-Coupled Device*). The spectral resolution (FWHM of slit function, see Sec. 3.3) of the UV and visible spectrometers are $0.5\,\mathrm{nm}$ and $0.85\,\mathrm{nm}$ respectively. The CCD's (both Roper Scientific, Princeton NTE/CCD-1340/400-EMB) are cooled down to -30° C to decrease dark current[4] and consist of semi conductors with 400×1340 pixels. In the pixel column's direction, the spectral signal is observed while the other dimension (rows) is used to increase the *signal-to-noise ratio* (SNR) by averaging all signals along this axis. In contrast to that, imaging instruments use the 2nd dimension to resolve all LOS instantaneously by projecting sorted light fibres on the different rows.

[4] Dark current means the spontaneous creation of free charges within semi-conductive units (e.g. CCD) when no external energy entered. The amount of dark current is strongly temperature depending.

3.3 Calibration

Data measured by the CCD is uncalibrated which necessitates calibration steps before the DOAS fit can be performed. This is done with the programs *nprepare* and *resolut* and is shortly explained below.
After the CCD read-out, data consists of a number of counts per pixel. This data is checked for saturation or for too low signals. The next step is the subtraction of the dark measurement to account for a possible offset of the spectrum, noise and the dark current (steps included in nprepare). These dark measurements are usually done on a daily basis during the nights by closing the telescopes Teflon shutter. The last step is the spectral calibration with the help of the before mentioned line lamp measurements (software: nkalib). The HgCd lamp provides well distinguishable absorption lines at known wavelengths. A measurement leads to high photon numbers on specific pixel columns with a smearing effect due to the limited spectral resolution of the detector onto nearby pixels. The emission lines can be assigned to the reference wavelength and the smearing is used to evaluate the instruments slit function (software: resolut). The final products are calibrated spectra with the unit counts per second and slit function files for the convolution of absorption cross sections explained in the next section.

3.4 Application of the DOAS fit

The DOAS fit (Eq. 2.47) is done with the Delphi-written in-house software package *NLIN_D* (RICHTER, 1997). Several parameters have to be defined for the individual fit settings and possible absorbers within the measurement. The most important parameters are explained in the following:

- **Fitting window:** The start and end wavelength of the spectral range used for fitting has to be selected. Ideally, this range should include the strongest absorption bands of the absorber of interest while minimizing the number of strong Fraunhofer lines and the absorption of other interfering species. If this is not possible, other absorption bands of the absorber of interest have to be selected.
- **Order of the polynomial:** It should be chosen large enough to account for Rayleigh and Mie scattering as well as broadband absorptions and cloud offsets. But orders too large should be avoided, because the number of spectral points within the fitting window might be limited and absorption lines might be decreased when the polynomials curvature/oscillation matches the specific wavelength of an absorption pattern.
- **Absorption cross sections:** Trace gases which might have contributed to the absorption within the measured spectral range have to be accounted for. This might also

involve several cross sections for one species at different temperatures or to account for higher order wavelength dependencies. These cross sections are then orthogonalized to prevent cross correlations (e.g. SANDERS, 1996, RICHTER, 1997 or PUĶĪTE et al., 2010). Furthermore, other absorber cross sections are frequently used even though there might not be a satisfying signal of the specie within the measurement, to reduce the effect of cross correlations of different absorbers (e.g BrO cross section within HCHO fits).

- **Background spectrum:** The reference spectrum I_0 has to be chosen to satisfy the needs of the user. This might be a temporally close zenith spectrum to reduce the stratospheric signal of some absorber or, vice versa, a spectrum around noon for the investigation of absorbers in the stratosphere.

Before applying the DOAS fit within NLIN_D, the convolution[5] of the absorption cross sections and a *Fraunhofer atlas* (KURUCZ, 1984) with the instruments slit function is done in order to resolve all spectral features on the instruments limited spectral resolution. Then, a second calibration of the reference spectrum I_0 with the high resolution Fraunhofer atlas is applied to gain further spectral accuracy. A shift and squeeze of I_0 is allowed here. Finally the DOAS fit is started. This fit consists of two individual steps which are repeated iteratively.
A non linear wavelength fit is used to fit $\ln I$ to $\ln I_0$ by applying another shift and squeeze. The absorption cross sections and polynomial coefficients are then linearly fitted to the resulting optical depth $\Delta\tau = \ln I_0/I$[6]. This linear fit can be represented as a minimization problem with a residual term $r(\lambda)$:

$$\left\| \ln\left(\frac{I_0(\lambda)}{I(\lambda)}\right) \; - \; \sum_i \sigma_i(\lambda)\Delta S_i \; - \; \sum_j a_j\lambda^j - r(\lambda) \right\|^2 \longrightarrow \min \tag{3.1}$$

Here, the first sum is over all absorbers i and the second sum includes a polynomial with order j. This minimization problem can be reformulated to $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 \longrightarrow \min$, where $\mathbf{y}$ is the measurement vector ($\Delta\tau$), $\mathbf{A}$ is the DOAS matrix consisting of all cross sections and wavelengths arrays with the orders $1, \lambda, \lambda^2, ...\lambda^j$ and $\mathbf{x}$ is the vector of interest including all ΔS_i and j polynomial coefficients. Vector $\mathbf{x}$ can be retrieved by inverting $\mathbf{A}$, so that $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$[7]leads to the final fitting quantities ΔS_i and a_j.
This minimization is solved iteratively and the root mean square (RMS) of the residual

[5] The convolution of two functions f and g is defined as: $(f * g)(\lambda) = \int_{-\infty}^{\infty} f(\lambda')g(\lambda - \lambda')d\lambda'$.
[6] In contrast to τ, $\Delta\tau$ indicates, that the reference spectrum I_0 includes also absorption features of the investigated species.
[7] Note that $\mathbf{A}$ is generally not invertible so that a Singular Value Decomposition (SVD) is applied for the calculation of a pseudo-inverse matrix.

term weighted by the number of wavelength is a frequently used fit quality criterion:

$$RMS = \sqrt{\frac{\sum_i r(\lambda_i)^2}{N}} = \sqrt{\frac{\chi^2}{N}}. \tag{3.2}$$

In the last step, χ^2 was introduced which is defined as follows (see e.g. PRESS, 2007 for further details):

$$\chi^2 = \sum_{i=0}^{N-1} \left(\frac{y_i - A_i x_i}{\sigma_i} \right)^2 \tag{3.3}$$

Here, the error of each individual measurement at wavelength i, σ_i is introduced but it is generally chosen to be $\sigma_i = 1$ because a quantification of all error sources is usually not possible. The errors of the individual fitting quantities are now defined with the help of χ^2 as

$$\sigma_{\Delta S} = \sqrt{\frac{\chi^2}{N - N_{par}} \cdot S_p(x_i)} \tag{3.4}$$

with the covariance matrix $S_p(x_i)$ of the specific parameter x_i retrieved during the inversion via singular value decomposition (SVD). Here, N_{par} is the number of fitting parameters (number of cross sections plus polynomial coefficients). Since the errors of individual wavelengths are not quantifiable, the fitting parameter errors include a weighted sum of all spectral residual points ($\sqrt{\sum_i (y_i - A_i x_i)^2/(N - N_{par})}$) and the variances of each parameter of the DOAS matrix ($\sqrt{S(x_i)}$).
In general, Equation 3.1 describes an over-determined system because the number of wavelengths is usually larger than the number of fitting parameter. However, if the fitting window is small and the number of cross sections and the order of the polynomial are large, the inversion might fail, leading to poor results.

Straylight correction

Light which is measured on the wrong pixels of the CCD is usually referred to as *straylight*. This occurs when the spectrometer or the CCD mounting are not fully light-tight or when light entering the spectrometer via the light fibre is reflected within the spectrometer instead of following the designed light path. Often, straylight creates an additive offset wit little wavelength dependency. With the assumption of a constant straylight c measured in addition to the sky intensity I, the optical depth can be adapted by assuming a Taylor

series expansion of the logarithm ($\ln(1+x) = x - x^2/2 + ...$):

$$\begin{aligned}\tau(\lambda) &= \ln\left(\frac{I(\lambda)+c}{I_0(\lambda)}\right) = \ln\left(\frac{I(\lambda)}{I_0(\lambda)} + \frac{c}{I_0(\lambda)}\right) \\ &= \ln\left(\frac{I(\lambda)}{I_0(\lambda)}\right) + \ln\left(1+\frac{c}{I(\lambda)}\right) \approx \ln\left(\frac{I(\lambda)}{I_0(\lambda)}\right) + \frac{c}{I(\lambda)}\end{aligned} \tag{3.5}$$

Here, the Taylor series was used in first approximation neglecting higher order terms. As a first guess, the constant c is set to $c = 0.03\, I_{max}$, which means that we assume it to be 3% of the maximal intensity. The straylight term is added as an additional cross section σ_{off} within the DOAS fit and its fitting quantity is frequently referred to as *offset*:

$$\sigma_{\text{off}}(\lambda) = \frac{0.03\, I_{max}}{I(\lambda)} \tag{3.6}$$

In addition, sometimes a wavelength depending straylight is assumed by introducing a second straylight cross section σ_{slo} which is created by multiplying the factor $(\lambda_2 - \lambda)/\lambda_2$ (λ_2 is the upper fitting wavelength) to σ_{off}. Note, that due to the dependency on $1/I(\lambda)$, the straylight correction is strongly correlated with the ring effect. Note that other groups include typically a 2nd order intensity offset cross section in addition to the above mentioned σ_{off}.

Ring effect

The *Ring effect* (see Sec. 2.3.2.2) is accounted for by including an additional cross section σ_{ring} to the DOAS fit. This cross section is usually calculated with the help of a RTM. Here, two SCIATRAN-simulated radiance spectra with and without Raman scattering are used and separately convolved with the instrument's slit function. The logarithm of the ratio of both spectra is used as an additional cross section.

I_0-correction

By applying Equation 3.1 it was assumed that the measured intensities can be written as convolved intensities $I^*(\lambda) = F(\lambda) * I(\lambda)$ and $I^*_{ref}(\lambda) = F(\lambda) * I(\lambda)$, so that a convolution of the cross sections yields in the following approximated:

$$I^*(\lambda) = F(\lambda) * \left(I_o(\lambda) e^{-S\sigma(\lambda)}\right) \approx (F(\lambda) * I_0(\lambda)) e^{-S(F(\lambda)*\sigma(\lambda))}$$

Here, one absorber only was considered and $I_0(\lambda)$ is a solar spectrum instead of the reference spectrum $I_{ref}(\lambda)$. This approximation is only valid, if the slit function is much smaller than the Fraunhofer lines and the absorption is not too large. However, for strong absorbers such as NO_2 and O_3, this might lead to uncertainties which are usually referred

to as I_0 effect. The problem is solved by applying an empirical solution scheme. The specific absorber's cross sections are replaced by cross sections which were calculated by solving highly resolved Lambert-Beer laws with assumed slant column densities for each absorber (see RICHTER, 1997 or ALIWELL et al., 2002). This approach needs a first guess for each absorbers slant column density which introduces further uncertainties but decrease problems due to the I_0 effect.

DOAS fitting example

In Figure 3.4, the fitting results of a MAX-DOAS fit for the investigation of NO_2 amounts in Bremen are shown. The first column depicts the individual retrieval steps starting with the logarithms of both intensity spectra I and I_0 (top left) and the resulting optical depth including the subtracted polynomial below. After the subtraction of the polynomial, the differential optical depth is shown in the third figure of column 1. The last graph depicts the residual after the DOAS fit, showing that most structures were successfully fitted and no obvious spectral features were overlooked. As a short summary of the DOAS approach, this plot can be described as follows:

- The Fraunhofer lines are the dominant spectral features but cancel out, when creating the logarithm of the intensity ratio.
- The overall slope of the optical depth can be described by a polynomial accounting for all broadband features.
- Within the differential optical depth, one can easily identify the most pronounced absorbing species by comparison with absorption cross sections.
- The residual is much smaller than the optical depth and it usually does not contain structured and specific spectral features indicating that it consists merely of noise.

The other columns in Figure 3.4 depict the *reference* ($\Delta S_i \sigma_i$) and *fitted* ($\Delta S_i \sigma_i$ + residual) spectral structures within the differential optical depth and the matching correlation between both curves as well as the individual ΔS_i values. One can easily see, that NO_2 is the most important absorber within the differential optical depth. The residual can hardly be seen within the NO_2 fit curve and the correlation coefficients[8] are high. H_2O and O_4 are also fitted well, with high correlations. The additional NO_2 cross section (NO_2 at $220\,K$ orthogonalized to $298\,K$ to account for stratospheric variations) and O_3 do not show good correlations. Since one fit is usually used for a long time span with different possible absorption events, daily and seasonal variations of absorbers and other influencing factors, cross sections should not be excluded only due to missing features in one specific measurement. Note, that Ring and Offset show similar absorption patterns (with a multiplication factor of -1) which was already discussed above.

[8] *Pearson's* correlation coefficient $r_{x,y}$ of two stochastic variables x, y is defined as $r_{x,y} = Cov(x,y)/(\sigma_x \sigma_y)$ with the individual standard deviations σ_i and the covariance matrix $Cov(x,y)$.

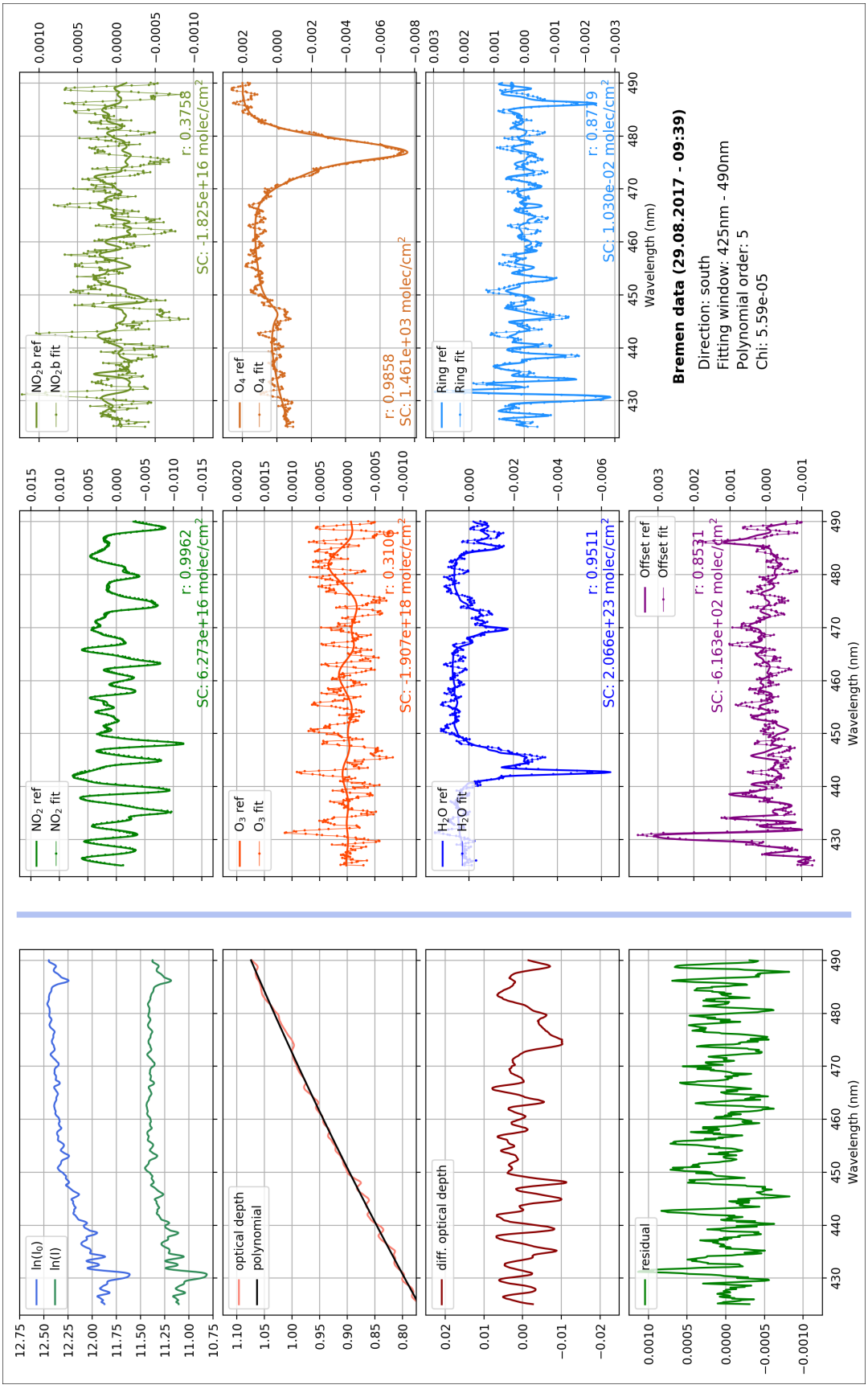


Fig. 3.4: DOAS fit example with the individual spectra $\ln(I)$, $\ln(I_0)$, the fitting parameters (ΔS, poly.) and the residual term.

3.5 Conversion to vertical column densities

In this section, the conversion from differential slant column densities to vertical column densities is introduced. The data was already calibrated (see Sec. 3.3) and successfully fitted (Sec. 3.4). The fit settings were used according to the CINDI-2 Semi-blind Intercomparison Protocol and are described by HENDRICK et al. [2016] and are shortly summarized in Table 3.3 (the campaign is shortly introduced in Sec. 4.4.1). The general data availability for the visible and UV spectrometer with abbreviations K and L respectively is shown in Figure 3.5.
Besides the time when the instrument measured in Cabauw (the Netherlands) during the CINDI-2 campaign, no longer disruptions from the daily measuring procedure took place. However, within the subsequent data analysis (Sec. 3.7), monthly and seasonal mean values are discussed. The October is in that sense poor, as only measurements from the year 2015 are available. The statistical significance for averaged values is therefore reduced.

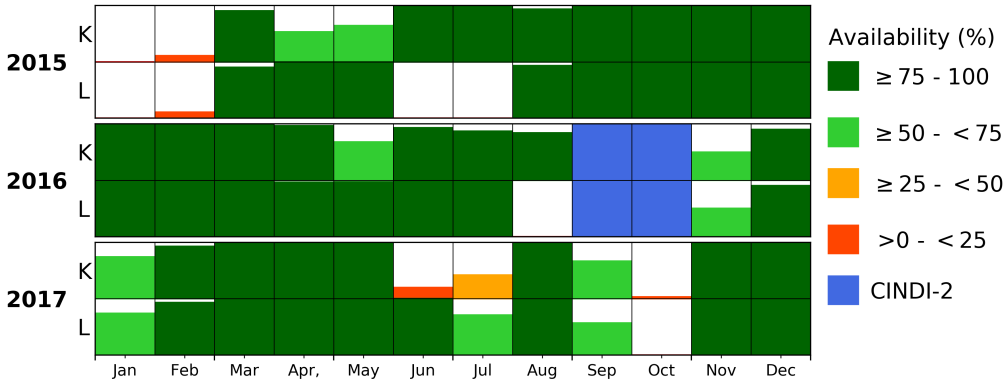


Fig. 3.5: *Availability of data in Bremen (in percent of all days per month) between 2015 and 2017 for the VIS and UV spectrometer K and L, respectively.*

The fit settings in Table 3.3 introduce the official UV (here UV1) and visible (VIS) fits of the CINDI-2 campagin and an in-between fit for which the settings were chosen as a compromise between UV and visible fit (UV2). The three fits are applied to the full dataset of the years 2015 until 2017 to all three azimuthal directions (b,c and s) introduced in Figure 3.2. The data is first converted to vertical column densities and then filtered for certain filtering parameters, explained in Section 3.5.2.
The conversion from ΔS to V is done in two different ways. A climatological AMF or a geometrical AMF (see Eq. 2.55) are applied. The climatological approach was used within the QA4ECV project for the harmonization of the conversion from slant to vertical column densities for all involved groups in order to create a common historical dataset of independent MAX-DOAS measurements. The approach is shortly summarized below (and can be found in HENDRICK et al., 2016a).

Parameter		UV1	UV2	VIS
wavelength range (nm)		338 - 370	367 - 387	425 - 490
polynomial order		5	5	5
intensity offset		const.	const.	const.
cross-section	**species [temp (K)]**			
VANDAELE et al. [1996]	**NO_2** [298]	**x**	**x**	**x**
VANDAELE et al. [1996]	**NO_2b** [220 ▷ 298]	**x**	**x**	**x**
SERDYUCHENKO et al. [2014]	**O_3** [223]	**x**	**x**	**x**
SERDYUCHENKO et al. [2014]	**O_3b** [243 ▷ 223]	**x**	**x**	–
THALMAN and VOLKAMER [2013]	O_4 [293]	**x**	**x**	**x**
ROTHMAN et al. [2010]	**H_2O** [296]	–	–	**x**
MELLER and MOORTGAT [2000]	**HCHO** [297]	**x**	–	–
FLEISCHMANN et al. [2004]	**BrO** [223]	**x**	–	–
RING_QDOAS_SAO2010	**ring**	**x**	**x**	**x**

Tab. 3.3: *Fit settings for all MAX-DOAS data discussed in this thesis for three different fitting windows. Two of them are the official CINDI-2 UV and visible fits (UV1, VIS). UV2 was chosen as an in between fit. The symbol ▷ indicates that a cross sections was orthogonalized to another temperature. The zenith spectrum closest in time to the measurement was used as a reference. The I_0 correction was applied to all NO_2 and O_3 cross sections. The ring cross section was provided during the CINDI-2 campaign by Michel van Roozendaal (BIRA-IASB. Belgium).*

AMF look up tables (LUT) were created by BIRA-IASB with the software packages *bePro/LIDORT* (CLÉMER et al., 2010; SPURR, 2008) including different geometries, wavelength, surface albedos and trace gas and aerosol profiles. Within the QA4ECV project, a *Fortran* based extraction tool was used which utilizes *Aeronet* (AErosol RObotic NETwork, DUBOVIK et al., 2000; HOLBEN et al., 1998) aerosol optical thickness's (AOT), planetary boundary layer heights (PBL) from ECMWF (DEE et al., 2011) and database albedo values (KOELEMEIJER et al., 2003) for the specific location to extract AMF. The profile shapes for the underlying trace gas profiles were exponentially, with fixed vertical columns for NO_2 and HCHO of VC = $2 \times 10^{16}\,\mathrm{molec/cm^2}$. The scale heights were selected by using the PBL values from ECMWF in combination with the fixed vertical column for trace gases and AOTs from AERONET for the aerosol profiles. An aerosol type using the *Henyey-Greenstein* approximation was defined with fixed aerosol properties ($g = 0.72$ and $SSA = 0.9$). Pressure, temperature and ozone profiles were taken from the AFGL database (PRICE and WALKER, 1976). Note, that only cloud-free scenarios are assumed within the AMF extractor tool.
The slant column densities of data from Bremen were converted with help of the QA4ECV extraction tool by applying the $477\,\mathrm{nm}$ AMF to the NO_2 VIS fit and the $360\,\mathrm{nm}$ AMF to UV1 and UV2. The HCHO AMF supports only one wavelength ($343\,\mathrm{nm}$).

In contrast to this more or less automatized approach, the geometric AMF application is easily done with usage of Equation 2.55. However, the resulting V need to be corrected,

because the geometric approach is only valid for low aerosol loads and shallow trace gas layers with scattering points above these layers (see Sec. 2.5.2). SHAIGANFAR et al. [2011] already introduced the implementation of correcting factors summarized in LUTs for the conversion of VC. The approach used within this thesis is explained in detail in the next subsection.

3.5.1 Correction of geometric vertical columns

In Figure 3.6, a correction scheme is depicted explaining all steps from the conversion of ΔS to VC used for the data shown in the following subsections.

First, the geometric DAMF (Eq. 2.55) is used to derive uncorrected vertical columns VC_{uncorr} from ΔS measurements at 30° LOS. Then, these VC_{uncorr} values are corrected by applying correction factors read from look up table which was calculated with SCIATRAN. The LUT was produced with parameters summarized in Table 3.4. Here, the assumption that aerosols and trace gases are well mixed within the planetary boundary layer was made. Therefore, only a box profile with different box heights is used. The general content of the LUTand AOTs are chosen with is: S at 30° and 90° LOS, ΔS_{30} and correction factors defined as ratio of true (V_{true}) to geometric vertical column density (V_{geom}). The proper concentrations, PBL heights the help of ECMWF and CAMS data. The PBL values are used from ECMWF (Era5[9]) and are assumed to match the box heights of the profiles within the LUT. AOT and NO_2 vertical columns were provided by CAMS[10]. The total vertical columns included in CAMS global datasets were corrected for the stratospheric column by using OMI overpass data of Bremen (L2 OVP, NASA, 2018), in order to calculate tropospheric columns in a simple way. Since CAMS and OMI do not necessarily show similar values, the daily stratospheric column is calculated by assuming that the ratio of stratospheric to total column during OMI's overpass time t_o (between 11 and 13 UTC) is similar for both datasets (see Eq. 3.7). The stratospheric column calculated from this equation is subtracted from each

	Parameter	Values
Aerosol	SSA	0.92
	Asym. Fac	0.68
	AOT	0, 0.1, 0.2, ... 1.0
Geometry	RAA (°)	0, 5, 10, ... 180
	SZA (°)	0, 5, 10, ... 90
	VZA (°)	30, 90
	Box height (km)	0.25, 0.5, 0.75, ... 1.5
General	wavelength (nm)	448
	Albedo	0.1
NO_2 **VC**	(1e15 molec/cm^2)	1, 5, 10, 50, 100, 500

Tab. 3.4: *LUT parameters for the geometric vertical column density correction.*

[9] Generated using Copernicus Climate Change Service Information [2018].
[10] As footnote above.

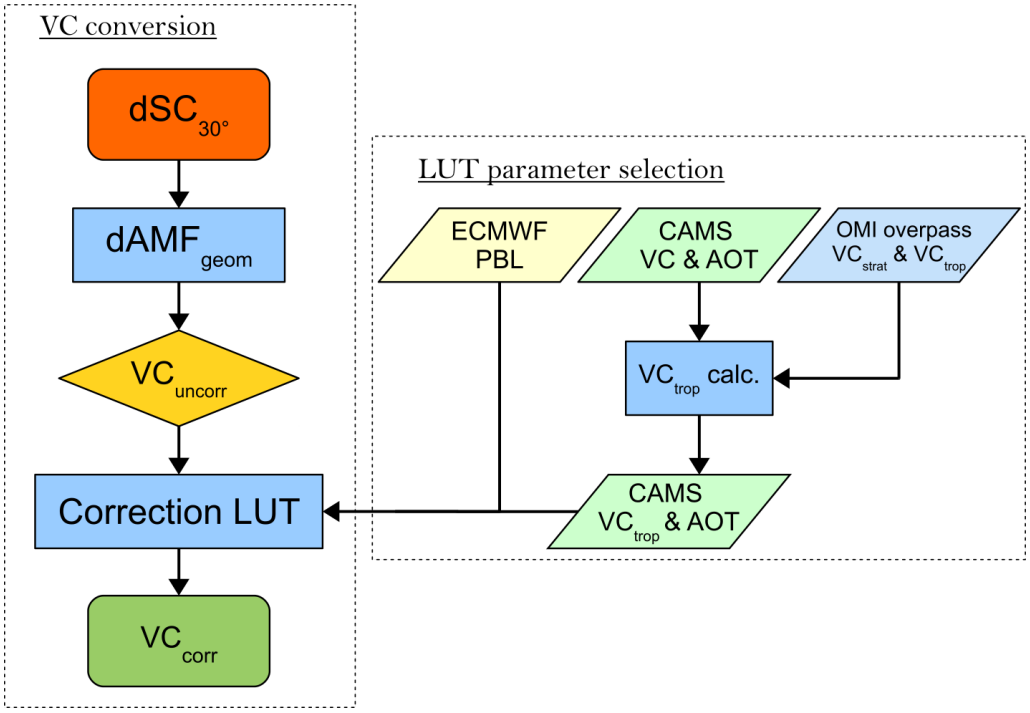


Fig. 3.6: *Flowchart of the conversion steps from differential slant column densities to vertical column densities by applying a corrected geometric air mass factor.*

daily vertical column value from CAMS to obtain tropospheric vertical column densities:

$$VC^{str}_{cams}(t_0) = VC^{tot}_{cams}(t_0)\,\frac{VC^{str}_{omi}(t_o)}{VC^{tot}_{omi}(t_o)} \;\Leftrightarrow\; \frac{VC^{str}_{cams}(t_o)}{VC^{tot}_{cams}(t_o)} = \frac{VC^{str}_{omi}(t_o)}{VC^{tot}_{omi}(t_o)} \tag{3.7}$$

This approach has the flaw, that diurnal variations of stratospheric columns are neglected. However, since the variability of NO_2 in the stratosphere over the day is usually small in comparison to tropospheric values, this approach should lead to smaller uncertainties than the overall negligence of the here introduced correction factor.

3.5.2 Data filtering

As mentioned above, several situations might have a negative impact on calibrated data or the fitting results. Potential issues are:

- Temperature stabilizing problems of the spectrometer or CCD leading to poorly calibrated data or temporal shifts in fitting results.
- Fog or clouds might have an impact on data and fitting quality by changing the light path and/or the intensity.
- The detector might have been in saturation when pointing close to bright surfaces (buildings, clouds, snowy mountains) or in the direction of the sun.

- Low signals of the specific absorbers deteriorate fitting results.
- *Forgotten* absorption cross sections increase fit errors and change retrieved ΔS.
- Insufficient spectral resolution of cross sections.

Some of these effects can be prevented and some might be correctable before the final analysis, but in general, filtering of bad data is a necessary step to maintain a constant level of data quality for longer time series. For the datasets analysed within this thesis, three different quality criteria were used for data filtering.

Uncertainty of DOAS fitting results

Errors of the fitted differential slant column densities (Eq. 3.4) as well as the RMS of the DOAS fit itself (Eq. 3.2) are used as a filtering options. Although the RMS was used during the calculation of the individual ΔS errors, both values are equally important as there is a general difference between the absolute error of a whole fit and a relative error which might be large only due to the specific ΔS values and their matching covariances. Thus, the RMS filter was set to 0.001 whereas slant column errors are not allowed to be larger than 4%.

Detection limit

For every retrieval quantity, there is a minimum value which can be measured (or detected). This minimum value is usually referred to as *detection limit* D. PLATT and STUTZ [2008] defined it as

$$D \approx \sigma \times \frac{6}{\sqrt{n-1}} \tag{3.8}$$

for a given noise level σ, with the number of the detectors pixel n. For the instrument used at IUP Bremen, the detection limit D was found to be $D = 3 \times 10^{14}\,\mathrm{molec/cm^2}$ by PETERS et al. [2012] who used instead an $\overline{D}$ detection limit D_{lim}^{up} defined as twice the RMS divided by the order of the specific absorber cross section N_σ:

$$\overline{D} \approx \frac{2 \cdot RMS}{N_\sigma} \tag{3.9}$$

With a typical fitting root mean square of $RMS \approx 10^{-4}$, these detection limits are $\overline{D}_{NO_2} = 2 \times 10^{15}\,\mathrm{molec/cm^2}$ and $\overline{D}_{HCHO} = 2 \times 10^{16}\,\mathrm{molec/cm^2}$ for NO_2 and HCHO, respectively. Since the HCHO limit is much too high and Equation 3.9 lacks in the dependence of the number of pixels within the fitting window, a new detection limit D' is defined

here as follows (see also the derivation of D' in the appendix A.3.1):

$$D' = \frac{3 \cdot RMS}{\sqrt{\sum_{\lambda} \sigma^2(\lambda)}} \tag{3.10}$$

This detection limit was found to be $1.9 \times 10^{14}\,\mathrm{molec/cm^2}$ for NO_2 and $1.4 \times 10^{15}\,\mathrm{molec/cm^2}$ for HCHO. Smaller values are filtered within the later analysis.

Colour indices

The third and last filtering criterion is the so called *Colour index* (CI). This index is defined as the ratio of two intensity values at different wavelengths λ_1 and λ_2, measured mostly in zenith direction (see WAGNER et al., 2014; GIELEN et al., 2014).

$$CI = \frac{I(\lambda_1)}{I(\lambda_2)} \tag{3.11}$$

The idea behind this approach is, that a cloud usually changes the slope of an intensity spectrum due to a change of the blue colour of the sky. In general, CI values should be larger than one under cloud free conditions, because of Rayleigh scattering, when $\lambda_1 < \lambda_2$ (see Sec. 2.3.2.2). Here, λ_1 and λ_2 should be sufficiently far away in a spectral sense and no strong spectral features should be present around these wavelengths. A cloud decreases these CI values drastically. Since the colour of the sky also depends on SZA, an analysis of CI is usually done for SZA $\leq 85°$. Furthermore, also aerosols change the CI value but can not necessarily considered as problematic for the usual measurements and ΔS evaluation. Therefore, an analysis for a specific location is done in a statistical way by analysing at least one year of data.

In Figure 3.7, results of five different CI ratios are plotted for the full time period from 2015 to 2017 in order to find the best CI value which will be used within this thesis. On the right hand side, the absolute frequency of each CI ratio is shown. Note, that each ratio displays a cloud peak which is, here in Bremen, the most dominant feature of the frequency curves. This peak is important to be filtered out because it mainly results from cloudy scenes. For higher CI values, most ratios tend to have a *nice weather plateau*. This plateau includes sometimes another peak (see $425\,\mathrm{nm}/440\,\mathrm{nm}$) which is due to clear sky days.

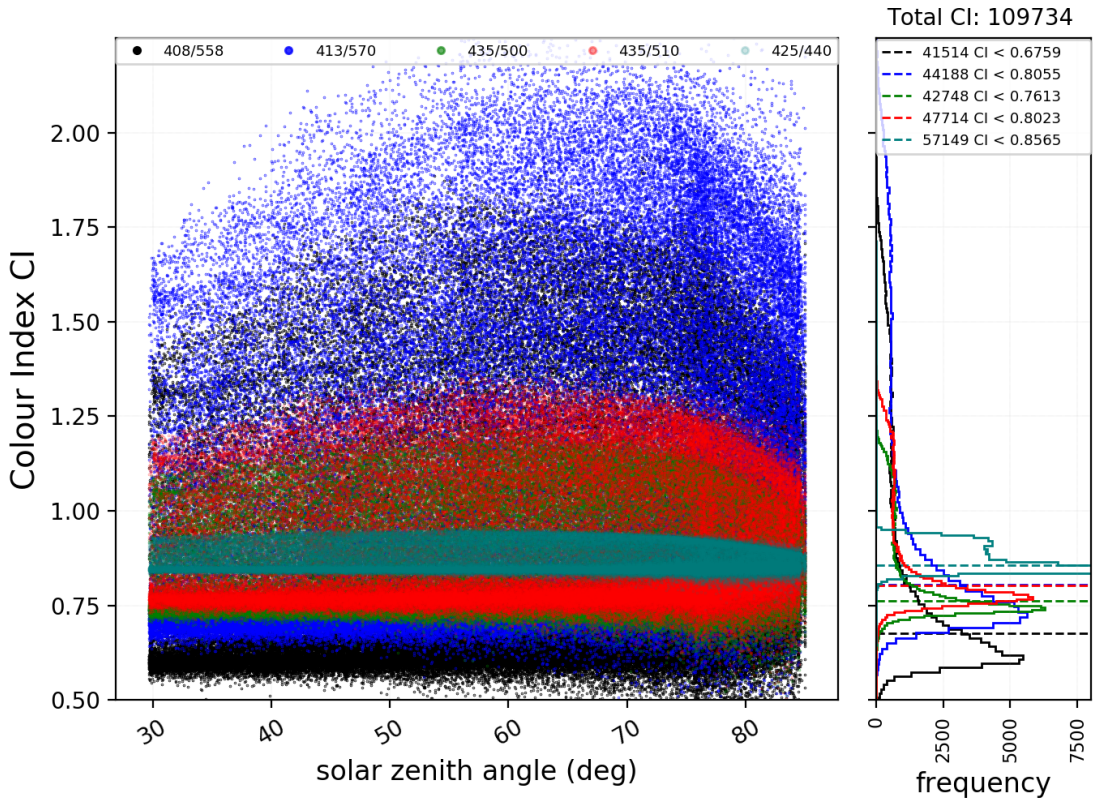


Fig. 3.7: *5 different colour indices for the years from 2015 to 2017 in Bremen. The frequency plot (**left**) shows the defined CI_{limit} as horizontal line calculated as FWHM of the individual cloud peaks of each ratio. The numbers show how many scenes have smaller CI values than the limit, calculated as FWHM of the individual cloud peak (dashed lines).*

The filtering is done by defining the upper full width at half maximum (FWHM) CI value of the cloud peak as the limit CI_{limit} for allowed colour indices (dashed lines). Every value lower than CI_{limit} is filtered. However, not all scenarios discarded in this way are necessarily cloudy. Also high aerosol loads might lead to low CI values. This is of course an unfortunate fact which can be prevented by applying a more complex cloud classification scheme, as introduced by WAGNER et al. [2014].
Nevertheless, here, CI's are only meant to be used for filtering cloudy scenes, which is done by applying the specific CI limit values to all zenith measurements. A scan which uses a zenith background spectrum with a CI lower than CI_{limit} is completely discarded within this thesis.

Another issue to consider is which CI ratio is the best. Five different ratios have been used and differ strongly in their specific frequency features. In general, the farther away both wavelengths of one ratio are from each other, the larger the spread of possible CI ratios. The histogram itself uses a binning factor to create the frequencies. When the

specific ratio depends strongly on this binning factor[11], then it is considered as unstable. This was found to be the case for ratio $425\,\mathrm{nm}$ to $440\,\mathrm{nm}$ which is the official CINDI-2 CI ratio for the visible fit.

In Figure 3.8, a pie chart depicts three different CI ratio comparisons. In each pie chart, blue pies show equally flagged scenes between both CI ratios, whereas red illustrate the unequally flagged CI's values in percent. "≤ 413" are the unequally flagged CI values smaller than the 413/570 limit. Therefore, this pie chart also shows the unequally flagged values larger than the other ratio (435/500). In general, unequally flagged CI's indicate that one or both CI ratios might be not reliable. The first comparison illustrates, that only $\sim 7\%$ of all scenes are not equally flagged by both ratios. On the other hand, in the second comparison, the CINDI-2 ratio (425/440) found 12.54% more scenes to be cloudy. In general, these values are mainly close to CI_{limit} and are therefore not necessarily cloudy. In this sense, CI ratios which leave more scenes unfiltered should be considered as better because they might also include higher aerosol loads or larger SZA measurements. This discrepancy between different CI ratios shows again that some ratios are not as well suited as others. Here, CI filtering is done with the ratio of $413\,\mathrm{nm}/570\,\mathrm{nm}$ because it is a compromise between a larger spread of CI values (see Fig. 3.7) and smaller numbers of filtered scenes. Note that this was mainly an empirical discussion which might be different for other instruments, locations and weather conditions.

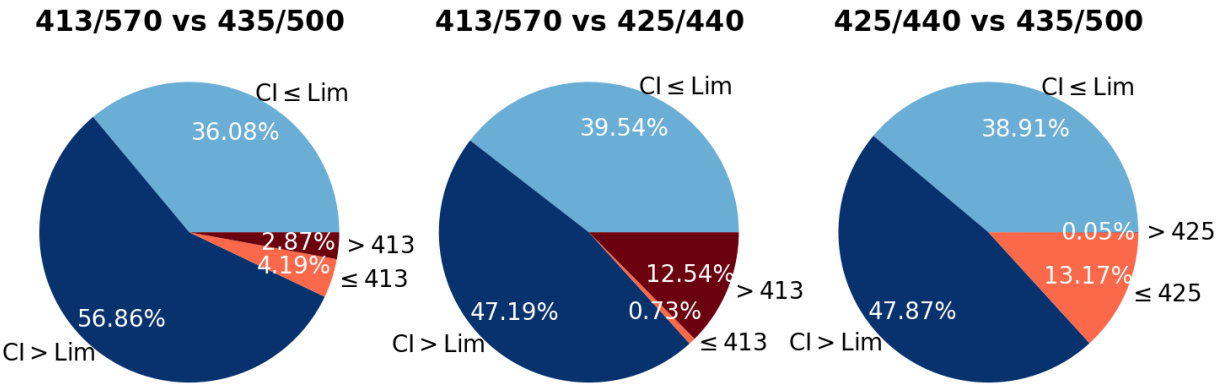


Fig. 3.8: *Pie chart of equally flagged CI in blue and unequally flagged CI in red, for different CI ratios. E.g. ≤ 413 means, that the unequal flagged CI's are cloudy for the ratio 413/570 and therefore not cloudy for the other ratio.*

[11] Here, the binning factor was chosen to be 150.

3.6 Discussion of slant column densities

Before the VC results are presented, a short introduction into NO_2 and HCHO slant column densities is shown in this subsection. In Figure 3.9, the ΔS values for the full time period and all three NO_2 fits are depicted[12].

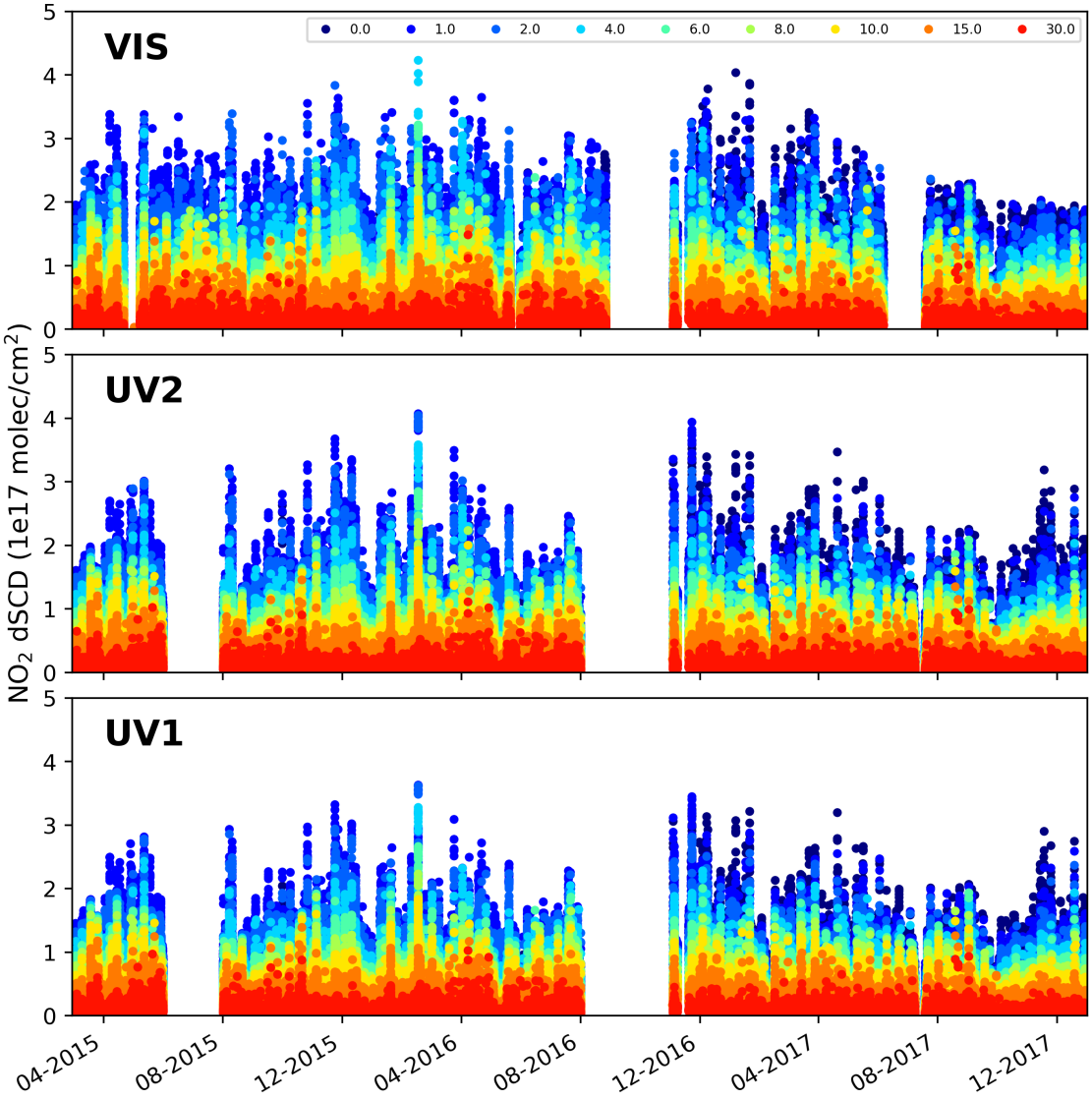


Fig. 3.9: *Time series of NO_2 differential slant column densities during the full period for different LOS, colour-coded (c direction). The three fits are shown in the individual subplots: VIS (**top**), UV2 (**mid**) and UV1 (**bottom**).*

Larger dSCD's are found in the visible fit compared to the UV fits. This is due to the larger effective light path lengths allowing more photons to reach the detector when the atmosphere is less opaque (smaller impact of Rayleigh and Mie scattering at larger wavelengths). Differences in the length of effective light paths will be discussed in Section 3.8

[12] A similar plot for O_4 dSCD can be found in the appendix (see Fig.AF.3)

and can be used to evaluate information about the distance of the specific concentration to the measurement location.
In all fits, larger ΔS occur predominantly in lower elevation angles because larger nitrogen dioxide concentrations are usually closer to their individual emission sources within the PBL and the light path is usually longer. One important feature can be found by comparing the 1° and 30° LOS around 12-2016 and 12-2017 for VIS. The 30° elevation shows similar values but the lower elevation dSCD are much higher around 12-2016 than at 12-2017. This indicates, that depending on the altitude of high NO_2 concentrations and wind direction, VC values derived by the geometric approach from the 30 ° LOS might show less NO_2 than one would find by using lower elevation angles and matching AMF, due to the difference in probed air masses. Of course, the vice versa case is possible, when higher plumes are transported above the field of view of lower elevation angles. However, these events can be considered to be rare. In Figure 3.10, the relative frequencies of NO_2 ΔS are depicted in a polar plot with the angles as a representation of the wind direction. As discussed before, smaller elevation angles show larger dSCD's in all directions whereas the 30° LOS does not see signal at all from e.g. PS_{mid} in north eastern directions (cp. Fig.3.2).

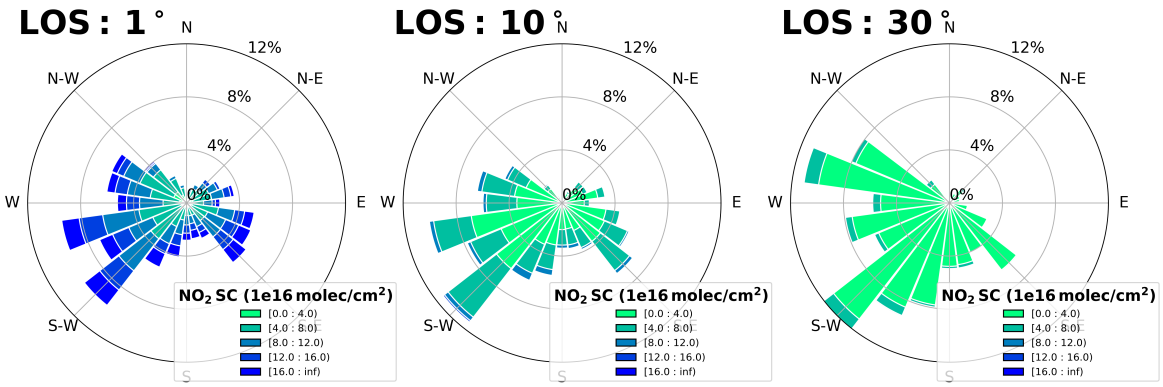


Fig. 3.10: *Windrose plot of NO_2 differential slant column densities during the full periode for three LOS with different dSCD ranges, colour-coded (c direction).*

On the other hand, the highest values for 30° are found in the south west but also in western regions indicating that the responsible emission sources might be PS_{west} and ArcelorMittal as well as traffic emissions in the city centre. Due to this possible angular and spatial deviations of different LOS, results from various approaches like the VC analysis of 30° LOS (Sec. 3.7), onion peeling (Sec. 3.8) or retrieval of vertical profiles (Sec. IV) might differ strongly, depending on the selected viewing azimuth angle. Note, that data shown in this subsection was filtered for the detection limit, RMS and ΔS error only.

In Figure 3.11, dSCD's for HCHO are shown. The seasonal variation of HCHO can be seen, with high values in the summer months and low ΔS during winter. The number of

data points is less than for to NO_2 for various reasons. 1. The HCHO amount is usually small in mid latitudinal rural regions compared to NO_2 so that the HCHO slant column densities are more frequently below the detection limit. 2. Since the signal is smaller, dSCD values might also be filtered due to high fitting RMS or large relative ΔS errors. Still, enough data points are left after filtering allowing a detailed discussion of seasonal variations in later sections. It can already be seen, that the elevation angle dependency of HCHO dSCD is not as pronounced as it was for NO_2 indicating a more homogeneous contribution on a vertical scale.

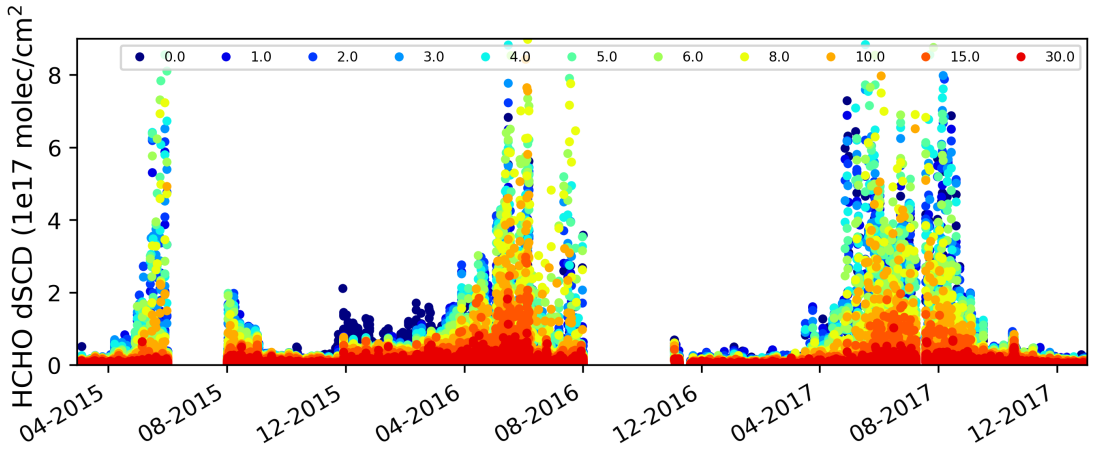


Fig. 3.11: *Time series of* HCHO *differential slant column densities during the full periode for different LOS, colour-coded (c direction).*

3.7 Discussion of vertical column densities

Two different ways to calculate VC from ΔS were already introduced. In this section, a comparison of both approaches and their individual dependencies is presented. In addition, the temporal and lateral variations of VC's in Bremen are shown. The main analysis is based on data from the VIS fit but a comparison of fits in the UV spectral range is introduced in Section 3.7.2 as well.

3.7.1 Comparison of different approaches for the vertical column density calculation

The details of the climatology AMF and the geometrical AMF were already explained in Section 3.5. As a short summary, differences might be due to aerosol parameters, profile shapes, surface albedos, geometry and wavelengths.
In Figure 3.12, ratios of the vertical columns calculated with different approaches and the geometric VC without correction (VC_{geom}) are depicted for direction s and c. In general, a symmetric behaviour for the different seasons of the year can be expected. Larger deviations from this symmetry are questionable and should be considered as problematic.

However, the correction approach should generally lead to more accurate results than the VC values without correction.
The light blue spots are the ratio of CAMS tropospheric columns and VC_{geom} which show a large spread over the full period. This could indicate two things: CAMS does not consider individual emissions of e.g. traffic and smaller power plants which can have a big impact on the measurement especially when gases accumulate at lower wind speeds, and VC_{geom} values are not a proper estimate of the true vertical column density. The latter point shows the need for a correction of geometric VC which is done here in two different ways.
First, the green curve shows the ratio of corrected geometric VC without aerosols (VC_{geom} corrected) and the uncorrected VC (VC_{geom}). Second, the red dots illustrate the difference when using a correction factor including aerosols (VC_{geom} corrected + Aerosols / VC_{geom}). The relative difference between uncorrected and corrected VC calculated with AMF's within a pure Rayleigh atmosphere only (green) goes up to $\sim \pm 20\%$ during summer (direction c), with similar high overestimations for direction s but less frequent underestimations of VC_{geom}, since not so many values are lower than 1. The difference between s and c should be due to the specific measurement geometries only. Especially during noon, RAA and SZA are small for the southern direction whereas c means measurements under small RAA only in the late afternoon. On average, no large differences in the actual amount of NO_2 measured in s and c direction are to be expected since both viewing directions point into more or less highly polluted areas (see also the discussion in Sec. 3.7.3 and the angle dependency plot in the appendix, Fig. AF.6).

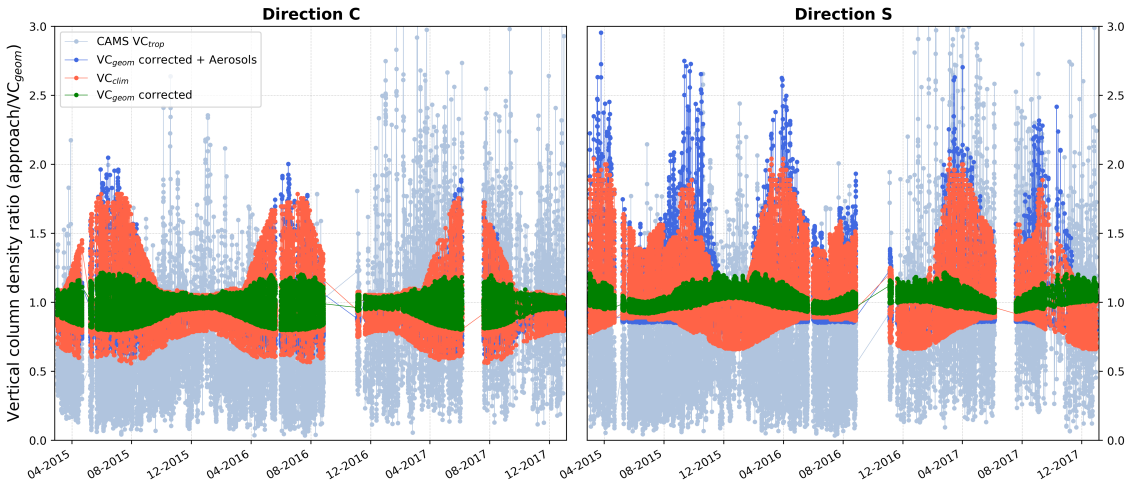


Fig. 3.12: *Comparison plot of the different VC calculation approaches.*

Comparing the results of corrections without (green dots) and with aerosols (blue dots) for direction c it is obvious that aerosols have a large impact on the measurement and VC_{geom}. The ratio goes up to values of 2-3 indicating that especially in the summer months, the geometric air mass factor is not a reliable quantity for calculating VC. The reason for

this can be separated into several contributions. First, the AOD in Bremen is larger in the summer months than in the winter months leading to a larger impact on NO_2 measurements (see Sec. 3.7.3). Second, the PBL is on average higher in summer enabling an ascending (uplifting) of trace gases to higher altitudes which might lead to smaller or larger slant columns depending on various factors (wind, distance to emission source, different height depending effects of aerosols and trace gases due to shielding and enhancements). Third, the individual scattering geometries have an impact since the SZA is on average smaller in summer than in winter, changing the effective length of light paths. Along with the last point, the aerosol phase function P_{HG} also leads to overestimations of geometric VC by deteriorating the results of the forward (and backward) scattering geometries with small relative azimuth angles.

Nevertheless, the correction leads to similar results when comparing the ratio of corrected VC (blue dots) and the climatological approach (red dots) for direction c. Only slightly larger ratios during summer for the correction with aerosols compared to the climatological VC can be identified. Even though the profile shapes between both approaches differ (box profile for geometric correction and exponential profile for the climatological approach) the final VC values agree well indicating that the climatological change of scale height for exponential profiles (VC_{clim}) produces similar results as the PBL, AOD and VC_{trop} depending selection of correcting factors.

Comparing the red and blue datasets for direction s, larger differences due to strong overestimations of the corrected geometric VC including aerosols can be identified, especially in the summer months. From the above mentioned reasons, a geometry problem seems to be most likely since direction c is not affected as strong as the southern direction s. Here, the used phase function seems to cause problems because especially for lower RAA (around noon), overestimations are found (compare also Fig. AF.6 in the appendix). Unfortunately, within the city of Bremen, no phase function measurements are available preventing the use of more accurate aerosol angular scattering distribution. In addition to the LUT calculation with P_{HG}, in Figure AF.5 in the appendix another dataset is shown with corrected VC calculated including a measured phase function from the MAD-CAT campaign in Mainz, Germany, in summer 2013[13] (PETERS et al., 2017). Although the used phase function can not be understood as a good representation of the aerosol angular scattering distribution in Bremen throughout the full time period, its use clearly shows an improvement as several overestimated values are strongly decreased. Remaining high ratios, especially for large scattering angles (compare Fig. AF.6), illustrate again the importance of an accurate phase function and a possible temporal variability of aerosol types throughout the year.

Here, as already pointed out, this problem can not finally be solved because of a lack in

[13] The aerosol phase function was measured by an AERONET station on 18.06.2013 and was kindly provided by Thomas Wagner and Meinrat O. Andreae (MPIC, Mainz, Germany). See MAD-CAT [2013] for further information on the campaign.

phase function measurements in Bremen and due to the computation time which is necessary when calculating LUTs for different phase functions. For the following discussion of vertical column densities in Bremen, the corrected VC values including P_{HG} are used and referred to as geometric VC (the type of correction is suppressed from now on) but RAA values smaller than $20°$ are removed instead of using possibly wrong correction factors and vertical column densities.

3.7.2 Comparison of vertical column densities retrieved from different fitting windows

According to Equation 2.53, the NO_2 vertical column density calculated with all three fit settings should be the same because it does not depend on the wavelength. In reality, due to uncertainties in the measurement, the DOAS fit and the conversion to vertical columns, VC values might differ. Especially when important spectral (interfering species in one fitting window), temporal (not important when routine daily measurements are available) or aerosol features (wrong wavelength dependency for the specific aerosol type or parametrization) were not covered during the conversion, deviations can be found.

In Figure 3.13, the time series and the scatter plots including the frequency distribution of VC_{clim}[14] for all three fitting windows and for directions c and s are depicted (direction b in the appendix, Fig. AF.7). Due to small differences in the temporal coverage, the datasets had to be resampled on 10 minute values without allowing extrapolation.

In general, the agreement of all three fitting windows is good with correlations larger than 0.82 and regression line slopes close to 1 with relative deviations of 5% and 7% (UV2, UV1) for direction c and 7% and 10% (UV2, UV1) for direction s. In the time series it is obvious that the visible data points tend more often to high values. The fitting window UV1 has usually smaller VC values. This can also be observed by comparing the slopes of the regression lines. Here, UV1 has for all directions the smallest slope. This indicates, that the conversion was not fully successful but the uncertainties are still small. Table 3.5 summarizes the mean and median relative deviation between each fitting window VC and the mean of the specific vertical column densities for each time step, for all directions (top rows). In addition, the fitting parameters of a skewed Gaussian distribution fitted to each frequency distribution within the histograms is shown in the bottom rows.
A table showing theses results without RAA filtered data can be found in the appendix in Table AT.1.
The comparison of the relative deviations (mean and median) show that UV2 has the largest deviation for all directions, and VIS the smallest. The deviations are always negative

[14] Unfortunately, VC_{geom} can not be used, because the correction LUT was only calculated for the visible spectral range. However, the comparison of VC_{geom} should be similar to this subsection.

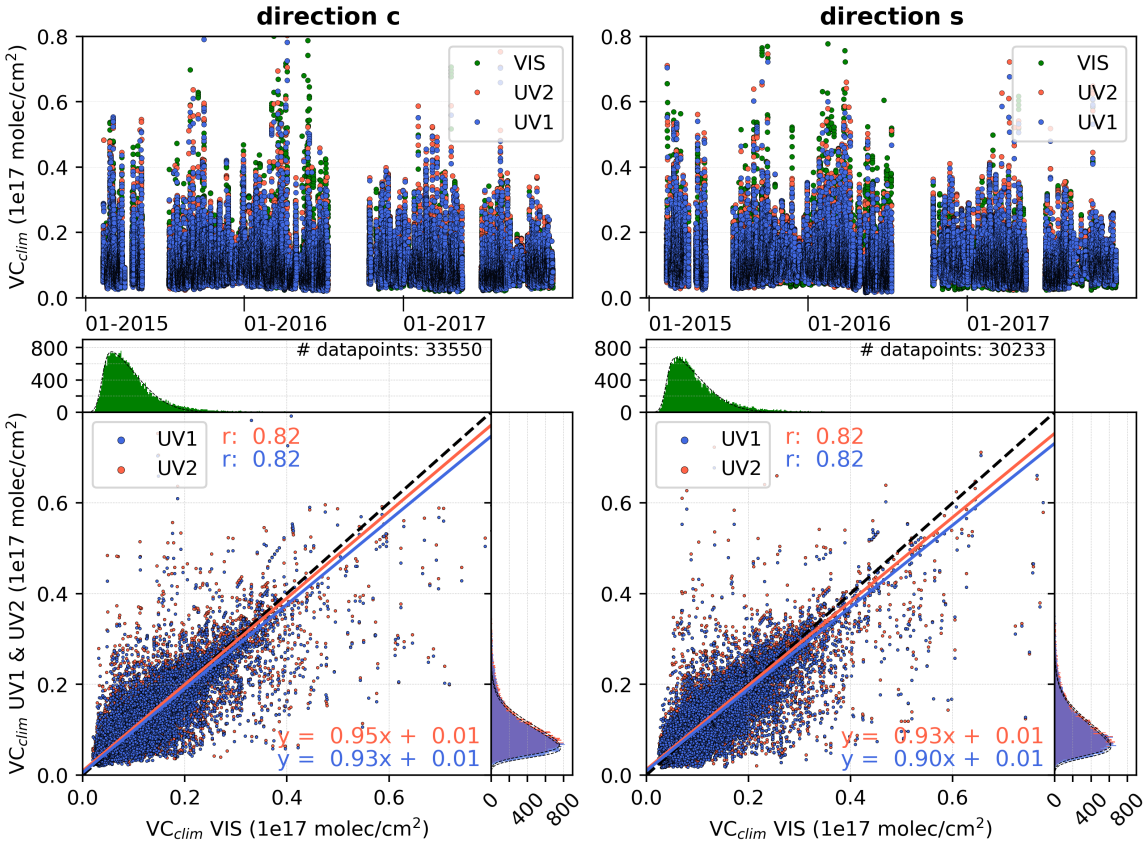


Fig. 3.13: *Comparison plots of the different fitting windows used for the calculation of VC_{clim} for the directions c and s.* ***Top:*** *Time series,* ***Bottom:*** *Correlation plots of visible and UV fitting windows including histograms depicting the frequency distribution of each data set. In addition, Pearson's correlation coefficient r as well as the parameter of the regression lines are given, colour-coded.*

for UV1 and positive for UV2. These findings indicate, that the visible fit is close to the mean and UV1 and UV2 are with a maximal average uncertainty of -3.6% and 4.7% smaller or larger than the mean value.

The fitting parameters of the Gaussian distributions enable more detailed analysis. Here, the absolute mean value μ (the centre point of the Gaussian) is the largest for UV2 and nearly similar for UV1 and VIS. This is in agreement with the discussion of the relative deviations above, which indicates that UV2 might overestimate the (true) vertical column density. The variance σ is a representation of the spread of each VC distribution, and it is marginally larger in the visible than in the UV for direction c but smaller for direction s. This deviation in σ for c and s indicates that lateral differences might play a role, i.e. when direction c could be affected by power plants, and direction s might be dominated by traffic emissions. The most interesting feature is the skewness γ, which shows how much the Gaussian distribution is tilted towards higher or lower values from the centre point (mean). Positive values show a tilting towards smaller VC with the largest skewness

for all directions in the visible fitting window. The smallest skewness can be found for direction c and UV2. Since a larger γ can also occur if high VC values affect the Gaussian distribution, VIS and UV1 tend to occasionally high VC. On the other hand, UV2 shows a more bell curve shape which shows a lack in high values and a possible problem for the attempt of separating individual NO_2 maxima from the temporal VC background.
The reader should note, that this was rather a relative discussion as the absolute truth is unknown. The analysis in the later sections should bring more light into the temporal and spatial variations of the NO_2 distribution in Bremen.

	direction c			**direction b**			**direction s**		
	UV1	UV2	VIS	UV1	UV2	VIS	UV1	UV2	VIS
Median (%)	-3.6	4.5	-0.8	-3.5	4.3	-0.8	-2.3	4.7	-3.1
Mean (%)	-3.2	4.2	-1.0	-3.3	3.8	-0.5	-2.0	3.7	-1.4
μ $(1e15\frac{molec}{cm^2})$	3.8	4.4	3.7	3.8	4.5	3.9	3.8	4.3	3.9
σ $(1e15\frac{molec}{cm^2})$	6.1	6.1	6.4	6.3	6.3	6.2	6.9	6.9	6.4
γ	5.7	5.2	7.1	6.3	5.3	6.9	5.9	5.5	7.4

Tab. 3.5: *Mean and median relative deviations of all fitting windows (UV1, UV2, VIS) to the mean value of the specific VC for all three fits (2 top rows). Fitting parameters (μ: mean, σ: variance, γ: skewness) of skewed Gaussian distributions fitted to the histogram of data points (3 bottom rows). RAA $\leq 20°$ were filtered.*

3.7.3 Temporal and azimuthal variations of vertical column densities in Bremen

In contrast to the discussion in the previous section, the corrected geometric vertical column density VC_{geom} is discussed here for NO_2 in the visible fitting window only. The differences between the individual fitting windows were found to be small so that the temporal and spatial discussions should be similar for all fitting windows. In addition, HCHO VC data converted with the climatological approach is analysed as well.

Vertical column densities of nitrogen dioxide

In Figure 3.14, the monthly mean values of VC for the south direction s and both westerly directions b and c are depicted, together with three different NO_2 concentration time series of BLUES in situ stations close to the measurement direction vectors (see Bremen map, Fig. 3.2). An annual cycle can be identified with large values during winter and lower vertical columns in the summer months. This temporal variation agrees well with the in situ stations indicating that emission sources are within or close to the city and

that they are more or less equally distributed. Since all three azimuthal viewing directions show similar patterns, single emitters such as power plants can not be considered as distinguishable, dominant emission source. Note that the relative deviation (not shown) between direction s and c is on average around 5% (20% on maximum on April 2015) but does not show a systematic pattern. Generally, the observed increase of NO_2 in the winter months is caused by three different reasons. First, a longer lifetime of nitrogen dioxide due to a decreased photolysis rate is expected in winter. In addition to this photochemical loss, NO_2 is also reduced due to reactions with OH and O_3 in the summer months and the loss of NO_2 is therefore not as pronounced in winter. As a third reason, a general smaller dispersion efficiency in winter leads to higher near-surface concentrations of NO_2 due to shallower PBL because of frequent thermal inversions in the lower troposphere (see also the time series of relevant quantities, Fig. AF.8). Similar seasonal cycles were also found

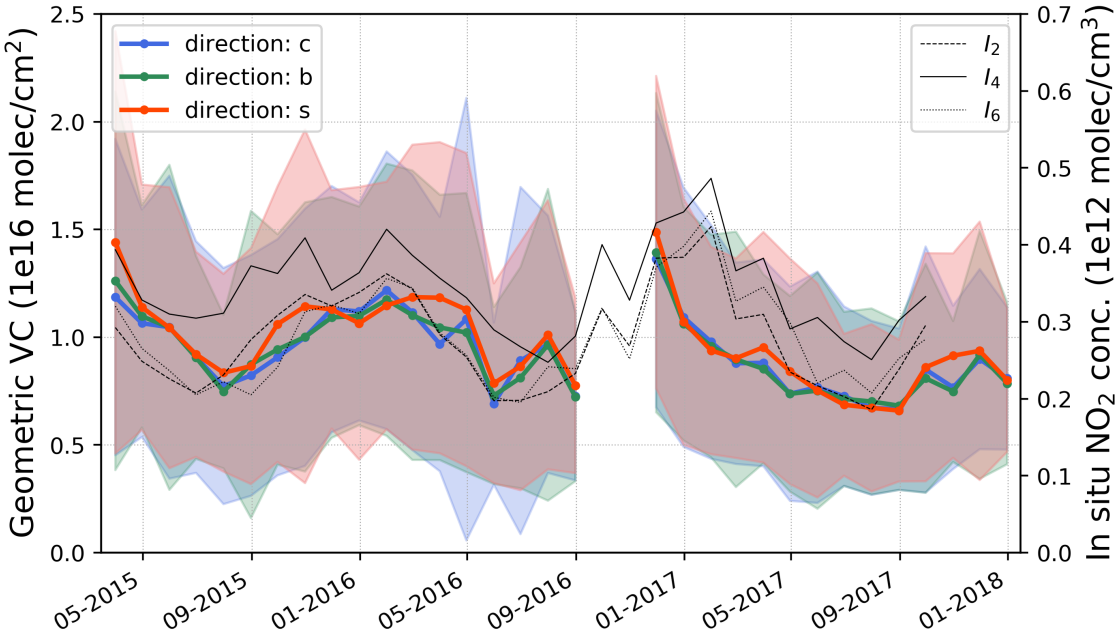


Fig. 3.14: *Time series of monthly mean NO_2 vertical column densities during the full period for the directions s (red) and c (blue). In addition, three different in situ time series of monthly mean NO_2 concentrations for the stations I_2, I_4 and I_6 are depicted. The shaded areas show the standard deviation of the vertical column monthly mean values.*

by MENDOLIA et al. [2013]; KRAMER et al. [2008] in Canada and England as well as in Asia by e.g. MA et al. [2013]; CHAN et al. [2015]; TIAN et al. [2018]. Although the cycles agree well with other studies of urban regions, the standard deviations are higher due to large monthly and daily variations. Note that these uncertainties are slightly smaller within the summer compared to the winter months which is due to the smaller number of measurements rather than a lesser spread of VC.

Figure 3.15 shows the daily variation of vertical column densities for the direction c (left) and direction s (right) averaged for each weekday (top), for each month (mid) and for the four seasons (bottom; direction b is depicted in the appendix, Fig. AF.9). Note that in contrast to the other plots in this section, the central European time (CET) was used for the x-axis because the changes of summer and wintertime smears the daily variational pattern too much.

For direction c, all plots show high values in the morning and in the afternoon, indicating a strong correlation with commuter traffic emissions. The top graph confirms this viewpoint as the weekend days show smaller values than the weekdays. It is interesting to notice that on Monday, the vertical column densities are slightly higher than on the weekend but not as large as on Tuesday, the day with the largest VC (see also daily means in the appendix, Fig AF.10). This behaviour was also found for several in situ sites (see e.g. I_2 concentrations in Fig. 3.16) and can be explained by two different effects. First, the NO_2 background signal, which increases throughout a usual week due to an accumulation of trace gases when the wind speed is low, is smaller at the beginning of the week. Second, changes in traffic emissions might play a role (lower emissions by power plants are not to be expected). Especially the weekend driving ban in Germany for larger lorries (with mass $\geq 3.5\,\mathrm{t}$) leads to lower numbers of vehicles[15] entering (or leaving) the area of Bremen on early Mondays as illustrated in Figure 3.17. Furthermore, all daily average curves peak around 6 PM except on Monday and Friday. Particularly on Fridays, this behaviour makes sense as jobholders tend to go home earlier. The reason for a temporal shift of the traffic peak on Mondays is unclear.

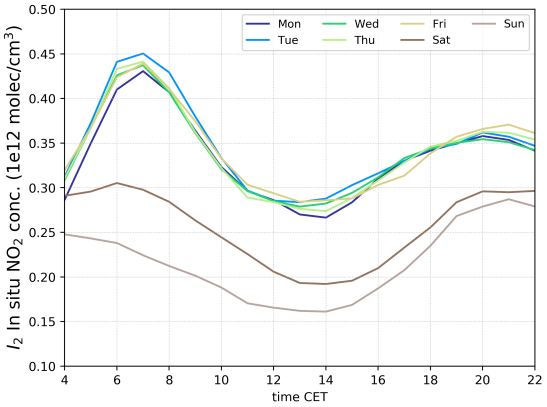
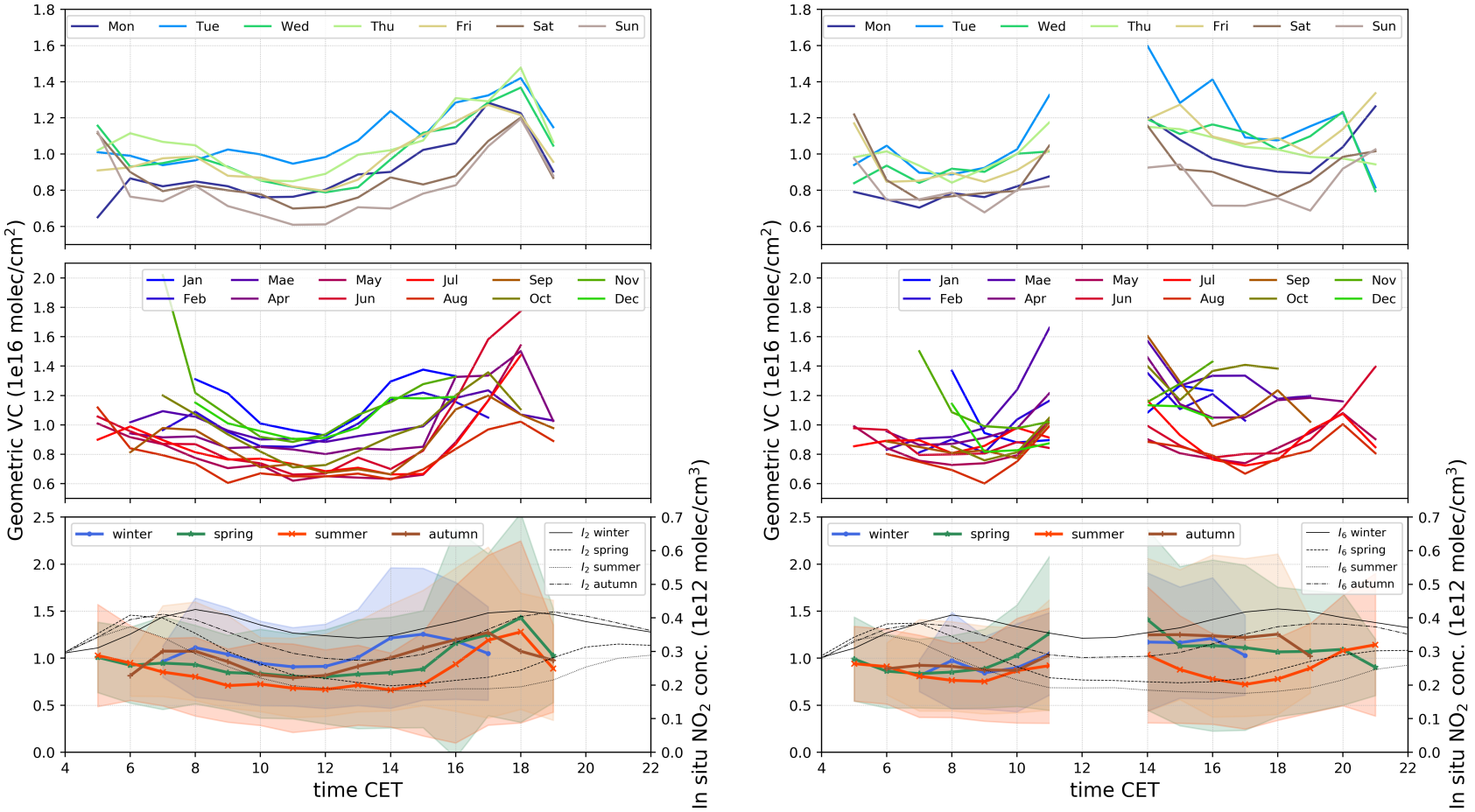


Fig. 3.16: *Daily variation of in situ average NO_2 concentrations for each day of the week, for site I_2.*

In Figure AF.11 in the appendix, the daily variations per weekday are split up into the four seasons of the year for direction c. The overall tendency for smaller values on Mondays can be found through all seasons but is more pronounced in spring and summer than in autumn and winter. Note that in general, the traffic emission cannot be clearly identified in the daily variation plots. Figure 3.17 shows a clear morning and afternoon peak on working days but a noon peak for Saturday and Sunday. These noon peaks can only be identified in the summer season. All other seasons show enhanced values in the morning

[15] The total number consists of the addition of datasets of 6 different traffic-counting-sites located on the major roads and highways in Bremen.



(a) *Direction C* ***(b)*** *Direction S*

Fig. 3.15: *Vertical column densities of NO_2 for direction c (left) and s (right) as daily variation for,* ***top****: each day of the week,* ***mid****: each month and,* ***bottom****: each season, averaged over all three years. Additionally, the bottom plots show in situ average NO_2 concentrations for site I_2. Winter: Dec, Jan, Feb; Spring: Mar, Apr, May; Summer: Jun, Jul, Aug; Autumn: Sep, Oct,Nov.*

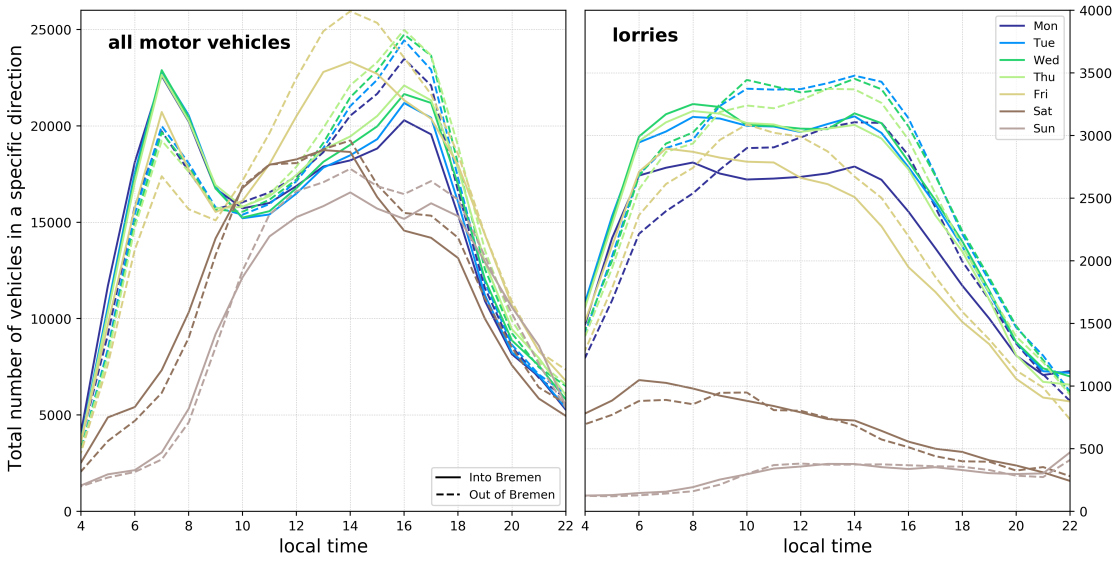


Fig. 3.17: *Daily variation of the total number of vehicles entering (solid) or leaving (dashed line) Bremen for each day of the week. Data provided by Bundesanstalt für Straßenwesen: Automatische Zählstelle 2015 (*FITSCHEN and NORDMANN, 2015*)*[15].

and afternoon but no strong distinguishable peaks which might indicate dominant traffic emissions. This was to be expected because the 30° LOS is not as suitable as lower elevation angles for the separation of near surface sources. Although several other studies found smaller values on Mondays as well (BEIRLE et al., 2003; MA et al., 2013; TIAN et al., 2018), the day with the maximum value differs, i.e. it was found to be Wednesday by MA et al. [2013]; TIAN et al. [2018].

The mid plots in Figure 3.15 show the VC daily variation for each month on average and confirm the results found in the VC time series plot before (Fig. 3.14). In general, the average values are higher in the winter months compared to summer for reasons explained above. August and July show the lowest values and, for August, the lowest daily variation cycle. The latter point can be explained by the summer holidays in the states of Bremen and Niedersachsen, usually in between July and August, which decrease the amount of commuter traffic drastically in addition to lower heating and energy production by power plants. The difference between winter and summer months can also be seen by comparing these results with the seasonal averages in the bottom plots. Here, the shaded areas indicate the standard deviation which was neglected for the other subplots for reasons of clarity. However, the highest peaks in the late afternoon traffic emissions can be found in spring but these results have larger standard deviations due to the relative azimuth filtering in the late afternoon for the westerly viewing directions. The highest values can be found around 6 PM for spring and summer, at 5 PM for autumn and at 3 PM for winter which might be due to enhanced energy production and in turn increased power plant emissions due to the earlier sunset in autumn and winter.

In contrast to the c direction, the southerly viewing direction shows a dominant noon peak which might be due the phase function issue discussed in the previous subsection. Even though, RAA $\leq \pm 20^{\circ}$ and the time between 11 AM and 2 PM was filtered, a clear tendency to higher values before and after this time period can be identified. Since this was not found for in situ station I_4, I_5, or I_6, this noon peak is considered to be an artefact. Besides this artefact, all curves are again dominated by morning and late afternoon enhancements. Here, a maximum in spring is not as pronounced as for direction c. For both directions, the nearby in situ stations show more pronounced morning peaks which can be explained by the different quantities (concentration vs. integrated column), the low PBL's in the morning, the high elevation angle for the VC calculation (leading to different proved air masses) and the distance to the sources and the dispersion time for the their individually pollutants. Direction b in the appendix is more dominated by RAA linked problems as direction c but the general daily variation for different weekdays, months and seasons agrees well.

Vertical column densities of formaldehyde

In contrast to NO_2, no correction LUT was calculated for $HCHO$. Here, the vertical column densities are discussed which were retrieved by applying the climatological AMF from the QA4ECV project (compare Sec. 3.5).

Figure 3.18 shows the time series for $HCHO$ vertical column densities, for all three directions. A seasonal cycle can be observed with high values in summer and lower VC in winter. Due to gaps, this cycle is more pronounced in 2017 but it is imaginable in 2015. The year 2016 seems to stand out with a clear minimum in spring instead of the steady increase observed in 2017. In general, $HCHO$ depends strongly on biogenic and anthropogenic emissions as it is an oxidation product of VOCs. The main sinks are photolysis, reactions with OH and wet deposition. However, the reason for this anomalous behaviour in 2016 can not be fully explained. In this time of the year, precipitation might lead to lower values but no enhanced or longer rainfalls were detected, in comparison to the other years.
The relative difference between the individual directions is in 2017 at maximum 25% which is similar to the results for NO_2. The $HCHO$ VC is in this year in summer one order of magnitude larger than in winter. Since the agreement between the different azimuthal directions is good (except in 2016), a homogeneous distribution of $HCHO$ in Bremen can be assumed with dominant biogenic emissions rather than anthropogenic. The observed seasonal cycle in Bremen with higher values in summer was also found in other studies e.g. in Asia (WANG et al., 2017; KHAN et al., 2018; TIAN et al., 2018).

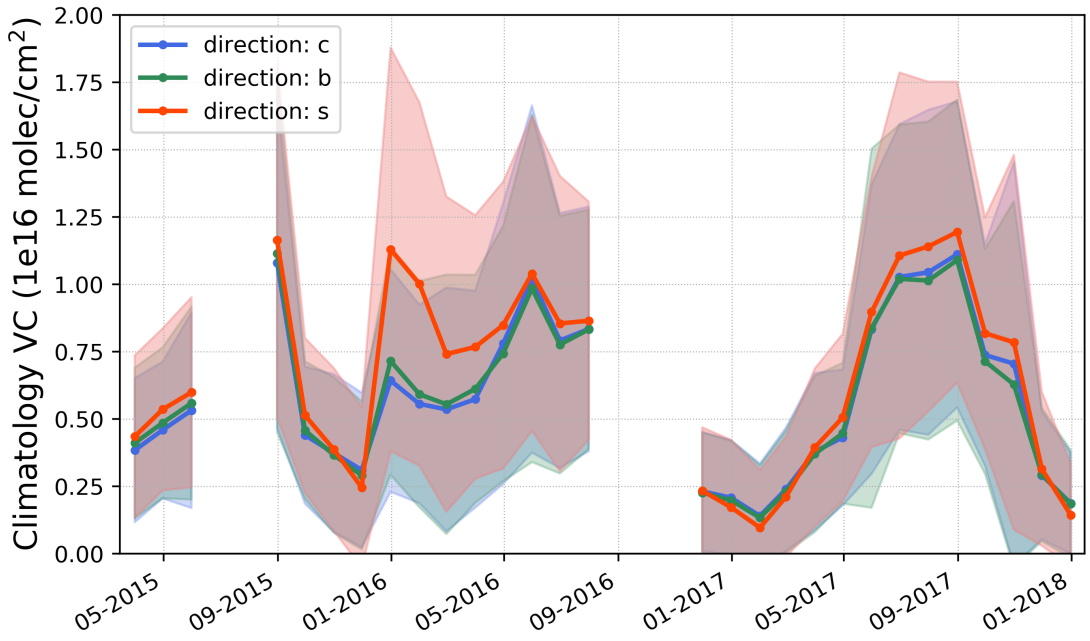
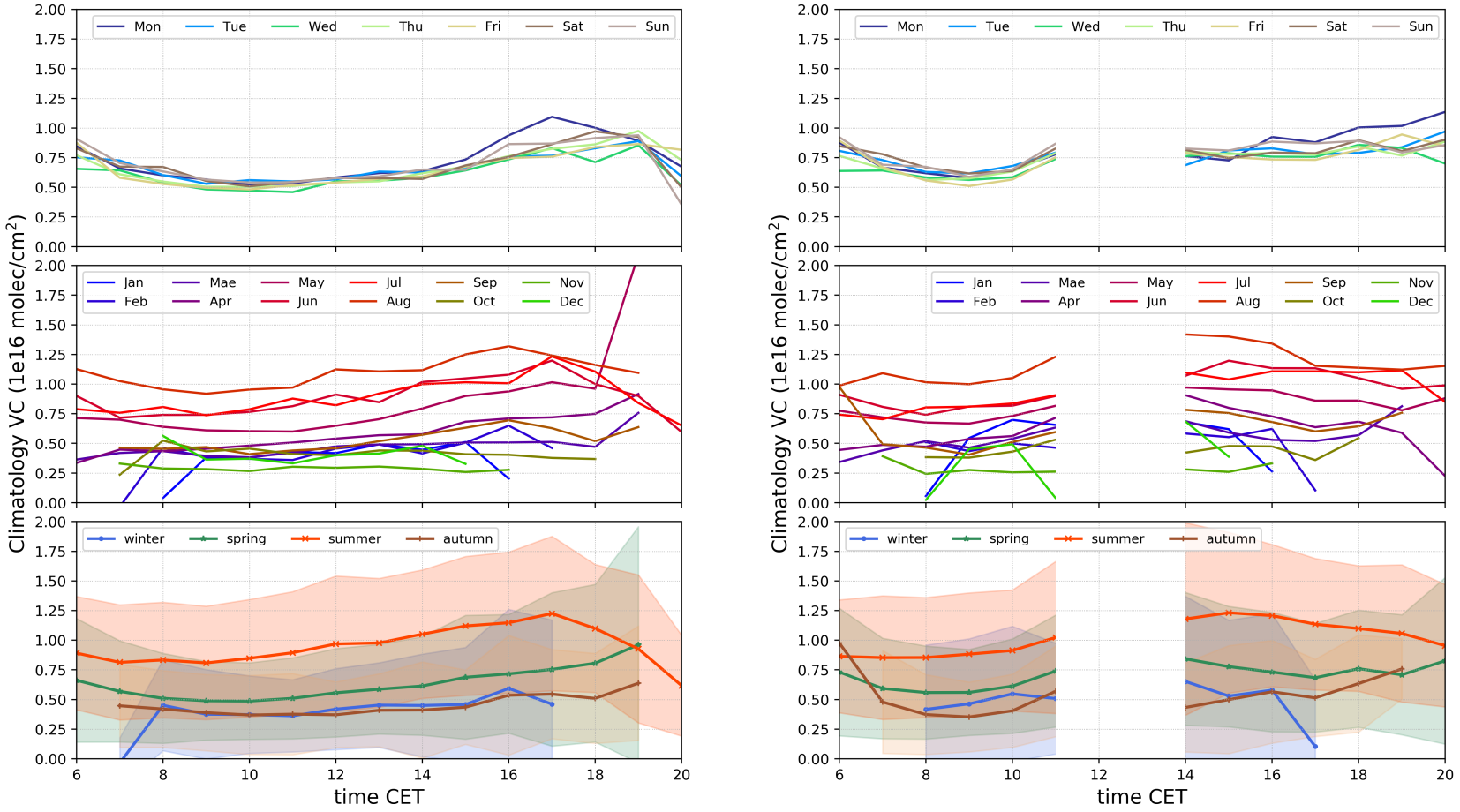


Fig. 3.18: *Time series of monthly mean* HCHO *vertical column densities during the full period for the directions s (red) and c (blue). The shaded areas show the standard deviation of the vertical column monthly mean values.*

Figure 3.19 shows the daily variation of vertical column densities of HCHO for the direction c (left) and direction s (right) averaged for each weekday (top), for every month (mid) and for the four seasons (bottom). In contrast to NO_2, the differences between the days of the week are smaller. This confirms the finding of the time series discussion above, that anthropogenic emissions are a minor contribution to HCHO in Bremen. The agreement of the different azimuthal viewing directions is good as well indicating a lateral homogeneity (direction b in the appendix, Fig AF.12). The only differences can be found for azimuthal viewing directions close to the sun (afternoon for direction b and c and noon for direction s).

In general, the daily variation shows large values in the morning and in the late afternoon with a decrease in the evening. This agrees well with studies of cities in Europe by Pinardi et al. [2013] and Gratsea et al. [2016] (the latter one for ΔS only, in winter) but not with Asian cities which show mostly a dominant noon and early afternoon enhancement (Wang et al., 2017; Tian et al., 2018). High values around noon are usually expected when biogenic emissions of hydrocarbons are dominant which are consequently oxidized to HCHO. On the other hand, morning and afternoon peaks usually indicate anthropogenic emissions. Here, neither noon nor morning peaks are dominant so that a mixture of biogenic and anthropogenic emissions is assumed. Note that even though clear maxima can be observed around 5 PM, morning peaks, which might have indicated traffic emissions, are missing. Another important point is that daily variations of the day of the



(a) *Direction C*

(b) *Direction S*

Fig. 3.19: *Vertical column densities of* HCHO *for direction c (left) and s (right) as daily variation for,* ***top****: each day of the week,* ***mid****: each month and,* ***bottom****: each season, averaged over all three years. Additionally, the bottom plots show in situ average* HCHO *concentrations for site* I_2*. Winter: Dec, Jan, Feb; Spring: Mar, Apr, May; Summer: Jun, Jul, Aug; Autumn: Sep, Oct,Nov.*

week curves are generally smaller than the standard deviations (not shown). However, since all azimuthal directions show a tendency of rising HCHO VC in the late afternoon, enhancements due to traffic emissions are considered as an additional source on top of the biogenic background signal (see also JUNKERMANN, 2009 for a discussion of a mixture of sources).

Figure AF.14 in the appendix shows the daily variation of the different days of the week for the four seasons of the year (direction c). Again, no clear indication for differences between the individual days can be observed (see also Fig. AF.13 in the appendix), Though, late afternoon peaks are present, especially in summer with a rapid decrease afterwards. This is in a good agreement with the monthly or seasonal averaged daily variations in Figure 3.19. Note that larger standard deviations for small RAA deteriorate the results in the evening for directions b and c (and around noon for direction s). The rapid reduction in the evening might also be an artefact of an inaccuracy in the conversion of vertical column densities due to the accompanied lack of proper aerosol information.

3.8 Horizontal separation of NO_2 with the help of onion peeling

The comparison of different viewing directions in the above discussion of vertical column densities does usually not enable the isolation of specific emitters like power plants. It was shown before that even the separation of traffic emissions from point sources is a difficult task due to the integration over long light paths and different air masses when pointing into an elevated LOS. More details on surface near emitters can be found when using lower elevation angles in addition to the analysis of different wavelengths as the effective light path length depends strongly on the wavelength (see also SC discussion in Sec. 3.6). Therefore, NO_2 VC retrieved in multiple fitting windows show a varying sensitivity to different distances to the measurement location but the final quantity is still an integrated value over the altitude under the assumption of a laterally homogeneously distributed trace gas concentration.

A simple approach for the discussion of concentrations (or VMR) instead of integrated quantities was introduced by SINREICH et al. [2013] and is shortly explained as follows. It was found that in a homogeneously distributed box-like near-surface trace gas layer, the scattering point (considering single scattering only) is always within this layer in similar (slant) distances for small and nearby elevation angles, when the aerosol load within this layer is high. Furthermore, since the scattering point for a zenith spectrum is also supposed to be close to or within this aerosol layer, the corresponding absorption signals from the

sun to the scattering points are nearly identical for the zenith and the lower elevation spectra. In this specific case, the light path length of different low elevation angle spectra are similar inside this box so that an effective light path L_{eff} can be calculated by

$$L_{eff} = \frac{\Delta S_{O_4}}{c_{O_4}} \cdot f_c, \tag{3.12}$$

with the O_4 concentration c_{O_4} at the instrument's altitude level. With this effective light path, the surface near trace gas concentration profile of e.g. NO_2 can be calculated as follows:

$$c_{NO_2} = \frac{\Delta S_{NO_2}}{L_{eff}} \frac{1}{f_c} \tag{3.13}$$

In both equations, f_c is a correction factor introduced by SINREICH et al., 2013 in order to compensate for the different trace gas profile shapes of O_4 and NO_2 and their wavelength dependencies:

$$f_c = \frac{\Delta M_{NO_2} \cdot PBL \cdot c_{O_4}}{\Delta M_{O_4} \cdot V_{O_4}} \tag{3.14}$$

In this equation, the planetary boundary layer height (PBL) was used and multiplied with the O_4 concentration in order to calculate a partial vertical column within which the trace gas and aerosol are assumed to be homogeneously distributed. V_{O_4} is the total O_4 vertical column density. The differential air mass factors for NO_2 and O_4 are evaluated with respect to the measurement geometry, wavelength, albedo and trace gas as well as aerosol profile (considered as box).

ORTEGA et al. [2015] used this approach at three wavelengths meaning the calculation of three different NO_2 concentrations with matching effective light paths. These three datasets are used within an *onion peeling* approach in order to retrieve distance depending trace gas concentrations. Figure 3.20 shows a schematic representation of the different quantities which are important for this approach. With Equation 3.12, the differential effective path lengths are calculated by using the difference of the individual effective paths lengths (e.g. $\Delta L_3 = L_{VIS} - L_{UV2}$)[16] for the three different fitting windows UV1, UV2 and VIS respectively (cf. Tab. 3.3). The differential concentrations are then calculated by

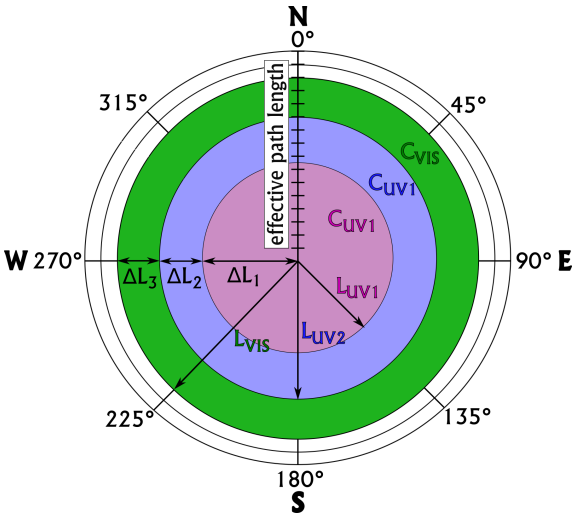


Fig. 3.20: *Schematic representation of the onion peeling approach including the fitting windows UV1, UV2 and VIS.*

[16] As a simplification, the subscripted "eff" is neglected.

using Equation 3.13 for different fitting windows so that the representing formula for the calculation of the outer ring (green) is:

$$\Delta c_{VIS,UV2} = \frac{c_{VIS} \cdot L_{VIS} - c_{UV2} \cdot L_{UV2}}{L_{VIS} - L_{UV2}} = \frac{\Delta S_{VIS} - \Delta S_{UV2}}{\Delta L_3} \tag{3.15}$$

with the NO_2 differential slant column densities ΔS_{VIS} and ΔS_{UV2}. The in-between ring is calculated respectively by using the UV2 terms for VIS and UV1 terms for UV2. The inner circle concentration is directly calculated with Equation 3.12 and 3.13 so that there is actually no differential concentration Δc_{UV1} but the UV1 concentration c_{UV1} only.
The calculation of near-surface concentrations (or VMR) was developed by SINREICH et al. [2013], the sensitivity further quantified by WANG et al. [2014] and the MAX-DOAS onion peeling approach introduced by ORTEGA et al. [2015]. All three publications showed tests on the sensitivity and the performance of this approach. A short summary of results which are important for the following discussion can be found below. The abbreviations S, O and W in brackets represent the publications of SINREICH et al. [2013], ORTEGA et al. [2015] and WANG et al. [2014] as sources, respectively:

- The ΔS_{O_4} values for the lowest elevations angles should be similar. Then a box profile can be assumed (S).
- The aerosol load must be high enough (AOD of 0.3 or higher) (S).
- When the aerosol load is high enough, the approach does not depend on the actual aerosol layer height or the specific AOD (S).
- However, when the AOD is high but the aerosol profile is wrongly considered for f_c, the uncertainties are even larger than for lower AOD's (W).
- The trace gas layer height and the RAA have a strong impact, when RAA and PBL are low (S).
- The approach leads to variable VMR and, therefore, unstable results for PBL $\leq 500\,\mathrm{m}$ when the true PBL is not accurately known (S).
- Due to the general sensitivity for the lowest layers, the approach is not sensitive to layer heights in high PBLs (S).
- When the mixing layer height is an accurate expression of the true trace gas layer, then uncertainties of the RTM calculations are reduced (in contrast to an inaccurate estimate of the PBL) (O) .
- The uncertainties are smaller for lower elevation angles because they probe altitudes where the profile shapes of aerosols and trace gases do not differ much (W).
- For an elevation angle of $1°$, the probed altitude range was found to be smaller than $200\,\mathrm{m}$ (W).
- The correction factor f_c has high uncertainties for RAA $\leq 50°$ and SZA $\geq 70°$ (W).
- The uncertainty of the VMR was found to be $30\,\%$ or smaller (S, W).

3.8.1 Discussion of the correction factor f_c

All three mentioned publications tested the retrieval sensitivity and the related dependencies on the individual parameters with synthetic data and showed only short analyses of real measurements. Here, an empirical discussion is presented which does not focus on specific scenarios or geometries but represents real measurements and conditions of the three years of Bremen data which were already introduced. First, in advance of the later case study, the importance of the correction factor f_c (Eq. 3.14) is highlighted. The differential air mass factors depend on the geometry (SZA, RAA), albedo, the climatology as well as the trace gas and aerosol profiles. While geometry, albedo and climatology are more or less known, the actual profiles are usually unknown. Here, a similar approach as in Section 3.5.1 is used. LUTs were calculated with the same properties as in Table 3.4 but for different wavelengths (O_4: $360.7\,\mathrm{nm}$ and $477\,\mathrm{nm}$, NO_2: $357\,\mathrm{nm}$ and $448\,\mathrm{nm}$). For the UV2-Fit, the dAMF values are linearly interpolated between both wavelengths. Note, that the specific ΔM values are chosen as before but the PBL has an outstanding importance here as it is directly accounted for within both, the selection of ΔM and the calculation of f_c. The O_4 concentration is calculated with the surface pressure and temperature from the DWD meteorology station close to the airport of Bremen[17] while $V_{O_4} = 1.358 \times 10^{43}\,\mathrm{molec}^2/\mathrm{cm}^5$ is the mean value of US standard atmospheric profiles (NASA, 1976) which were scaled with the before calculated O_4 concentrations.
Similar to the mentioned publications, filtering is done to account for inaccuracies of the approach, the RTM calculations and the individual conditions. The ΔS filtering is done as introduced in Section 3.5.2. Additionally, RAA $\leq 50°$ are removed. This was introduced by WANG et al. [2014] and was confirmed as smaller RAA lead to variations of the individual f_c values and outliers in the later shown scatter plots (e.g. Fig. 3.21). Furthermore, SZA $\geq 85°$ are excluded and measurements are removed when the corresponding AOD's (found from CAMS) are smaller than 0.1.

In order to evaluate the largest dependencies on the correction factor, Figure 3.21 shows scatter plots of f_c for all three wavelength regions (colour-coded) and several parameters for data of direction c. For this plot, a constant PBL of $750\,\mathrm{m}$ height was used with one O_4 vertical column ($V_{O_4} = 1.358 \times 10^{43}\,\mathrm{molec}^2/\mathrm{cm}^5$) instead of the scaling approach explained above. The O_4 concentration calculation and ΔM extraction was done as stated before. Note that no RAA filtering was applied when creating this plot. For these scatter plots, especially extremely high or low f_c values are questionable as they indicate a stronger correction of the calculated quantities meaning that the underlying assumptions are maybe not valid or that in these cases, the applied parameters for calculating f_c should be known accurately.

[17] Source: FTP Server - DWD [2018] - `ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany/climate/hourly/`, Date of download: 14.03.2018.

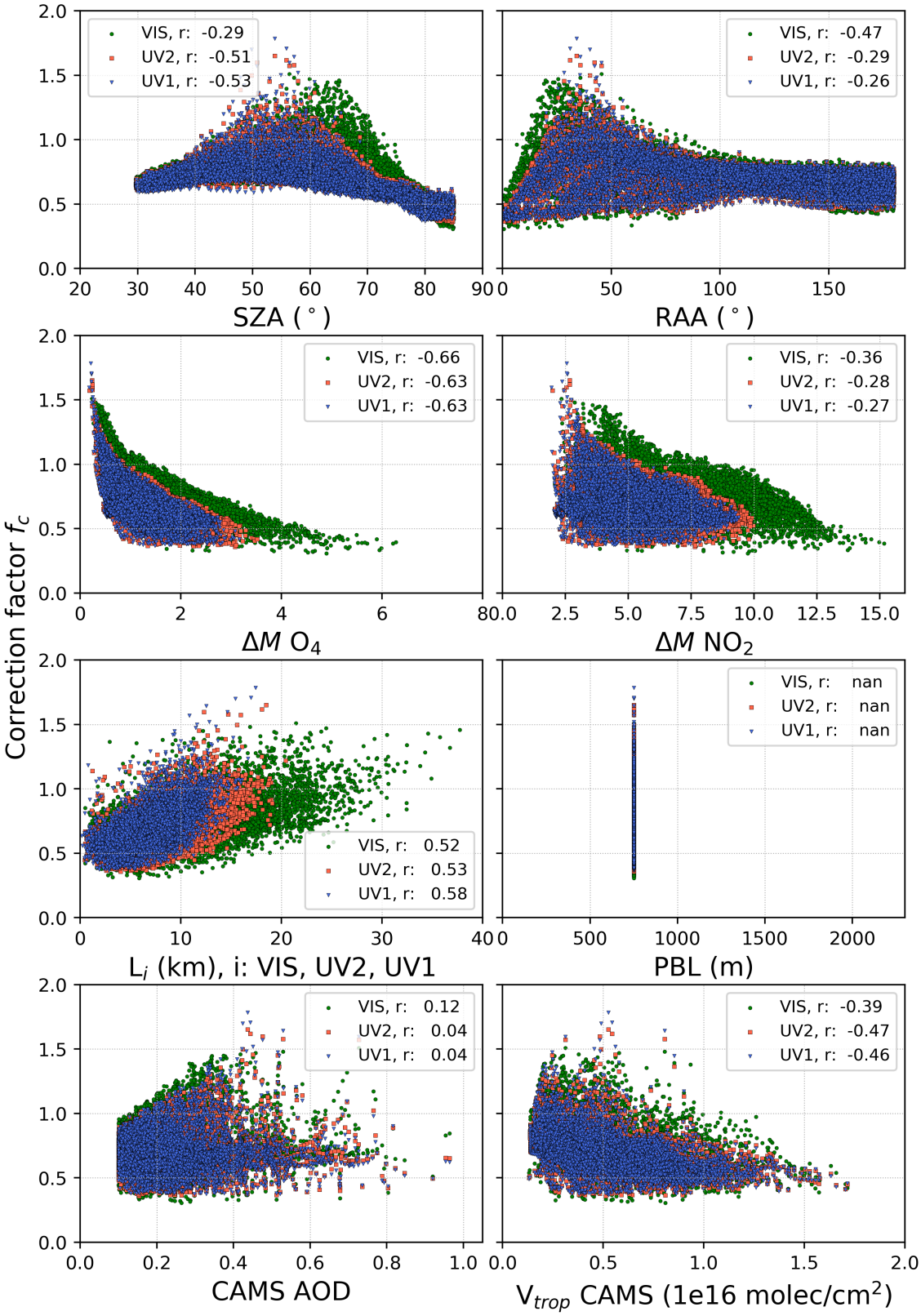


Fig. 3.21: *Scatter plot of f_c with various parameters for three different fitting windows and direction c (colour-coded). Pearson's correlation coefficient (r) is given in the legend of each sub-plot. PBL and V_{O_4} were kept constant.*

It can be seen that there are clear minima and maxima of f_c for both geometry angles (SZA, RAA) with maxima between $50°$ and $65°$ for SZA and for $20°$ and $40°$ RAA. SINREICH et al. [2013] already showed that the diurnal shape of f_c depends strongly on SZA and RAA meaning that the azimuthal viewing direction has also an impact (cf. with dir s, Fig. AF.15 in the appendix). The minimum in the SZA plot for large angles implies a problem during twilight indicating that the approach might not be reliable for SZA $\geq 80°$. Especially under these conditions, the assumption that ΔS values for different LOS and the absorption of the zenith measurement and another elevation angle measurement before the scattering point are similar, might be wrong. This indicated a stronger dependence on different profile shapes of NO_2 and O_4. The RAA minimum was found at $0°$ and $180°$ showing again that aerosols might have a strong impact in the forward and backward scattering direction. This argument can be supported by comparing direction c with s (Fig. AF.15 in the appendix). Here, a larger variation of f_c can be observed leading to higher uncertainties when pointing towards the sun. The next row of subplots in Figure 3.21 shows the dependence of f_c on the differential air mass factor of O_4 and NO_2. The largest correction factors can be found for the lowest ΔM. The lowest f_c values belong to small RAA or large SZA while the largest f_c can be accompanied with the range of SZA angles which were responsible for the peaks in the SZA vs. f_c subplot. The NO_2 ΔM shows a similar pattern but the range of f_c is larger indicating a smaller dependency on the NO_2 dAMF. The next row shows how the resulting effective light path length depend on f_c and the constant PBL. In general, the larger f_c the larger is L but the spread of L values indicates a smaller importance of f_c for the calculation of effective light path lengths. The lowest row shows the scatter plots of f_c vs. AOD and the tropospheric NO_2 vertical column which are both input parameters for the ΔM extraction. Since the scatter is large for the AOD, no strong dependency can be expected which is in agreement with the findings of SINREICH et al. [2013] and WANG et al. [2014]. The NO_2 tropospheric column has a slightly higher impact but especially for low values the spread is large.
In general, the correlation (anti-correlation) is largest for O_4 ΔM and smallest for the AOD. The other quantities show medium large anti-correlations indicating that the dependence is more or less equally shared between all input quantities. When comparing the correlation values r with the s direction in the appendix, distinctly larger r can be found for the geometry angles and both ΔM while the NO_2 V_{trop} shows a decrease. Note that the dependence on the O_4 surface concentration c_{O_4} (not shown) was found to be small.

Figure AF.18 in the appendix shows the same dataset with variable ECMWF PBL values used for the calculation of f_c (or dir s, Fig. AF.19). A strong correlation for the PBL can be found while all the other variables show larger spreads. The dependency of L on f_c is now stronger as low PBL values are directly linked to smaller effective path lengths showing the limitations of the approach. This indicates that the knowledge of the actual box

height of the specific trace gas layer is crucial. Unfortunately, the PBL values are not an accurate measure of the prevalent trace gas layering condition during the measurement. Due to this reason, ORTEGA et al. [2015] used the mixing layer heights (MLH) retrieved from a MAX-DOAS profiling algorithm as a better estimate. Since this is a rather time consuming step for a large dataset as analysed here, the following subsections show data calculated with a constant PBL of $750\,\mathrm{m}$. This value was close to the mean value ($787\,\mathrm{m}$ with a standard deviation of $468\,\mathrm{m}$) of all ECMWF PBL used in the filtered dataset, which will be analysed in the next two subsections. Figure AF.16 and AF.20 (or dir s, Fig. AF.17 and AF.21) in the appendix show the result for a scaled V_{O_4} as input for the f_c calculation. The dependencies stay similar but the spread of f_c is generally larger.

Figure 3.22 shows on the left-hand side mean values for the different seasons of the year of f_c as diurnal variation. The correction factors are the smallest in winter and the largest in summer. The overall diurnal shape depends strongly on the geometry and shows a maximum for all curves at a different time for another azimuthal direction (cf. Fig. AF.22 in the appendix). The reason for the strong diurnal variations is caused by large SZA and RAA values (cf. with Fig. 3.21). Note that the correction factors in the morning and evening are much smaller than 1 indicating that the effective light paths are much too long without correction. The right-hand side of Figure 3.22 shows the time series of monthly mean f_c values for all three fitting windows and direction c with the standard deviation as shaded area. Clear winter minima and summer maxima can be identified. This shape is independent of the azimuthal viewing direction but the range of f_c values is larger in direction s than for c. Note that this plot for direction s shows a noisier annual variation which is due to the smaller number of data points.

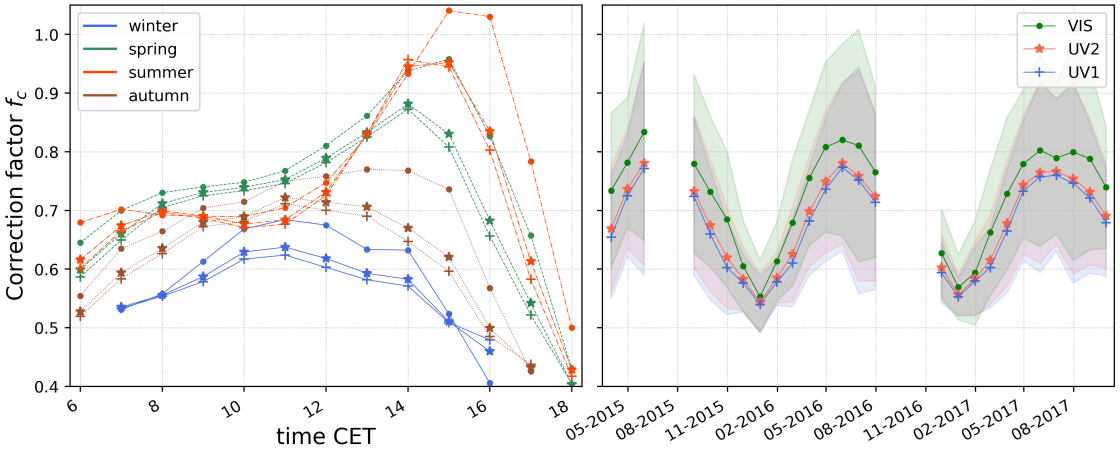


Fig. 3.22: Left: *Mean f_c values for different seasons of the year colour-coded and for direction c. Different marker styles depict the fitting windows (see right legend).* ***Right:*** *Time series of monthly averaged f_c with the fitting windows colour-coded. PBL and V_{O_4} were kept constant.*

3.8.2 Case study

In this subsection, a case study based on the example of three different onion peeling time series for direction c is presented. Since this viewing direction pointed directly to different emitters, a correlation with emission values is expected. Furthermore, two in situ stations are close to this viewing direction but only I_2 is used as a validation source here because I_1 found rather small values throughout the year due to the prevalent wind direction. Figures for directions b and s can be found in the appendix (AF.23 - AF.28). These cases were selected as the number of suitable onion peeling days is quite limited after filtering and the specific high emissions and wind directions show interesting features which can be understood as representative for the measurement location.

Case 1

Figure 3.23 shows the onion peeling result for 10 days in the year 2015. Time is plotted on the x-axis while the y-axis represents the distance s from the instrument. The horizontal, black, dashed lines show the estimated distance to large emitters (cf. map of Bremen 3.2). The black arrows indicate the specific wind directions and velocities at 6, 12 and 18 o'clock, local time. The x-axis ticks are arranged to match these times. The maximal wind velocity (the longest arrow) is noted at the lower right corner of the plot. The continuous colour-coded dataset at the smallest distance shows the measurement results of the in situ stations for the NO_2 concentration divided by 2 in order to enable the use of the same colour bar (vf. correlation and discussion of in situ and onion peeling results in Sec. 3.8.3). Note that station I_2 is close to PS_{west}. Each vertical line is separated into three different coloured segments which represent the NO_2 concentration of the specific fitting windows calculated with Equation 3.13 at different differential effective light paths (Eq. 3.12).

Figure 3.23 shows interesting features which can be associated to different conditions and spatial NO_2 inhomogeneities. The wind direction and velocity varies strongly in the discussed time range with winds from south easterly directions at the beginning and south westerly winds in the mid. At the end of the time period, variable wind directions with low wind speeds prevail. The first four days show the impact of the wind direction on the onion peeling with wind coming from the south-east, accompanied by low NO_2 concentrations and long light paths. When the wind changes to variable directions with low velocities on the third day, the concentrations are larger for all three distance ranges. Furthermore, it can be seen that the largest concentrations are found for the mid fitting window accompanying these high values with distances close to the industrial area with the steel manufacturing facility and PS_{west}. Larger values on the first two days in the morning might be due to commuter traffic. Smaller distances confirm these results.

The in situ station shows a similar pattern in the morning indicating that a larger spatial

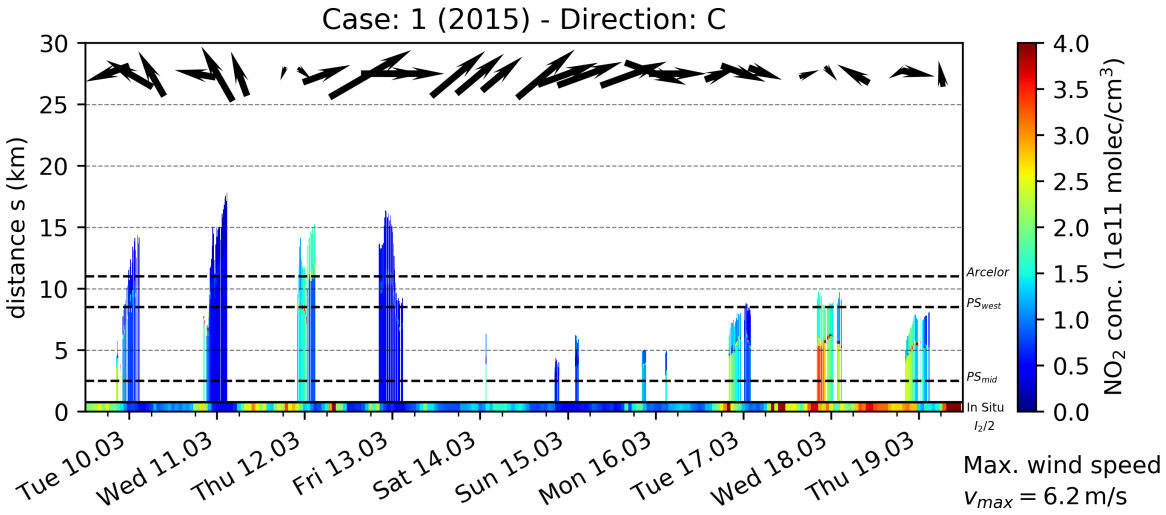


Fig. 3.23: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ values divided by 2 are shown at the smallest distance. The black dashed lines show the distance to strong emitters.*

airmass is affected by similar near-surface emissions rather than a single emitter (with a certain emission height). In contrast to that, the agreement with the high values for the mid fitting window is not as good as expected showing that either different air masses were probed or vertical and/or horizontal inhomogeneities were present. The remaining days are strongly affected by ΔS-filtering due to cloudy conditions which lead to less data points. Note that clouds might lead to smaller effective light paths but are not necessarily problematic for the NO_2 ΔS measurement. However, since Equation 3.13 uses with L a quantity strongly affected by a decreased O_4 ΔS measurement in cloudy scenes, an impact on the concentration is expected for data which was not successfully cloud-filtered. On the other hand, the last days show a good agreement with the in situ station even though the effective light paths are short. The prevailing wind directions support the results as low wind velocities prevent fast transport of polluted air masses keeping the emissions of traffic as well as industry in the area of Bremen. The large concentrations for the mid fitting window have a small differential effective light path so that it is unclear if this is a real plume or a problem of the ΔS measurement. The fitting windows UV1 and UV2 led to quite similar O_4 and NO_2 ΔS on these days.

Case 2

Figure 3.24 shows 10 days for which many measurements remain even though filtering was applied. Here, longer lights paths but smaller NO_2 concentrations can be observed,

in comparison to Case 1[18]. The wind is variable on the first day. For the other days, the wind changes from north-westerly directions to the south-east with more or less increasing wind speeds. The maximum wind speed on the 12th of May was larger than that for all days of Case 1. Interestingly, the mid fitting window shows the smallest concentrations for the full time period but with larger differential effective light paths than in Case 1.

North-westerly winds are accompanied with higher concentrations close to the instrument (UV1) and for the largest differential effective light paths in the mornings. Here, power plants and emissions from commuter traffic cross the light path of the instrument when shortly transported into this direction. On days when the visible fitting window led to clearly larger concentrations than UV2, the matching distances show high concentrations where the industrial area is located. In the afternoons, larger distances and lower NO_2 concentrations can be found for all days. On the last day, the wind came from southerly directions implying transport of emissions from the city centre into the FOV. The concentration is small and close to the instrument. VIS and UV2 values are negligible compared to UV1.

The in situ station shows a good agreement with the UV1 (and sometimes VIS) concentrations throughout the full time period with the exception of two strong concentration peaks in the morning of the 9th and 10th. These maxima can not be found with any fitting window which might be explained by surface near NO_2 in the morning when the PBL is thin. These air masses are most likely not probed by the 1° elevation angles measurement.

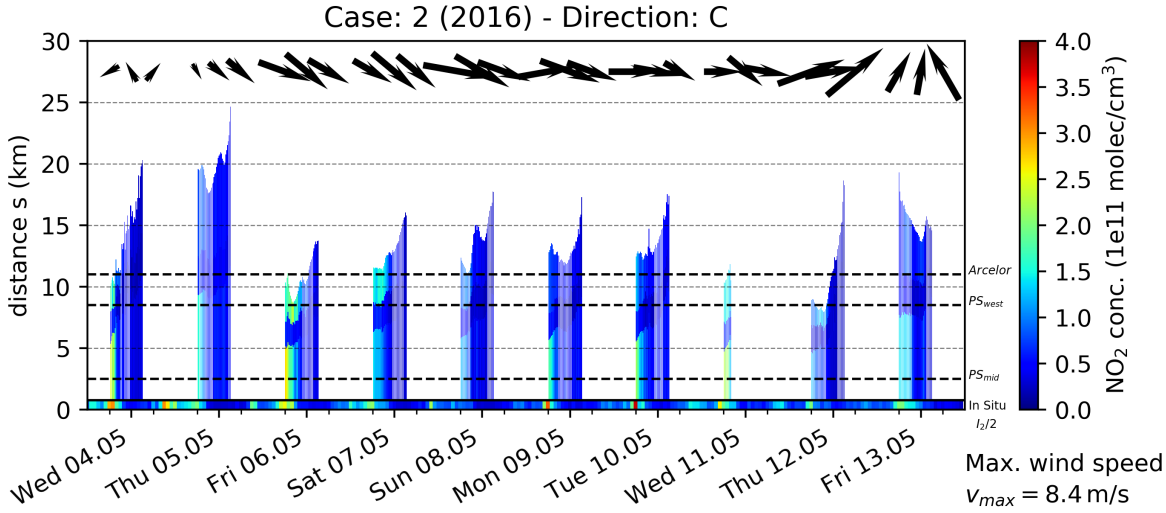


Fig. 3.24: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ values divided by 2 are shown at the smallest distance. The black dashed lines show the distance to strong emitters.*

[18] Note that for this case and direction c, Figure AF.29 in the appendix shows the results without application of the correction factor f_c. Long effective light paths can be found in the morning and late afternoon indicating the need for a proper correction.

Case 3

Case 3 is shown in Figure 3.25 and represents 10 days with highly variable wind speeds and directions. On the first two days, the concentrations are more or less low with some higher values close to the instrument in the morning of Monday the 18th of July. This day is followed by high concentrations close to the instrument from the morning until the early afternoon of Tuesday. It is interesting to see that low wind speeds from the north-east lead to these high concentrations while the afternoon values, accompanied with a change in the wind direction, lead to smaller NO_2 concentrations. This could also be explained by decreased emissions. However, the power plants in this azimuthal directions are expected to emit constantly throughout the day. The next day shows again high early morning values but also an interesting pattern in the UV2 range. Here, the values are high throughout the whole day. The wind direction might indicate that a single emitter (e.g. the steel manufacturing facility) could be responsible for this peculiar signal. The shape of these values over the day shows that there might be still an unaccounted feature in the correction factor since the distance to this emitter should not change, when the wind speed and direction stays constant (which was the case). Since O_4 concentration, PBL and the vertical column of O_4 do not yield drastic elongations of the light path (cf. Sec. 3.8.1), aerosols might play an important role. The AOD used for the ΔM extraction was not found to be an important dependency of f_c. However, the inaccurate aerosol phase function might still be a problem and additional tests should be considered in the future. The remaining days show not a good temporal coverage throughout each day but high NO_2 concentrations can be seen on the 21st and the 26th of July. Especially on the

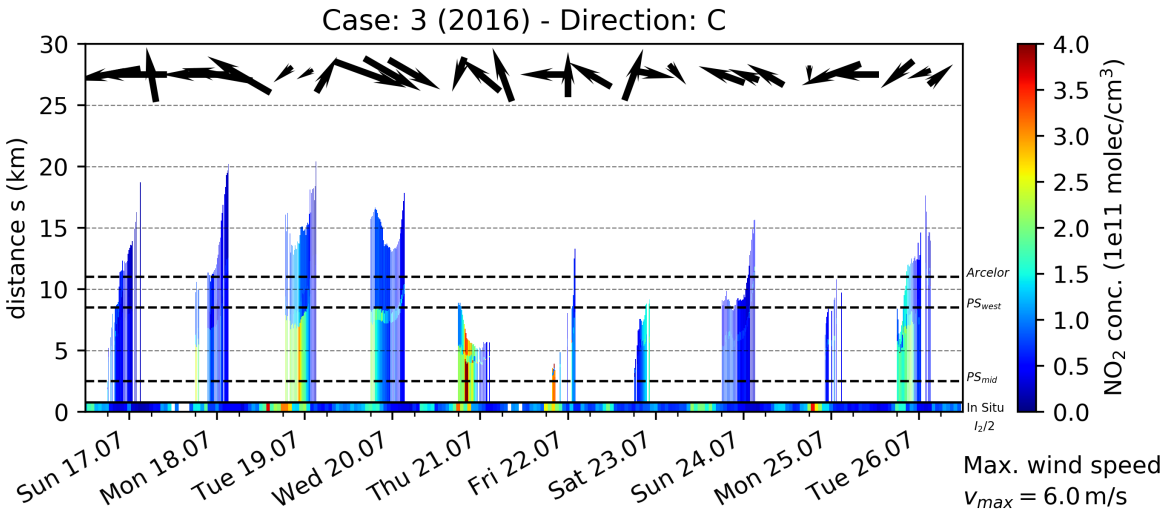


Fig. 3.25: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ values divided by 2 are shown at the smallest distance. The black dashed lines show the distance to strong emitters.*

21st, the UV1 values show the largest concentrations of the week which were not found by the in situ instrument. Since all days show highly variable winds, the correlation with the in situ measurement site can be considered as weak for this case.

3.8.3 Statistical analysis and emitter back-tracing

The previous subsection dealt with cases because longer time periods with continuous data are not available due to the necessary filtering steps. It was already shown that on some days, correlation with ancillary in situ measurements seems to be high while for other days, no agreement was found. Furthermore, even when the concentrations seemed to be reasonable, the effective light path lengths showed a high variability throughout the investigated time period. For the following analysis, all filtered data points are considered rather than specific cases in order to analyse the full dataset and giving the results a more statistical meaning. Furthermore, an attempt for the creation of a NO_2 concentration map of Bremen is made from the onion peeling results.
Figure 3.26 shows the correlation of all data points from 2015 - 2017 with the in situ NO_2 concentrations of I_2 for direction c. The colour-coding indicates the results of the different fitting windows in addition to data points which are referred to as HIGH. For these points, the highest value of the three NO_2 onion peeling concentrations was selected for this specific time in order to validate if there is a dominant high concentration emission source.

Statistical analysis

It can be seen that the correlations for the HIGH dataset and the UV1 concentrations are the highest while the VIS values show only a fair correlation coefficient. All four regression lines were found to have a slope much smaller than 1 indicating that the onion peeling concentrations are usually smaller than the in situ measurements. A closer look to the HIGH dataset shows that the highest NO_2 concentration is more often found for UV1 rather than for VIS or UV2. However, the correlation for HIGH is not larger than for UV1 meaning that the idea of dominant emissions from the industrial area ($ArcelorMittal$ and PC_{west}) can not be confirmed. This supports the findings from Section 3.7 and shows that pollution in the city of Bremen is dominated by the sum of all emitters.
These correlation plots for direction b and s can be found in the appendix (Fig. AF.30 and AF.31) and show slightly smaller correlations with similar underestimations of the NO_2 concentrations calculated with the onion peeling approach.

In Figure 3.27, temporal average values for direction c and the three concentrations retrieved from VIS, UV1 and UV2 are shown[19]. In the top left sub-plot, mean daily variations for the four seasons of the years (colour-coded) are depicted for the three fitting windows

[19] In the appendix, Figure AF.32 and AF.33 show similar plots for the directions b and s, respectively.

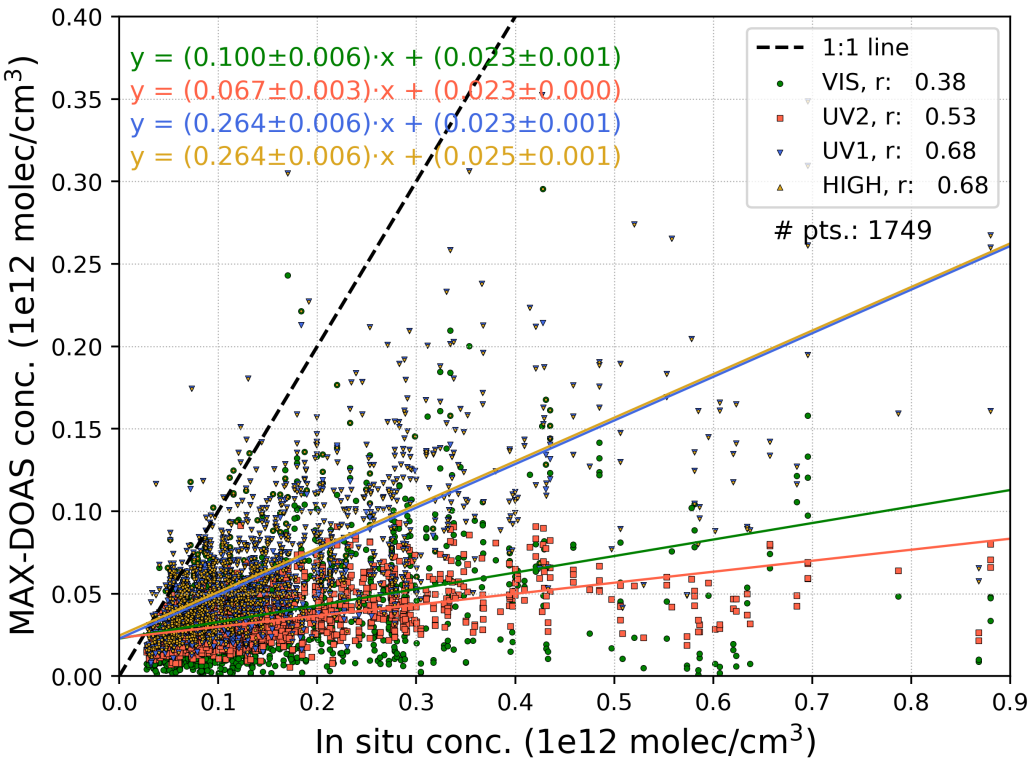


Fig. 3.26: *Correlation plot of onion peeling NO_2 concentration datasets with in situ measurements for direction c. The dataset HIGH summarizes the highest concentrations of the three onion peeling values per time step.*

(different marker styles, cf. top right legend). It can be seen that for all seasons the UV2 values are much smaller than those for the other two fitting windows. When comparing the curves for VIS only, the winter season (dots) shows much higher NO_2 concentrations than spring, autumn and summer. This finding is similar for UV1 while the UV2 concentrations do not show a clear seasonal variability. This might be due to different reasons: 1. the UV2 fit is maybe not as reliable as the other fits, 2. More filtering is applied because the differential effective light path lengths for UV2 are usually the smallest (and sometimes negative), 3. there might be specific emissions which dominate the UV2 results so that no seasonal pattern can be identified when the emissions are constant throughout the year. Unfortunately, a final conclusion for this behaviour cannot be given at this point.
The top right sub-plot presents monthly mean values for the full time period for the three fitting windows (different colours and marker styles) and for the in situ station I_2 (black dashed line). VIS and UV1 show high values in winter and smaller values in summer. The in situ station confirms these results for the full time period except in winter 2015/16 when no clear peak can be seen. UV2 has a good agreement with all curves in the year 2017 while the remaining years do not agree well. Note that the standard deviation for UV2 is the smallest indicating smaller deviations in the averaged time periods while VIS

and UV1 show a larger spread. The in situ dataset is on average 48 % larger than UV1, 70 % larger than UV2 and 57 % larger than VIS.

The bottom left sub-plot shows the average diurnal cycle for each month of the year. Note that the x-axis is equidistant but the shown data points are not. Within one month, each point is the mean value of all three years for the specific hour. The vertical black lines represent the separation line between each month with the starting time as x-label. Again, it can be seen that the winter months are accompanied with the highest NO_2 concentrations for VIS, UV1 and the in situ dataset. UV2 shows an increase for the last months but not the first, which might be due to a smaller number of data points. Furthermore, all curves show the largest values in the early morning and a decreasing trend throughout the day in the summer months with a more chaotic daily variation in the winter. However, the general difference between the in situ dataset and the onion peeling results seems to be larger in summer than in winter. This can be explained by both, the total number of measurements (which is much smaller in winter), and the differently probed air masses. The bottom right

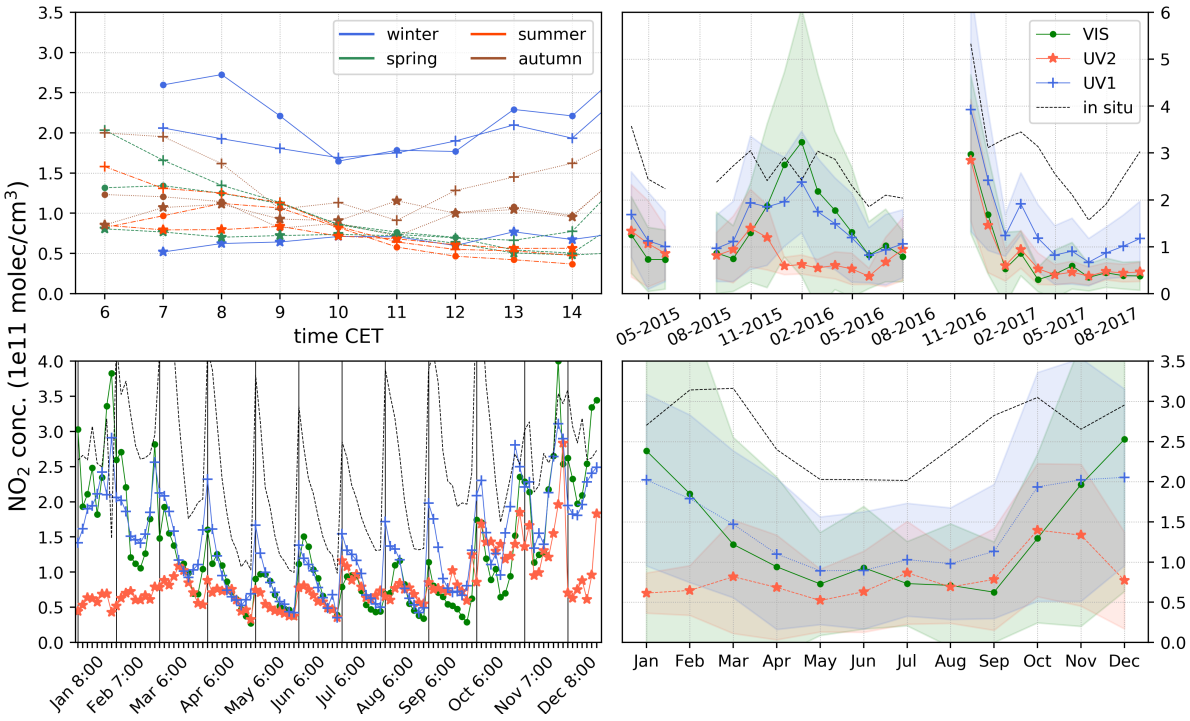


Fig. 3.27: NO_2 *concentrations for the three fitting windows as temporal mean values and for direction c.* ***Top left****: Daily variation for the four seasons of the year.* ***Top right****: Time series of monthly mean values.* ***Bottom left****: Mean daily variations for each month of the year.* ***Bottom right****: Monthly mean values averaged over the full time period. Different marker styles indicate the three fitting windows while the shaded area shows the standard deviations. Black, dashed lines represent in situ concentrations.*

sub-plot shows the monthly mean values for the full time period. Again, the highest values can be found in summer for VIS and UV1 while UV2 does not confirm this trend. The in situ station shows not the decreasing trend as VIS and UV1 does indicating again larger

variations in the NO_2 concentration in winter. There are different potential explanations for the larger deviations in winter. First, the number of remaining data points after the filtering is much smaller in winter indicated by the larger standard deviation. Also the lifetime of NO_2 is usually longer in winter which might lead to more transport meaning a higher spatial inhomogeneity of the NO_2 distribution in Bremen. Furthermore, not only enhanced traffic emissions can be expected but also larger emissions due to heating in residential areas and a higher energy consumption leading to larger emissions from power plants. These might increase the spread between near surface and elevated emissions because of the higher injection altitude due to the individual stacks.

Emitter back-tracing

The high correlation with the in situ complementary measurements (cf. Fig. 3.26) and the reasonable results in the temporal analysis of mean values (cf. Fig. 3.27) support the accuracy of the NO_2 concentration values itself but the reliability of the effective light paths is questionable. Unfortunately, no measurements are available which can be used to validate the distance of the specific concentration. The Figures AF.34 and AF.35 in the appendix show a similar plot as Figure 3.27 but for the effective light path and the differential effective light path for direction c, only. However, these plots do not support the accuracy of the presented values and no further conclusions can be found in addition to the discussions above.

It was already stated that no emitter can be clearly matched with specific NO_2 concentrations and light paths so that an empirical approach is presented here in order to assess the accuracy of the distance to the measurement location and to assess if some concentrations can be attributed to a certain emitter in Bremen. Furthermore, since three different azimuthal viewing directions are available, an attempt for the creation of a NO_2 distribution in Bremen is made. For that, onion peeling results from the three different azimuthal viewing directions, the wind velocity and direction from the DWD meteorology station close to the airport, and NO_2 lifetimes are utilized.

A schematic representation of the approach used for the validation of the effective light paths is shown in Figure 3.28. Even though single emitter were not found to

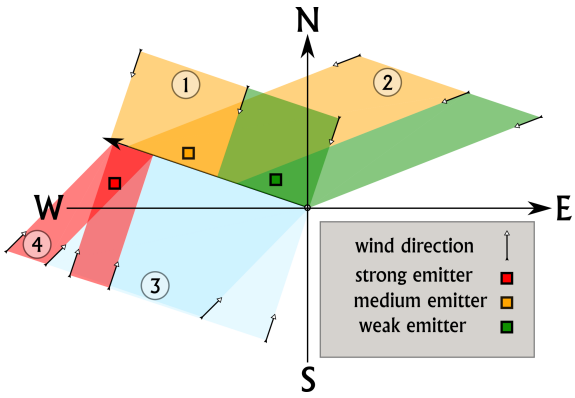


Fig. 3.28: *Schematic representation of the approach for creating a map of the NO_2 distribution in Bremen. The telescope's viewing direction is indicated by the long arrow. The numbers name different measurements cases.*

contribute to the vertical columns in a dominant way (cf. Sec. 3.7), some results of the case study indicated that, depending on the wind direction, higher (or lower) concentrations are more often measured in a certain viewing direction. This is the main idea behind the following approach: When a certain wind direction leads to higher NO_2 concentrations due to specific emitters only (under the assumption that the background NO_2 level and other emissions sources are negligible), this concentration should be measurable with the onion peeling approach. Consider case 1 in Figure 3.28 where the wind transports NO_2 from the north-east into the field of view of the instrument. If the next wind direction (case 2) traverses these two emitters as well, a second measurement with higher concentrations farer away from the instrument (and lower close to the instrument) would be measured. The two wind vectors as well as the differential effective light paths create areas in which the average concentration is assumed to be constant. If the wind vector is known, the mean values of both measurements would reveal that there are triangles around the emitters with higher concentration values than the surrounding areas. If no emitter is present (case 3 for small effective light paths: blueish colour) the onion peeling would have found a smaller signal which cancels out during the averaging (consider the cases 3 and 4). However, the length of the wind vector and therefore the second dimension of the limiting area, is usually unknown. Here, NO_2 lifetimes calculated with the chemical transport model (CTM) TM4-ECPL (KANAKIDOU et al., 2012; DASKALAKIS et al., 2016) are used in order to create wind vectors. The idea is, that a certain source emits NO which is then transformed into NO_2 (here, with the assumption of an immediate conversion, see Sec. 2.6) and which is transported by the wind along the vector *wind speed* $\times$ *lifetime*.

This approach is evaluated for all onion peeling measurements in the investigated time period as follows: Wind vectors are created for each onion peeling result with the wind directions and wind speeds from the DWD measurement site near the airport along with lifetimes from the CTM (sampled in $30\,\mathrm{min}$ values). A map of Bremen is divided into 81×81 grid cells with a width of $500\,\mathrm{m}$ each, meaning that an area of $1640.25\,\mathrm{km}^2$ was covered within a coordinate system of $\pm\ 20\,\mathrm{km}$ in all directions, with the IUP Bremen in the origin. The areas covered by the wind vector times the differential effective light paths (for each fitting window individually) are calculated according to this grid and projected onto the map of Bremen. For each (x,y) grid cell, a mean value and the standard deviation was calculated based on all maps which showed non-zero concentrations on this specific cell, for all three directions and for the different fitting windows individually. Here, for statistical reasons, at least 6 data points are needed for the calculation of mean values and standard deviations. Grid cells having smaller number of measurements with concentrations unequal to zero were set to zero.

In general, a better spatial resolution can be achieved the higher the number of measurements is and when the matching wind vectors cover a wide range of wind directions.

Unfortunately, there are cardinal points which will never lead to a reliable onion peeling result, for instance, when the wind direction angle and the viewing azimuth angle are equal or with an offset of $180°$ to each other. This leads to the requirement of using different instrumental viewing directions in order to create a more accurate map of the surrounding area of Bremen.

Figure 3.29 shows a map for direction c with the approach explained above. The top row shows the mean NO_2 concentration values for the three different fitting windows and the mean map of the first three maps. The second row shows the standard deviations according to the first row. In all four mean maps, stripe-like features with higher concentrations can be identified together with wider areas with medium concentrations. These stripes point always to certain pivot points which are closer to the measurement location for UV1 and farer away for UV2 and VIS. Note that for UV1 and its standard deviation another colour-map maximum value was used. Some stripes in UV1 seem to focus on a point between PS_{mid} (yellow rectangle closest to the origin) and the IUP Bremen. When comparing UV2 and VIS it can be seen that some stripe like features point directly towards the other power plants. For example, the three high concentration stripes in the lower left corner of the UV2 sub-plot seem to focus on a location close the power plant PS_{west} (westernmost yellow square). Also the top right corner stripes point onto a similar location. This result can be confirmed by looking at the VIS plot which shows similar stripes with a pivot point close to PS_{west}. Here, narrow point-like stripes can be seen in the lower right corner which are artefacts due to the finite resolution of the grid, when small numbers of wind vectors together with high concentrations are averaged. These features are not discussed due to the missing statistical meaning. Note that direction c was pointing to a viewing azimuth angle of $-75°$ with respect to the North (ccw), meaning that the direction of $105°$ (cw) is *blind* for this approach (blue triangle-shaped area).
The *ALL mean* map shows an average of all three fitting window maps on which some stripes can not be identified any more. However, the overall tendency stays the same. Depending on the wind direction, measured NO_2 concentrations can be attributed to the area around PS_{west} indicating that there is the dominant emission source located, for viewing direction c. The matching standard deviations show often high values where the stripes are located. However, similar concentrations for two stripes do not necessarily mean similar standard deviations. This indicates that some wind directions lead more often to good onion peeling results than others (cp. the two stripes for VIS, at the coordinates (3, -20) and (5, -20)).

Figures AF.36 and AF.37 in the appendix show these maps for direction b and s, respectively. The results for direction b support the graph above as concentrations in UV1 point often towards the area around PS_{mid} while UV2 and VIS NO_2 stripes point mostly in the

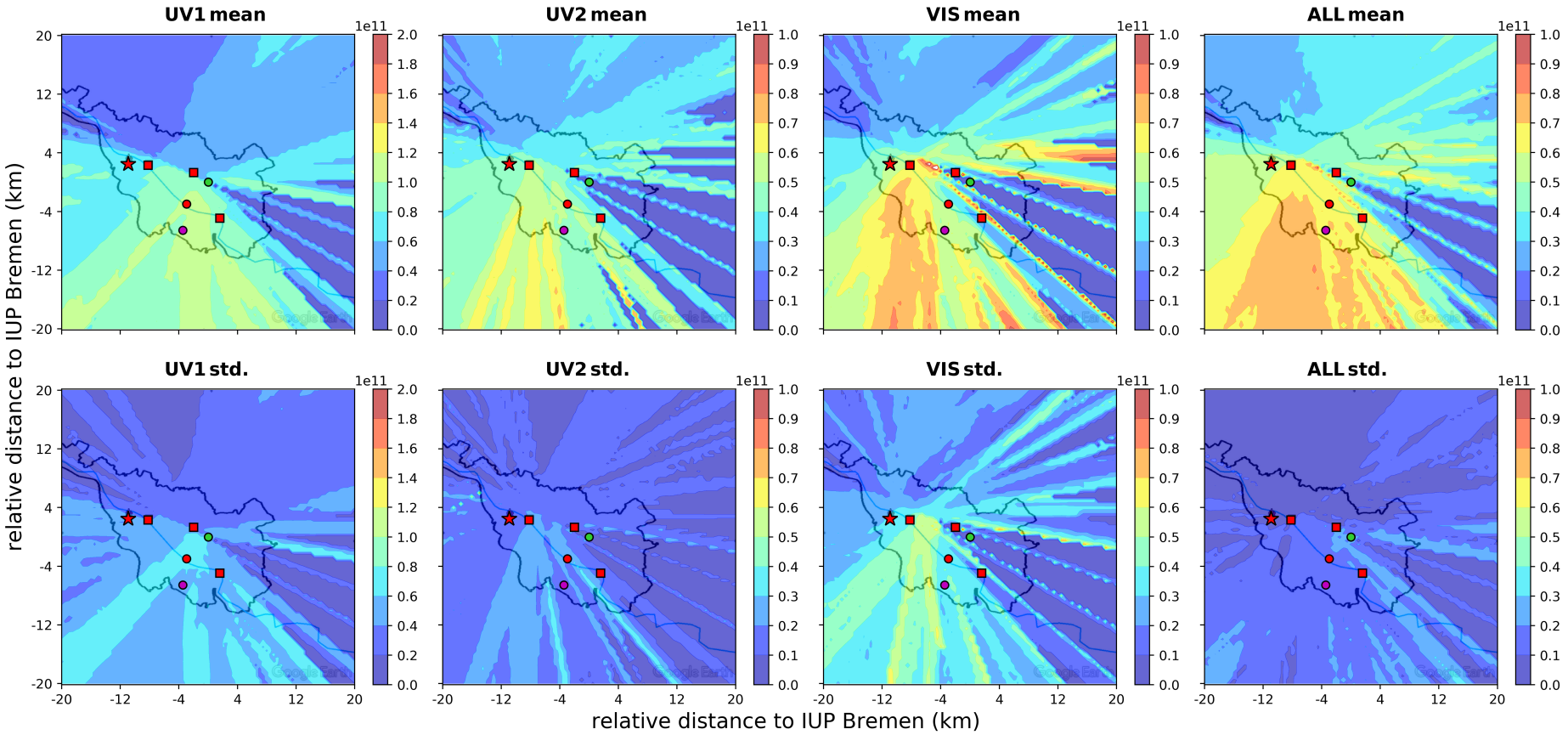


Fig. 3.29: *Mean NO_2 concentration maps of Bremen, with the concentration colour-coded in molec/cm^3. In the top row, three different fitting windows and the average of these three maps for is shown, for direction c. Matching standard deviations can be seen in the bottom row. Red squares depict power plants, the red star shows the steel manufacturing facility. The dots in green, red and purple represent the IUP, the city centre, and the DWD airport meteorology station, respectively.*

direction of a location close to PS_{west}. In this figure for direction s, also stripes pointing towards PS_{south} can be identified while features focussing on PS_{west} are not as frequent any more.
Figure 3.30 shows the average maps of all three azimuthal viewing directions. Still, stripe-like features which point towards certain pivot points can be identified but the angular distribution around the origin shows larger values all over the map. UV1 concentrations remain focussing on a location close to PS_{mid} while for UV2 and VIS, stripes pointing to every emitter in Bremen seem to exist. This indicates that areas with high concentrations depend strongly on the wind direction and the distance to certain emitters. In the VIS and ALL mean map, some hotspots can be identified between PS_{west} and PS_{mid} as well as close to PS_{south} but with lower concentrations here. It is interesting that the largest emitter for NO_2 in Bremen, the steel manufacturing facility ArcelorMittal (cp. Tab. 3.1), does not show significant values. This could be either due to wrong effective light paths length with an underestimation of the real length or because of characteristic emission features of ArcelorMittal. In contrast to power plants, it might be possible that the steel-manufacturing facility does not emit continuously but intermittently. The E-PRTR inventory only provides annual mean values so that this problem can not finally be solved here. However, since the prevalent concentration stripes point more or less accurately to specific power plants, the average effective light path can be considered as accurate. Unfortunately, this does not mean that it is accurate enough for specific case studies so that the allocation of high NO_2 concentration events to specific emitters is not suggested, when several possible sources are close to each other.

The main idea behind this approach was the validation of effective light path lengths and the creation of maps with a reasonable NO_2 concentration distribution. While the validation can be understood as a success as the stripes point onto the more or less precise emitter locations, an accurate map of Bremen's NO_2 distribution was not found. The stripes indicate that either not enough azimuthal viewing directions were considered, the angular spacing was not applied fine enough or the number of cases where the wind comes from other directions than the dominant south-westerly was not sufficiently high enough. However, a finer spacing was not possible as the number of data points was limited. Additionally, the larger orange areas around (-10, -12) in Figure 3.30 indicate that more data points accompanied with wind directions where the smaller stripes were found might be enough to give the map a smoother appearance.

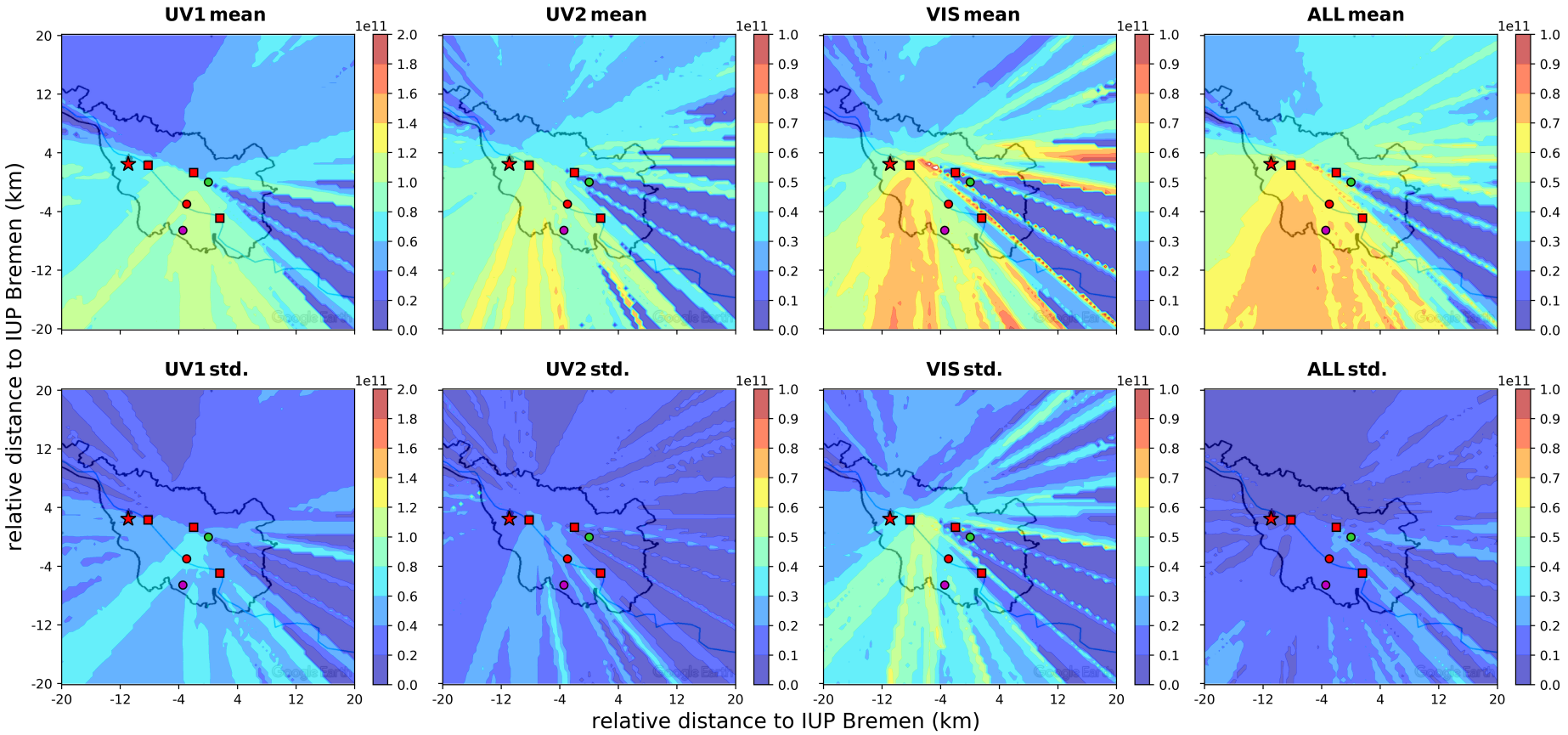


Fig. 3.30: *Mean NO_2 concentration maps of Bremen, with the concentration colour-coded in molec/cm^3. In the top row, three different fitting windows and the average of these three maps is shown for all viewing directions. Matching standard deviations can be seen in the bottom row. Red squares depict power plants, the red star shows the steel manufacturing facility. The dots in green, red and purple represent the IUP, the city centre, and the DWD airport meteorology station, respectively.*

CHAPTER

IV

PROFILE RETRIEVAL OF MAX-DOAS DATA

Note that in this chapter, an asterisk symbol (*) in the section title indicates parts which have already been published in BÖSCH et al. [2018]. The first sub-section is a short introduction into the retrieval of vertical aerosol and trace gas profiles from MAX-DOAS measurements (Sec. 4.1). It is followed by the description of the retrieval algorithm (Sec. 4.2). Sub-section 4.3 focuses on synthetic data and shows sensitivity studies on important parameters and dependencies. The last sub-section (Sec. 4.4) shows profiling results of real data measured during the CINDI-2 campaign (Cabauw, 2016) and in Bremen, for the years 2015 - 2017.

4.1 Introduction into the retrieval of vertical profiles*

For more than 15 years, Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) measurements have been used to investigate the chemical composition of the troposphere (HÖNNINGER and PLATT, 2002; BOBROWSKI et al., 2003; LESER et al., 2003; WITTROCK et al., 2004; WAGNER et al., 2004; van ROOZENDAEL et al., 2004). This passive remote sensing method is based on absorption spectroscopy applied to measurements of scattered sunlight. The advantages of ground-based MAX-DOAS when compared to satellite observations are the high sensitivity for the lowermost layers of the troposphere, the high temporal and spatial resolution of measurements, and the lower cost.

The strength of the absorption signal detected in scattered sunlight depends on the absorber amounts and their vertical distribution but also on the length of the light path. In general, this length in a certain altitude layer is a function of the measurement geometry. Therefore, a set of MAX-DOAS radiance measurements taken at different elevation angles (lines of sight, LOS) contains information about the vertical distribution of trace gases, which can be retrieved. However, the retrieval of absorber profiles from MAX-DOAS measurements is an ill posed problem that needs additional constraints (see Sec. 4.3). Thus, the inversion, which is applied for retrieving vertical concentration profiles, is done by elaborated mathematical methods such as optimal estimation (OE). In addition to the trace gas profile, the aerosol extinction profile needs to be retrieved as well as it has a non-linear effect on the trace gas retrieval and it is too variable to be approximated by climatologies.

* This section has already been published in BÖSCH et al. [2018].

In contrast to the trace gas retrievals, aerosol retrievals are strongly non-linear, necessitating iterative inversion schemes such as the Gauss-Newton algorithm or the Levenberg Marquardt algorithm (RODGERS, 2004).
Profile retrieval algorithms used in the scientific community are either inversion algorithms or parametrized approaches. Inversion algorithms directly link the measurement quantity to the vertical profile of the target absorber with the help of a forward model (WITTROCK et al., 2004; HENDRICK et al., 2004; WAGNER et al., 2004; FRIESS et al., 2006; WITTROCK, 2006; CLÉMER et al., 2010; WANG et al., 2017). This model is usually calculated with a radiative transfer model (RTM) assuming a selected set of atmospheric conditions. For trace gases, the model also includes a sensitivity matrix which can be considered as derivative of the measurement quantity with respect to the trace gas concentration for each altitude level. For aerosols, this sensitivity matrix is normally calculated via perturbation theory. This perturbation approach is shortly summarized as follows: by consideration of a trace gas with a well known vertical distribution such as the oxygen dimer O_2-O_2 (or short O_4), the aerosol extinction for each layer is changed gradually and the resulting modelled observations for different LOS are compared with measurements until the iteration converges (e.g. the difference between measurement and modelled quantity is small enough).
Parametrization on the other hand means, that a forward model, based on a limited set of parameters, is used to describe the measurement quantity. Frequently used forward model parameters are integrated values and profile information such as shape and height of certain absorber layers (SINREICH et al., 2005; LEE et al., 2009; LI et al., 2010; VLEMMIX et al., 2011; WAGNER et al., 2011; SINREICH et al., 2013). The forward model results are calculated with an RTM and are least squares fitted to the measurements. Generally, look up tables (LUT) are pre-calculated for a set of different scenarios, parameters and geometries covering all relevant atmospheric conditions, avoiding high computational efforts during near real time calculations.
Inversion algorithms have the advantage that they are not limited to the scenarios used when creating the LUT but unrealistic profiles are possible when the measurement, inversion, or regularization (weighting between the information from measurements and a priori informations) is poor. On the other hand, parametrized approaches evaluate profiles much faster, as the slow forward model computation was already done when creating the LUT.

Although efforts for deriving concentration profiles from MAX-DOAS measurements have been made for more than one decade, profiling is still a difficult task and results of different algorithms can differ strongly when the absorber of interest is highly variable and inhomogeneous on spatial and temporal scales (ZIEGER et al., 2011; VLEMMIX et al., 2015; FRIESS et al., 2016). Comparison studies of different profiling algorithms for synthetic as

well as for real data summarizing the current state of the art and including results from the algorithm introduced here will be reported by Frieß et al. (2018) and Tirpitz et al. (2018).
The purpose of this study is the introduction of IUP Bremen's new MAX-DOAS profile retrieval algorithm for aerosols and trace gases BOREAS (Bremen Optimal estimation REtrieval for Aerosols and trace gaseS) which has been developed to improve on the earlier profile retrieval algorithm (WITTROCK, 2006). BOREAS uses a novel approach for the retrieval of aerosols but a similar optimal estimation technique for the retrieval of trace gases. In contrast to perturbation based inversion algorithms, BOREAS uses the change (from an a priori state) of depth in an absorption band to get information on the aerosol content which caused this change. To the authors' knowledge, this approach has never been used within an operational profiling algorithm and it complements the variety of methods with another promising technique. The development of BOREAS aimed at several key properties:

1. Flexibility: the algorithm should retrieve aerosol and trace gase profiles from any MAX-DOAS measurement with pre-filtering options for the data.

2. Accuracy/stability: the algorithm should be stable in terms of varying atmospheric conditions when retrieving profiles for several years of data. The profiling results (modelled observation) should fit the measured observations with a high accuracy.

3. Automation: the retrieval should respond to problematic data/settings (e.g. low information content, wrong regularization) automatically with an included problem solution scheme.

4. Fast: the algorithm should be fast enough to allow near real time profile retrievals.

4.2 Retrieval algorithm description*

The profile retrieval algorithm BOREAS was developed in order to retrieve aerosol and trace gas vertical profiles from MAX-DOAS measurements. The aerosol retrieval is fully implemented within the RTM SCIATRAN (ROZANOV et al., 2014) to decrease computation time for the iterative minimization scheme. BOREAS is a Python written analysis script which calls SCIATRAN for the aerosol retrieval and for calculations of BAMF matrices which are then used within an optimal estimation based trace gas retrieval. A flow chart depicting BOREAS is shown in Figure 4.1. In the next subsections, we give an overview of the individual steps of the algorithm.

* This section has already been published in BÖSCH et al. [2018].

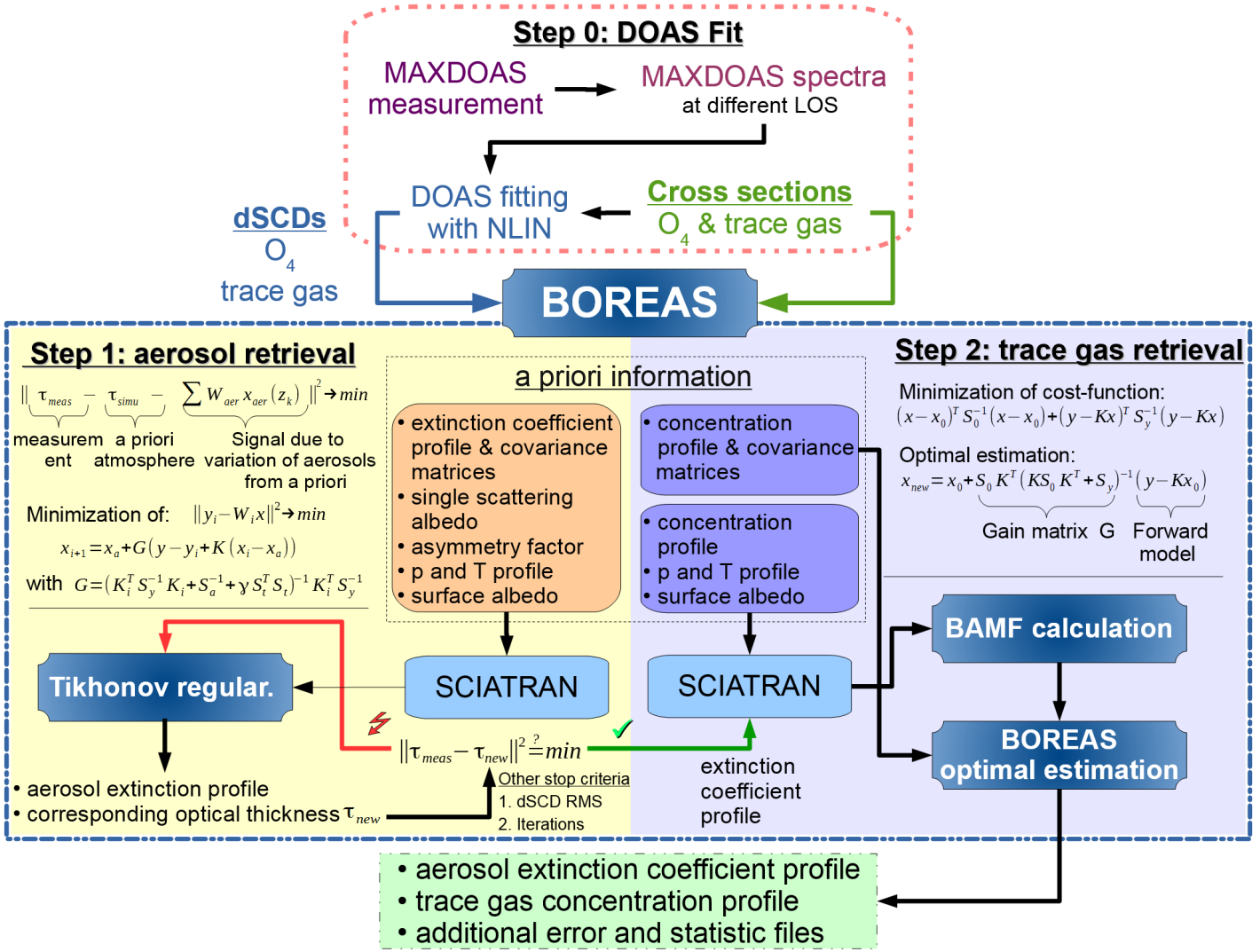


Fig. 4.1: BOREAS analysis flow chart. The algorithm is separated into two steps. Step 1: Aerosol retrieval within a Tikhonov-Regularization. Step 2: Optimal estimation based trace gas retrieval (BÖSCH et al., 2018).

4.2.1 Retrieval of aerosol profiles*

The standard DOAS fit does not provide direct information about aerosols present in the atmosphere. However, the scattering and absorption properties of aerosols have an impact on the measured differential slant column density $\Delta S(\mathbf{\Omega})$ because scattering processes can significantly modify the light path. In general, the absorption of photons by aerosols plays a small role in contrast to scattering effects, which lead to the modification in photon path length. The single scattering albedo ω (SSA)

$$\omega = \frac{\sigma_s}{\sigma_e} = \frac{\sigma_s}{\sigma_s + \sigma_a} \tag{4.1}$$

quantifies the ratio of scattering σ_s to total extinction efficiency σ_e. Here, σ_e is the sum of scattering and absorption efficiency $\sigma_e = \sigma_s + \sigma_a$. For urban pollution, SSA values are in the range of 0.90 to 0.99 for the visible spectral range (see e.g. DUBOVIK et al., 2002) and they stay more or less constant when the aerosol type does not change over time and altitude. Therefore, the SSA is not a quantity to be retrieved in BOREAS and is kept constant with typical values for the prevalent aerosol type (e.g. urban pollution: $SSA = 0.92$). The angular scattering distribution of aerosols is fully quantified by the scattering phase function. Within BOREAS, we use the Henyey-Greenstein approximation (HENYEY and GREENSTEIN, 1941) with constant values for the asymmetry factor g which quantifies the amount of forward and backward scattering (e.g. urban pollution: $g = 0.68$). In this parametrization, the optical properties of aerosols are fully defined with g, SSA and an extinction coefficient profile $\sigma_e(z)$. The latter is the retrieval parameter for BOREAS aerosol retrievals. Note, that the usage of measured phase functions will be implemented in BOREAS in the future, since Henyey-Greenstein is in some situations not an accurate representation of the atmospheric aerosol scattering distribution.

If $\sigma_e(z)$ is known, then the aerosol optical thickness (AOT) can be determined as $\tau_{aer} = \int_0^H \sigma_e(z)dz$, where the integration is performed over the entire atmosphere. If the vertical profile of an absorber number density is known, differences between modelled and measured S for this absorber are the result of differences between the assumed and real aerosol profile. As observations at different LOS have varying sensitivity to the presence of aerosols at different altitudes, this can be used to retrieve an aerosol profile.

Generally, the vertical profiles of species in the troposphere are unknown because of the temporal and spatial variability of emission sources, transport and conversion processes. However, the oxygen monomer O_2 is in this respect an exceptional species because it only depends on pressure and temperature. Furthermore, as the oxygen dimer O_4 concentration is proportional to the squared monomer concentration, its profile is exponentially decreasing with altitude as well. The O_4 slant column density can easily be determined, because O_4 has spectral absorption features in the wavelength regions of most DOAS fits.

* This section has already been published in BÖSCH et al. [2018].

Detailed sensitivity studies of O_4 $\Delta S(\mathbf{\Omega})$ measurements regarding changes in atmospheric aerosol properties can be found in WAGNER et al. [2004] and FRIESS et al. [2006].
In the BOREAS aerosol retrieval algorithm, the difference between modelled and measured O_4 differential slant optical thickness's $\Delta\tau(\lambda, \mathbf{\Omega})$ (dSOT) is used to retrieve aerosol extinction profiles in an iterative process. The measured $\Delta\tau(\lambda, \mathbf{\Omega})$ is calculated using Eq. (2.52), where $\Delta S_i(\mathbf{\Omega})$ is the dSCD retrieved in the framework of the standard DOAS fit, and the simulated O_4 dSOT ($\Delta\tilde{\tau}(\lambda, \mathbf{\Omega})$) is calculated as follows:

$$\Delta\tilde{\tau}(\lambda, \mathbf{\Omega}, N_a(z)) \;=\; \ln \frac{I(\lambda, \mathbf{\Omega}, N_a(z))}{I_{ref}(\lambda, N_a(z))} \;, \tag{4.2}$$

where $I(\lambda, \mathbf{\Omega}, N_a(z))$ and $I_{ref}(\lambda, N_a(z))$ are the intensities calculated using SCIATRAN under the assumption that O_4 is the only absorber. The dependency on the reference geometry $\mathbf{\Omega}_{ref}$ is summarized with the index ref and will be neglected from now on. The vertical profiles of pressure and temperature are used according to the US standard atmosphere model (NASA, 1976) but can be replaced by measured atmospheric conditions when available. The aerosol loading is described using an a priori concentration number density vertical profile $N_a(z)$.
The inverse problem with respect to the aerosol optical depth is then formulated as

$$\Big\| \Delta\tau(\lambda, \mathbf{\Omega}) \;-\; \Delta\tilde{\tau}(\lambda, \mathbf{\Omega}, N_a(z)) \;-\; P(\lambda, \mathbf{\Omega}) \Big\|^2 \;\longrightarrow\; \min \;, \tag{4.3}$$

where $P(\lambda, \mathbf{\Omega})$ is a polynomial of lower order and its argument $\mathbf{\Omega}$ emphasises that polynomial coefficients depend on the LOS angle. The assumptions made in this formulation are that 1) the O_4 absorption derived from the measurements does not depend on the concentration of other trace gases and 2) that the optical depth in an atmosphere without other absorbers can be described as the sum of the O_4 optical depth and a low order polynomial. Since $\Delta\tilde{\tau}(\lambda, \mathbf{\Omega}, N_a(z))$ is a non-linear and complicated functional of $N_a(z)$, an analytical solution of this minimization problem does not exist. To simplify the solution of the minimization problem given by Eq. (4.3), let us consider the variation of dSOT caused by the variation of the aerosol number density. The variation of dSOT in a linear approximation can be represented in the form of the following functional Tailor series (ROZANOV and ROZANOV, 2010):

$$\Delta\tilde{\tau}(\lambda, \mathbf{\Omega}, N_a(z)) \;=\; \Delta\tilde{\tau}(\lambda, \mathbf{\Omega}, \bar{N}_a(z)) \;+\; \int_0^H W(\lambda, \mathbf{\Omega}, z)\, \delta N_a(z)\, dz \;+\; \varepsilon_{lin}(\lambda) \;, \tag{4.4}$$

where

$$W(\lambda, \mathbf{\Omega}, z) \;=\; \left. \frac{\delta\, \Delta\tilde{\tau}(\lambda, \mathbf{\Omega}, N_a(z))}{\delta\, N_a(z)} \right|_{\bar{N}_a}$$

is a functional derivative of dSOT with respect to the aerosol number density profile $N_a(z)$ around the initial guess $\bar{N}_a(z)$. $W(\lambda, \mathbf{\Omega}, z)$ is usually referred to as weighting function and

$\varepsilon_{lin}(\lambda)$ is the linearisation error. It follows that the weighting function provides a linear relationship between the variation of dSOT and the variation of an aerosol number density N_a around the initial guess $\bar{N}_a$. Substituting now Eq. (4.4) into (4.3) and approximating the integral with a finite sum, we have

$$\left\| \Delta\tau(\lambda,\mathbf{\Omega}) - \Delta\tilde{\tau}(\lambda,\mathbf{\Omega},\bar{N}_a(z)) - \sum_{k=1}^{L} \mathcal{W}(\lambda,\mathbf{\Omega},z_k)\,(x_k-\bar{x}_k) - P(\lambda,\mathbf{\Omega}) \right\|^2 \longrightarrow \min\,. \quad (4.5)$$

Here, $\Delta x = x_k - \bar{x}_k = N_a(z_k) - \bar{N}_a(z_k)$, L is the number of altitude levels, and $\mathcal{W}(\lambda,\mathbf{\Omega},z_k)$ is defined to satisfy the trapezoidal integration rule.
The solution of the minimization problem given by Eq. (4.5) is performed on a discrete wavelength grid. In fact, only five wavelengths over the O_4 absorption band centred at e.g. 477 nm are used.
The minimization problem can be reformulated in the following vector-matrix form:

$$\left\| \Delta\mathbf{y} \;-\; \mathbf{K}\Delta\mathbf{x} \right\|^2 \;\longrightarrow\; \min\,. \quad (4.6)$$

Here, the vector $\Delta\mathbf{y} = \mathbf{y} - \bar{\mathbf{y}} = |\Delta\mathbf{y}_1, \Delta\mathbf{y}_2, ..., \Delta\mathbf{y}_N|^T$ has the dimension $N \cdot M \times 1$ and describes the difference between measured and simulated dSOT at N LOS angles and M wavelengths. The m-th element of vectors $\mathbf{y}_n$ and $\bar{\mathbf{y}}_n$ is the differential slant optical depth of O_4 at wavelength m, given by

$$\{\mathbf{y}_n\}_m \;=\; \Delta\tau^{-}(\lambda_m,\mathbf{\Omega}_n)\;,\; m \;=\; 1,\ldots,M\;, \quad (4.7)$$

$$\{\bar{\mathbf{y}}_n\}_m \;=\; \Delta\tilde{\tau}^{-}(\lambda_m,\mathbf{\Omega}_n,\bar{N}_a(z))\;, \quad (4.8)$$

where the superscript "-" denotes that a polynomial is subtracted. The state vector $\Delta\mathbf{x}$ has the dimension $L \times 1$ and matrix $\mathbf{K}$ has the dimension $N \cdot M \times L$, i.e., the matrix consists of $N \cdot M$ rows and L columns. The r-th row of matrix $\mathbf{K}$ is then given by

$$\{\mathbf{K}\}_r \;=\; \left\{\mathcal{W}^{-}(\lambda_m,\mathbf{\Omega}_n,z_1), \mathcal{W}^{-}(\lambda_m,\mathbf{\Omega}_n,z_2), \ldots, \mathcal{W}^{-}(\lambda_m,\mathbf{\Omega}_n,z_L)\right\} \quad (4.9)$$

with $r = (n-1)M + m$ and contains weighting functions for the m-th wavelength, n-th LOS in L layers.
The minimization problem given by Eq. (4.6) is solved employing an iterative Tikhonov regularization technique (RODGERS, 2004):

$$\mathbf{x}_{i+1} = \mathbf{x}_0 + (\mathbf{K}_i^T\mathbf{S}_y^{-1}\mathbf{K}_i + \mathbf{S}_0^{-1} + \gamma\mathbf{S}_t^T\mathbf{S}_t)^{-1}\mathbf{K}_i^T\mathbf{S}_y^{-1}(\mathbf{y} - \mathbf{y}_i + \mathbf{K}_i(\mathbf{x}_i - \mathbf{x}_0)) \quad (4.10)$$

Here, i is the iteration number, $\mathbf{S}_0$ and $\mathbf{S}_y$ are the a priori and measurement covariance matrices, $\mathbf{K}_i$ is the weighting function matrix calculated using the estimated aerosol number density profile $\mathbf{x}_i$, $\mathbf{S}_t$ is the first order derivative matrix with Tikhonov parameter γ

and $\mathbf{x}_0 = \bar{\mathbf{x}}$ and $\mathbf{y}_0 = \bar{\mathbf{y}}$ are the a priori profile and the representing a priori measurement vector, respectively.
There are three criteria to stop the iteration process:

- Convergence in parameter space, i.e., the maximum difference between the components of the state vector at two subsequent iterative steps does not exceed the selected criterion (e.g. 0.0001 km^{-1}).
- The root mean square difference between measured and simulated O_4 dSOT is less than selected (e.g. 0.001).
- The maximum number of iterations is reached.

In addition to the equations above, we introduce here other quantities which are useful to describe the retrieval. The gain matrix

$$\mathbf{G}_i = (\mathbf{K}_i^T \mathbf{S}_y^{-1} \mathbf{K}_i + \mathbf{S}_0^{-1} + \gamma \mathbf{S}_t^T \mathbf{S}_t)^{-1} \mathbf{K}_i^T \mathbf{S}_y^{-1} \tag{4.11}$$

describes the sensitivity of the solution to the measurement. Note, that this formulation includes also the Tikhonov term, leading to differences for $\gamma \neq 0$ with the definition given in RODGERS [2004]. The averaging kernel $\mathbf{A}_i = \mathbf{G}_i \mathbf{K}_i$ (AK) characterizes the sensitivity of the solution to the true state. The trace of AK quantifies the degrees of freedom of the signal $d_s = \mathrm{tr}(\mathbf{A})$ (DOF). This quantity is commonly understood as the number of individual pieces of information which, can be retrieved.

4.2.2 Retrieval of trace gas concentration profiles*

Compared to aerosols, the inverse problem for trace gases is easier to solve because, under the assumption of an optically thin atmosphere, the relationship between trace gas concentration and measured differential slant column density is linear. Then, the forward model $\mathbf{F}(\mathbf{x}, \mathbf{b})$ is equal to a set of measurements $\mathbf{y}$,

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \epsilon \tag{4.12}$$

where ϵ includes the error of measurement and forward model. $\mathbf{F}(\mathbf{x}, \mathbf{b})$ depends on the retrieval quantity vector $\mathbf{x}$ (trace gas concentration profile) and on an additional parameter vector $\mathbf{b}$. The latter one includes quantities which have an impact on the measurement and are known with some accuracy (e.g. a priori information).
From the perspective of MAX-DOAS, this relationship is strongly ill-posed as the number of retrieval parameters is usually much higher than the number of measurements. Furthermore, as the light paths of two geometrically close elevation angles traverse similar

* This section has already been published in BÖSCH et al. [2018].

vertical layers near the surface, measurements cannot be considered as independent. Consequently, the retrieval parameter vector $\mathbf{x}$ is not fully constrained by the input vector $\mathbf{y}$ which introduces the need for additional a priori knowledge which constrains the solution. With the assumption of a Gaussian error distribution, the optimal estimation method (OE) is often used to prevent unstable solutions when dealing with ill-posed problems (RODGERS, 2004).

$$\begin{aligned} \mathbf{x}_n &= \mathbf{x}_0 + (\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} + g^{-1}\mathbf{S}_0^{-1})^{-1}\mathbf{K}^T\mathbf{S}_y^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x}_0) \qquad \text{or} \\ &= \mathbf{x}_0 + \mathbf{S}_0\mathbf{K}^T(\mathbf{K}\mathbf{S}_0\mathbf{K}^T + g\mathbf{S}_y)^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x}_0) \end{aligned} \tag{4.13}$$

Again, $\mathbf{S}_0$ and $\mathbf{S}_y$ are the covariance matrices, $\mathbf{y}$ is the measurement and $\mathbf{x}_0$ is the a priori profile. Note that we introduced a scaling factor g which gives the possibility of regulating the weighting between a priori and measurement information. Here, the matrix $\mathbf{K}$ consists of BAMF values for every layer and measuring geometry (Eq. (2.56)) instead of weighting functions. The covariance matrix of measurements has only diagonal elements which are the absolute errors of the spectral DOAS fit of this absorber. $\mathbf{S}_0$ has constant variances on its diagonal elements but consists of additional non-diagonal elements based on a Gaussian distribution which accounts for a possible correlation of different profile layers (BARRET et al., 2002).

4.3 Sensitivity study with synthetic data

The analysis in this section is done on synthetic datasets which were created by using SCIATRAN (v.3.8.4). The RTM performed in its spherical geometry mode using multiple scattering and atmospheric refraction but without polarisation, the consideration of an opening angle of the telescope, and additional Raman scattering. Further settings are summarized in Table 4.1. The calculations were done on $25\,\mathrm{m}$ grid steps from the surface up to $6\,\mathrm{km}$, $250\,\mathrm{m}$ steps up to $10\,\mathrm{km}$ and $1\,\mathrm{km}$ steps up to $60\,\mathrm{km}$. The output quantities are O_4 and NO_2 ΔS with a zenith reference geometry. More LOS values were selected for lower elevation angles where the sensitivity is expected to be the highest. This set of LOS should be considered as a representation of a real measurement. Pressure and temperature profiles

parameter	value
trace gases	O_4, NO_2
wavelength (nm)	477, 461
albedo	0.06
SZA(°)	40, 60, 80
RAA (°)	0, 90, 180
LOS (°)	1, 2, ..., 6, 8, 15, 30
asym. fac.	0.68
SSA	0.92
climatology	US standard

Tab. 4.1: *SCIATRAN settings for the calculation of differential slant column densities.*

were taken from a US standard atmosphere (NASA, 1976). In general, this is a poor climatology in comparison to real conditions. However, since synthetic datasets and the retrieval were done with the same atmospheric profiles, the results should be consistent. The Henyey-Greenstein parametrization is used with a fixed asymmetry factor and single scattering albedo. The albedo was kept fixed as well. SZA and RAA were chosen as a representation of typical values for real measurements. Although the BOREAS algorithm can perform retrievals in different wavelengths regions, the synthetic study is focused on the $477\,\mathrm{nm}$ O_4 absorption band and the NO_2 retrieval in the visible fitting window (VIS, compare Tab. 3.3). Note that all datasets were created without adding noise.

4.3.1 Sensitivity of the aerosol retrieval

For the aerosol retrieval, different extinction coefficient profiles were used within the RTM calculation of ΔS. Five exponentially decreasing profiles, 3 Box-profiles as well as 17 different Gaussian shaped profiles are shown in Figure 4.2, with matching profile shape parameters in Table AT.2 and AT.3 in the appendix A.2.

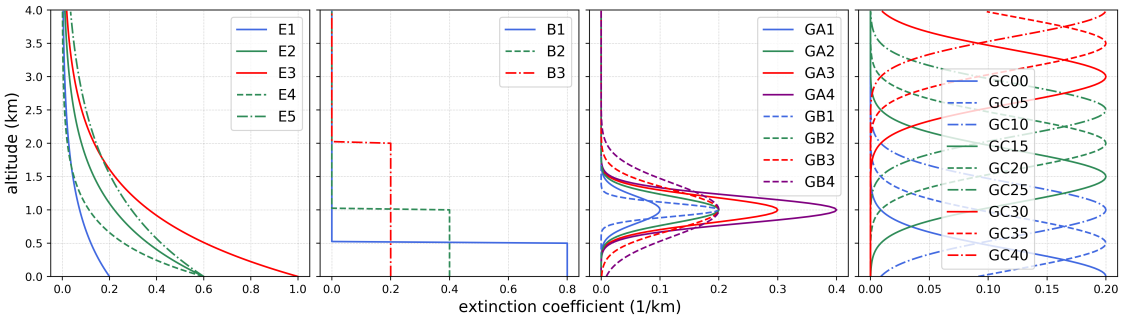


Fig. 4.2: *Aerosol extinction coefficient profiles used for the creation of synthetic ΔS datasets.*

These true profiles were chosen to test the retrieval for ideal situations (e.g. a priori profile with similar shape as true profile), and unideal conditions (edgy boxes as well as narrow Gaussian profiles should be hard to retrieve). The Gaussian profiles with different maximum altitudes will be used, in addition to the sensitivity discussion in Section 4.3.1.3, to quantify the height depending sensitivity of the retrieval algorithm.

4.3.1.1 Optimal aerosol retrieval settings

It was already stated that the retrieval via Equation 4.10 depends on many a priori parameters, which have to be chosen carefully in order to retrieve an *optimal* profile.
In this sense, *optimal* might be interpreted differently, depending on the retrieval user, the true profile and the measurement conditions. Here, *optimal* is referred to as an improvement in the total AOT of the matching profile, the degrees of freedom, the bottom

extinction coefficient value, the RMS between measurement and modelled dSOT as well as between true and retrieved profile, and the overall profile shape. Since the latter point is hard to quantify it will be only considered as an intuitive point of view without a real scientific meaning.
As a short summary of Section 4.2.1, the retrieval depends on the following a priori parameters:

- a priori extinction coefficient profile $\mathbf{x}_0$
- a priori covariance-matrix $\mathbf{S}_0$
- measurement covariance-matrix $\mathbf{S}_y$
- Tikhonov parameter γ
- vertical grid steps and maximum retrieval height
- number of LOS used for one profile
- actual values of LOS angles
- retrieval a priori mode (previous, a priori or zero)
- convergence criteria and iteration limit
- atmospheric climatology profiles

The first four points are obvious as they can be considered as main parameters of the solution scheme:
The a priori profile is often assumed to be a mean state of the atmosphere, e.g. calculated by averaging a set of ancillary measurements (RODGERS, 2004). Unfortunately, the availability of such measurements is limited introducing the need for estimating a priori profiles based on the experience of the profiler or by comparing with other studies and measurements.
The definition of the a priori covariance matrix focuses on the knowledge of the accuracy of the a priori profile. Here, it will be considered as a relative quantity whose actual value depends strongly on the measurement covariance matrix and the smoothing constraint via Tikhonov parameter. In general, smaller a priori covariances or larger Tikhonov parameters mean that the solution profile will be more constraint in the direction of the a priori profile. This can be prevented by decreasing the measurement covariances. The dependencies of these parameters are crucial and will be discussed in detail at the end of this chapter in **Test 7**.
The optimal number of grid steps and the maximum retrieval height are a direct result of the vertical sensitivity of the retrieval. In addition to the vertical grid, the sensitivity also depends on the number of elevation angles and the actual value of each LOS. The dependency of the retrieval result on the vertical grid will be discussed in **Test 2**. The importance of a proper LOS selection is shown in **Test 3**.
The retrieval a priori mode can be set to *previous*, *a priori* or *zero* within BOREAS. The difference is that in the first case, the previous profile will be used as a priori profile for the

next iteration. In contrast to that, iterations from one a priori profile only is the commonly used approach in the DOAS community (e.g. FRIESS et al., 2006; CLÉMER et al., 2010; VLEMMIX et al., 2015). In addition, a zero a priori profile can be selected which means, that the starting point of the iterations is an aerosol free atmosphere. The outcome of the usage of all modi is shortly analysed in **Test 4** and discussed in detail in **Test 7**, together with the regularization optimization.
The second last point of the bullet list above are the convergence criteria which should usually have a small impact on the result. When the retrieval parameters are defined so that the solution is close to a minimum in solution space than convergence should be reached fast and stricter convergence criteria will not change the resulting profile strongly. A discussion on this matter can be found in **Test 5**. The last test in this subsection deals with the vertical resolution of the retrieval which depends on several parameters (**Test 6**). Additionally, discussions on the vertical resolution can be found in **Test 7** during the analysis of height depending Gaussian shaped profiles.
The last bullet point was included because it was shown by WANG et al. [2017] that the actual pressure and temperature might differ strongly from a fixed climatology as the US standard atmosphere. This has an impact on the O_4 concentration and, therefore, on the retrieved aerosol profile. There will be no synthetic data test on this matter within this thesis but a simple approach is introduced and shortly discussed in Section 4.4.
As a start, a perfect retrieval is considered in order to find the maximum amount of information retrievable by BOREAS. Additionally, further quantities are introduced which are of importance for the following tests.

Test 1: Maximum number of pieces of information and retrieval errors

In Section 4.2.1, the degrees of freedom d_s were introduced as a quantity describing the information content of the signal. In addition to that, the *information content* H was defined by SHANNON [2001] as follows:

$$H = S(P_1) - S(P_2) = -0.5 \ln |\mathbf{I} - \mathbf{A}| \tag{4.14}$$

Here, $S(P) = -\sum_i p_i \ln p_i$ is the entropy for information systems with the probability p_i of state i so that the information content is defined as the entropy $S(P_1)$ (depending on the probability density function[1] P_1) before the measurement minus the entropy after the measurement $S(P_2)$. In the second part of the above equation, a Gaussian distribution was assumed as P and the vertical bars indicate the determinant of the matrix. H can now be calculated with help of the averaging kernel matrix **A** (and the identity matrix **I**).

[1] If there is a real function f for all $a \in \mathbb{R}$ so that $p(X \leq a) = \int_{-\infty}^{a} P(x)\mathrm{d}x$, then P is the probability density function of p. Definition from wikipedia: `https://de.wikipedia.org/wiki/Wahrscheinlichkeitsdichtefunktion`, Date: 19.07.2018.

From a theoretical perspective, the averaging kernel matrix is considered as ideal when it equals the identity matrix. In this case, the trace of **A** is exactly the number of rows (or columns) and therefore, the number of retrieval grid levels. Whereas this simple consideration makes sense for d_s, Equation 4.14 is not defined for $\mathbf{I} = \mathbf{A}$.
This ideal situation is only possible, when the measurement is perfect, so that its covariance matrix $\mathbf{S}_y \to \mathbf{0}_{mn}$. Since a perfect measurement needs neither Tikhonov smoothing ($\gamma = 0$) nor an a priori profile ($\mathbf{S}_a \gg \mathbf{S}_y$), Equation 4.10 leads to

$$\mathbf{x}_{i+1} \approx \mathbf{x}_0 + \mathbf{K}_i^{-1}(\mathbf{y} - \mathbf{y}_i + \mathbf{K}_i(\mathbf{x}_i - \mathbf{x}_0)), \tag{4.15}$$

when the matrices are invertible and transposable[2]. Thus, $\mathbf{A} = \mathbf{G}_i\mathbf{K}_i = \mathbf{K}_i^{-1}\mathbf{K}_i = \mathbf{I}$. Unfortunately, the general problem is ill-posed so that the above assumptions are usually not valid. For getting closer to a more realistic retrieval, the covariance matrices are now assumed to be identity matrices multiplied with scalar values ($\mathbf{S}_y = \epsilon\mathbf{I}$ and $\mathbf{S}_0 = \alpha\mathbf{I}$). With the neglection of Tikhonov-smoothing ($\gamma = 0$), the resulting iteration equation is:

$$\mathbf{x}_{i+1} \approx \mathbf{x}_0 + \frac{\alpha}{\alpha + \epsilon}\mathbf{K}_i^T(\mathbf{y} - \mathbf{y}_i + \mathbf{K}_i(\mathbf{x}_i - \mathbf{x}_0)), \tag{4.16}$$

Then, the averaging kernel matrix is $\mathbf{A} = \alpha(\alpha + \epsilon)^{-1}\mathbf{K}_i^T\mathbf{K}_i = \alpha(\alpha + \epsilon)^{-1}\mathbf{I}$. In other words, the degrees of freedom are decreased by the factor $\alpha(\alpha + \epsilon)^{-1}$.

In Figure 4.3, the two cases above are similarly analysed as follows:
The retrieval was started on a grid from the surface up to $4\,\mathrm{km}$ in $100\,\mathrm{m}$ steps on scenario E1 (see Fig. 4.2), with SZA = 40°, 60°, 80° and RAA = 90°. The signal to noise ratio[3] was varied from 0.2 to 2×10^6, with different a priori variances, with and without Tikhonov smoothing, colour-coded. In the four sub-plots, the degrees of freedom, the information content, the RMS of the difference of measured and retrieved dSOT, as well as the RMS of the difference of the true and retrieved profiles are shown as mean values for the three geometries.
The red curve represents the first discussed case with the maximum amount of information retrieved for the largest SNR value (top plots). The grid consists of 40 levels, meaning that the maximum d_s should also be 40. This value was nearly reached for the largest SNR but it decreases fast as the measurement is given less weight (smaller SNR). On the other hand, the information content shows a clear maximum around SNR $= 2000$. A

[2] In general, this assumption is not correct as **K** is usually not quadratic or invertible. But as a first and simple consideration, lets assume that the following rules can be applied: 1. $\mathbf{B}^T\mathbf{B} = \mathbf{I}$; 2. $\mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$; 3. $(\mathbf{BC})^{-1} = (\mathbf{C}^{-1}\mathbf{B}^{-1})$ and $(\mathbf{B}^T)^{-1} = (\mathbf{B}^{-1})^T$.
[3] Note that the SNR values should be understood as an assumed uncertainty for the measurement rather than a proper description of the measurement noise. For the synthetic dataset, no noise was added. However, the ill-posedness of the inversion problem introduces the need for appropriate measurement variances unequal to zero.

comparison with the other curves shows that already a small Tikhonov factor ($\gamma = 2$) decreases d_s drastically which is accompanied with an in turn reduction of both RMS quantities. Additional reduction of the a priori covariance from $1 \times 10^8\,\%$ to $1 \times 10^5\,\%$ and $100\,\%$ leads to lower RMS values but also reduces d_s and H. The RMS of the dSOT differences shows clear minima indicating that there is an ideal range of SNR values which is usually far away from the maximum degrees of freedom. Note that the constant RMS values for small SNR are due to the forcing of the profiling result into the direction of the a priori profile[4], when giving less weight to the measurement.

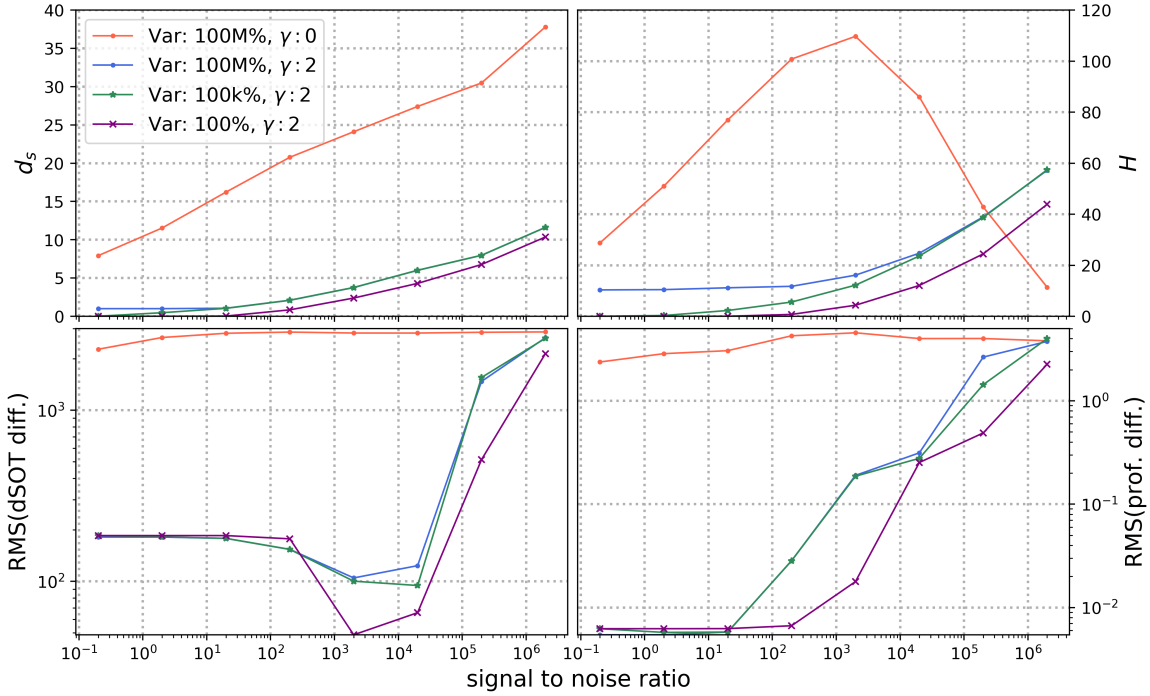


Fig. 4.3: *Degrees of freedom d_s (**top left**), Shannon information content H (**top right**), RMS of the measured and simulated dSOT difference (**bottom left**) and the RMS of the true and retrieved profiles (**bottom right**) in dependence of different signal to noise ratios, shown for four different retrieval settings.*

As a short summary it can be concluded that due to the underdetermination of the underlying problem, the maximum information content leads to unstable and oscillating retrieval results (indicated by the large RMS values). An optimal set of settings can be found with the help of the RMS quantities but this optimal range means a strong reduction in d_s and H (see also Fig. 4.12 for informations on the vertical resolution).

[4] In this test, an exponential profile was used with a scale height of $1\,\mathrm{km}$ and a bottom extinction of 0.18. Note that for this large range of regularization parameters, the actual a priori profile is of minor importance for the shape of the presented curves.

Test 2: Vertical grid

Generally, the grid steps are of minor importance as long as the grid is not too coarse, so that important profile features are smoothed due to the reduced vertical resolution. On the other hand, too fine grid steps lead to oscillations because retrieval noise is introduced and/or the inversion in Equation 4.10 fails. Furthermore, for some algorithms, the computational time is a sensitive quantity as multi-core processing (as used within BOEAS) is not possible or the algorithms are too slow for highly resolved vertical grids. In order to find the optimal range between too coarse and too fine grid steps, the following test is introduced.

The BOREAS aerosol retrieval was applied to the synthetic measurement scenarios E1, B3 and GA2 (see Fig. 4.2) with SZA = 40°, 60°, 80° and RAA = 90°, for different sets of LOS and vertical grids. For the results below, the analysed quantity is averaged for these three SZA. The a priori variance was set to $100\,\%$, while, within the measurement covariance matrix, different signal to noise ratios $(1000, 2000, ..., 9000)$ were used. The Tikhonov parameter was fixed to $\gamma = 2$. The a priori profile was exponentially decreasing with a scaling height of $1\,\mathrm{km}$ and a bottom extinction of $0.18\,\mathrm{km}^{-1}$.

In the first test, the vertical grid was constrained from surface up to $4\,\mathrm{km}$ with different grid width: $12.5\,\mathrm{m}$, $25\,\mathrm{m}$, $50\,\mathrm{m}$, $75\,\mathrm{m}$, $100\,\mathrm{m}$, $125\,\mathrm{m}$, $150\,\mathrm{m}$, $175\,\mathrm{m}$, $200\,\mathrm{m}$, $300\,\mathrm{m}$, $400\,\mathrm{m}$, $500\,\mathrm{m}$[5].

In Figure 4.4, d_s and H are shown for all profiles. Since the a priori settings stayed fixed for all retrievals and one set of LOS (1°, 2°, ..., 10°, 15° and 30°), the curves are the same for all scenarios presented in this test.

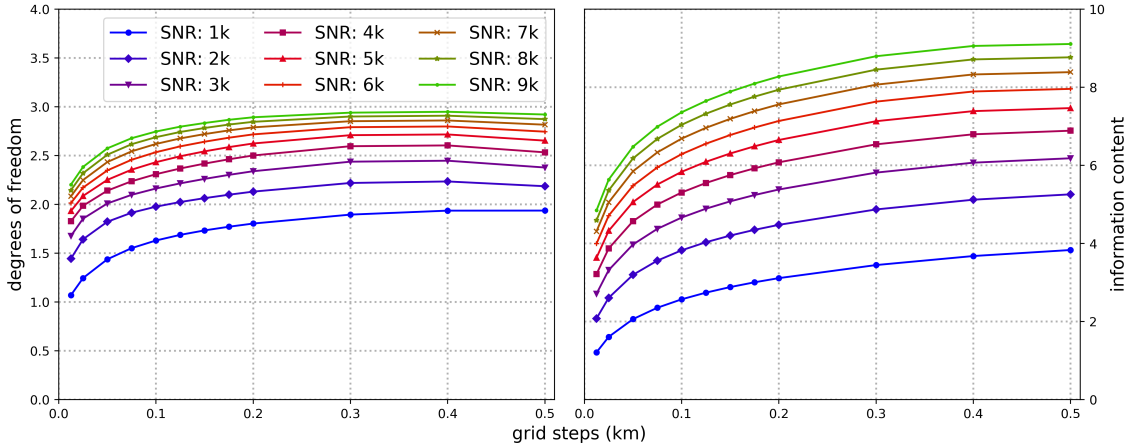


Fig. 4.4: *Degrees of freedom and information content for different signal to noise ratios (colour-coded) and grid steps with a fixed Tikhonov-parameter ($\gamma = 2$), a priori profile and covariance.*

[5] Note that the upper altitude limit might have changed because of the equidistant spacing within the boundary values (e.g. $3.9\,\mathrm{km}$ for $300\,\mathrm{m}$ grid steps instead of $4.0\,\mathrm{km}$).

Both information quantities converge for larger SNR (the greenish curves are getting closer to each other). A slight decrease for $\Delta z = 500\,\mathrm{m}$ indicates that even coarser grids might not be suitable for a successful retrieval because the coarse vertical resolution misses important profile features. On the other hand, if the grid is too fine, the information content is poor because retrieval noise might be introduced, leading to highly oscillating profiles. In general, d_s and H show similar patterns but a small decrease in d_s can be observed at $\Delta z = 500\,\mathrm{m}$ for all SNR except 1k. This can not be found for H. From this plot alone, an optimal grid step width cannot be found as the quality of the retrieval products is not presented.
In addition to the already used quality quantities in **Test 1**, the error of the profile itself is interesting. Therefore, the total error $\mathbf{S}_{tot}$ of the retrieved profiles

$$\mathbf{S}_{tot} = \mathbf{S}_{sm} + \mathbf{S}_{ns} + \mathbf{S}_{fw} \tag{4.17}$$

is introduced to quantify the accuracy of the retrieval results. It is split up into three different error sources. The *smoothing error* is defined as $\mathbf{S}_{sm} = (\mathbf{A} - \mathbf{I})\mathbf{S}_a(\mathbf{A} - \mathbf{I})^T$ and represents the error of the with **A** smoothed true state of the atmosphere[6]. The second term is the *retrieval noise* (or noise error) $\mathbf{S}_{ns} = \mathbf{G}_y\mathbf{S}_y\mathbf{G}_y^T$, with the gain matrix $\mathbf{G}_y$ (compare Eq. 4.11). It quantifies the propagation of the measurement error through the retrieval onto the retrieved profile. The last term is called *forward model error* $\mathbf{S}_{fw}$, and it means the error of the accuracy of the forward model in describing the true atmosphere. In real measurements, this error is hard to quantify as the true state of the atmosphere is usually unknown. In general, this error source can be neglected, when the forward model is estimated well.

In Figure 4.5, RMS values (as mean of the three SZA) are depicted, for relevant quantities and for the three scenarios introduced above.
In the three rows, the RMS values of the differences of 1. measurement and simulated differential slant optical thickness, 2. true and retrieved profile, and 3. the RMS of the smoothing error S_{sm} and retrieval noise S_{ns} is shown. In the first column, RMS values for E1 are presented (top left). This profile was chosen to be close to the a priori profile. It can be seen that the measurement (dSOT) is described best for $\Delta z >= 100\,\mathrm{m}$ and $\Delta z <= 300\,\mathrm{m}$, depending on SNR. The grid step width with the highest information content ($\Delta z = 400\,\mathrm{m}$) shows one of the worse results. This behaviour can be confirmed by looking at the RMS of the difference between the true and the retrieved profile. Here, the smallest grid steps lead to the best profiles. The RMS of the error sources show clear maxima for nearly all SNR values and error types. Note that the smoothing error is ge-

[6] Note that in general, the actual smoothing error is defined with the covariance matrix of the true state of the atmosphere $\mathbf{S}_t$. Since this is mostly unknown, it is assumed that the a priori state is a representation of an appropriate ensemble of states close to the true conditions.

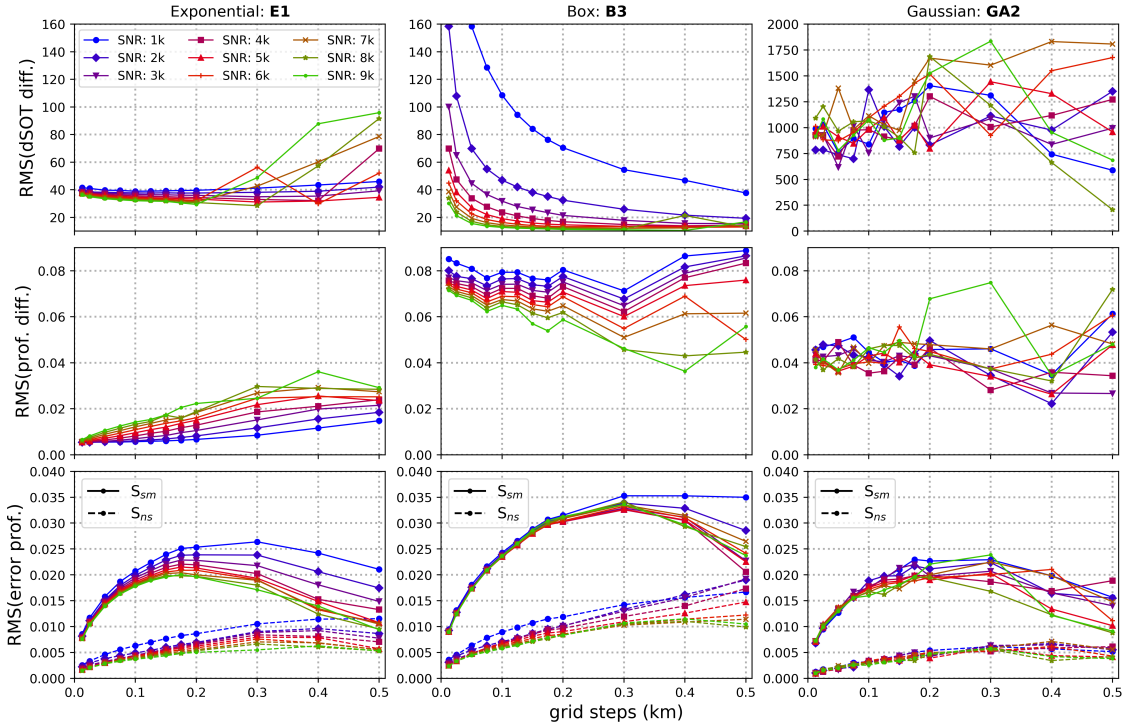


Fig. 4.5: *RMS values for different quantities with varying SNR values colour-coded and for three different scenarios (E1, B3, GA2) depending on the grid steps width.* ***Top row****: RMS of the difference of retrieved and simulated dSOT;* ***Mid row****: RMS of the difference of the true and retrieved profile;* ***Bottom row****: RMS of the smoothing (*$\boldsymbol{S}_{sm}$*) and noise error (*$\boldsymbol{S}_{ns}$*) profiles.*

nerally larger than the noise error due to the limited amount of information within the measurement rather than the measurement noise itself.

Together with Figure 4.4, a clear tendency for the best grid step width of E1 is hard to find. The RMS values show the lowest uncertainties for 1. a width range between $\Delta z >= 100\,\mathrm{m}$ and $\Delta z <= 300\,\mathrm{m}$ (RMS(dSOT diff.)), 2. the lowest Δz values (RMS(prof. diff) and RMS(error prof.)) while the information content shows the maximum amount of information for the largest width.

The second column of Figure 4.5 leads to other conclusions. The RMS of dSOT differences is now the smallest for the largest grid step width. In addition to that, the RMS of the profile differences shows varying results with clear minima at $\Delta z = 75\,\mathrm{m}$ and $\Delta z = 175\,\mathrm{m}$. The shape of the error RMS is similar to E1 but the overall values are larger. When comparing the first two columns with the last one (Gaussian at $1\,\mathrm{km}$ altitude), it can be seen that all results are more or less unstable indicating a bad retrieval. The last point is due to the fact, that the sensitivity for higher aerosols is lower (see **Test 6** and **Test 7** for further details) so that the applied settings might not be optimal.

The conclusion of the grid step width analysis is that there is no such thing as perfect grid. It depends strongly on the true aerosol profile but also slightly on the proper regulariza-

tion (ratio of a priori smoothing/covariance and measurement covariance). Furthermore, a comparison of the RMS values and the information content shows that the best grid seems to be a trade-of between the highest information content and the lowest RMS values. Here, possible minima in RMS of dSOT and profile differences indicate local minima in the solution space. Considering that an accurate profile is usually more important than some artificial information content definitions, a grid step width between $100\,\mathrm{m}$ and $200\,\mathrm{m}$ shows small RMS values without introducing too much retrieval noise. Since the retrieval noise RMS is larger for higher Δz and the smoothing error RMS is close to a maximum for $\Delta z = 200\,\mathrm{m}$, within this thesis, a grid step width of $\Delta z = 100\,\mathrm{m}$ is used.

To end up the discussion of the vertical grid, Figure 4.6 shows the dependence on the maximum retrieval height with $100\,\mathrm{m}$ grid steps widths. A constant SNR $= 2000$ was used with the same settings as in the test above but for three different sets of LOS (see **Test 3** for details on the LOS sets). It can be seen that the RMS of the dSOT differences shows varying behaviour depending on the scenario. For the exponential profile (E1), a higher altitude leads to smaller RMS values while for the box profile an increase in the maximum retrieval height shows larger differences. Note that this is to be expected as the exponential profiles for the calculation of synthetic data went up to $6\,\mathrm{km}$ whereas the B3 and GA2 cases are limited to a certain altitude range. Furthermore, since the a priori profile is automatically calculated until the maximum retrieval height by BOREAS, an exponentially decreasing profile means a forcing to zero (or the a priori profile) on more altitudes for B3 and GA2 while for E1, the algorithms needs to adapt the resulting profile into the other direction. In contrast to E1 and B3, the Gaussian shaped profile leads, as before, to highly variable results. The RMS of dSOT differences shows no clear ten-

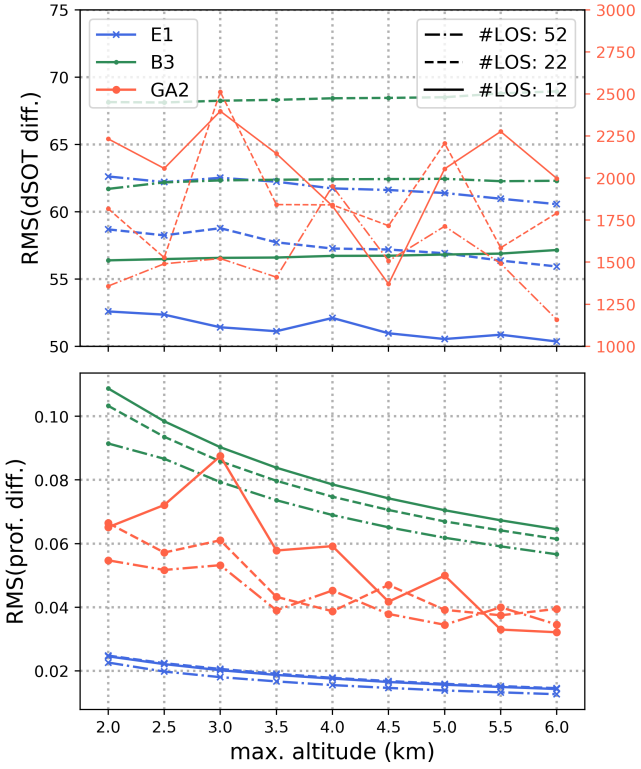


Fig. 4.6: *RMS of different quantities depending on the maximum retrieval altitude, for the scenarios E1, B3 and GA2.* ***Top****: RMS of the difference of retrieved and simulated dSOT;* ***Bottom****: RMS of the difference of the true and retrieved profile. The line style indicates the set of LOS with 52 (pointed dashes), 22 (dashed) and 12 LOS (solid).*

dency to a certain maximum altitude. On the other hand, the RMS of the profile shapes shows decreasing values for higher upper grid limits with a local maximum for LOS12 at $3\,\mathrm{km}$. A larger number of LOS used within the retrieval means always a lower profile RMS but a varying tendency for the dSOT RMS. For the latter RMS quantity, E1 has the lowest RMS for 12 LOS and the highest for 52. B3 shows again the lowest RMS for 12 LOS but the highest for 22. The RMS of the error profiles and d_s (not shown) stayed more or less constant for all retrieval heights.
As a summary, the maximum retrieval height depends strongly on the real profile and should comprise the most important profile features. Here, a maximum altitude of $4\,\mathrm{km}$ will be used because the differences in profiling accuracy for exponential profiles is small but Box and Gaussian shaped profiles are located within the lowest $4\,\mathrm{km}$.

Test 3: Number and spacing of LOS angles

To my knowledge, the optimal number of LOS angles and the spacing of each angle for MAX-DOAS profile retrievals have not been discussed by the DOAS community so far. For imaging DOAS instruments, these quantities are fixed as the LOS angle spacing is constrained by the optical features of the telescope. For 2D MAX-DOAS instruments, the pointing to several elevation angles is possible showing the need for a detailed analysis.

In this test, the retrieval was started with the following settings: $\gamma = 2$, a priori Variance was $100\,\%$, SNR $= 2000$, a priori profile used as above and the vertical grid was again set to $100\,\mathrm{m}$ steps from the surface up to $4\,\mathrm{km}$. The geometries were kept constant with a RAA of $90°$ and three SZA angles ($40°$, $60°$ and $80°$), which were averaged for the calculation of the discussed quantities.
Three different sets of LOS angles were used with the additionally included $15°$ and $30°$ elevation angle for all sets but with a different spacing from $0°$ to $10°$. 52 LOS angles mean a $0.2°$ spacing, 22 a $0.5°$ spacing and 12 numbers of LOS a $1°$ spacing between $0°$ and $10°$. The $1°$ LOS spacing for lower elevation angles can be understood as a representation of MAX-DOAS measurements, commonly used in the community. Each LOS angle is considered as a single beam. A discussion about the impact on an opening angle (or field of view, FOV) can be found later in this test sub-section.

The first figure (4.7) shows the already introduced RMS values and d_s for the three scenarios (E1, B3 and GA2, see Fig. 4.2) and for the three sets of LOS, as a relative difference to the calculation without $0°$ LOS only, in percent. The reason for the neglection of the $0°$ is to avoid a cross dependence impact of the spectral surface reflectance and to avoid for possible problems of the RTM when pointing towards the surface (when dealing with an additional FOV).

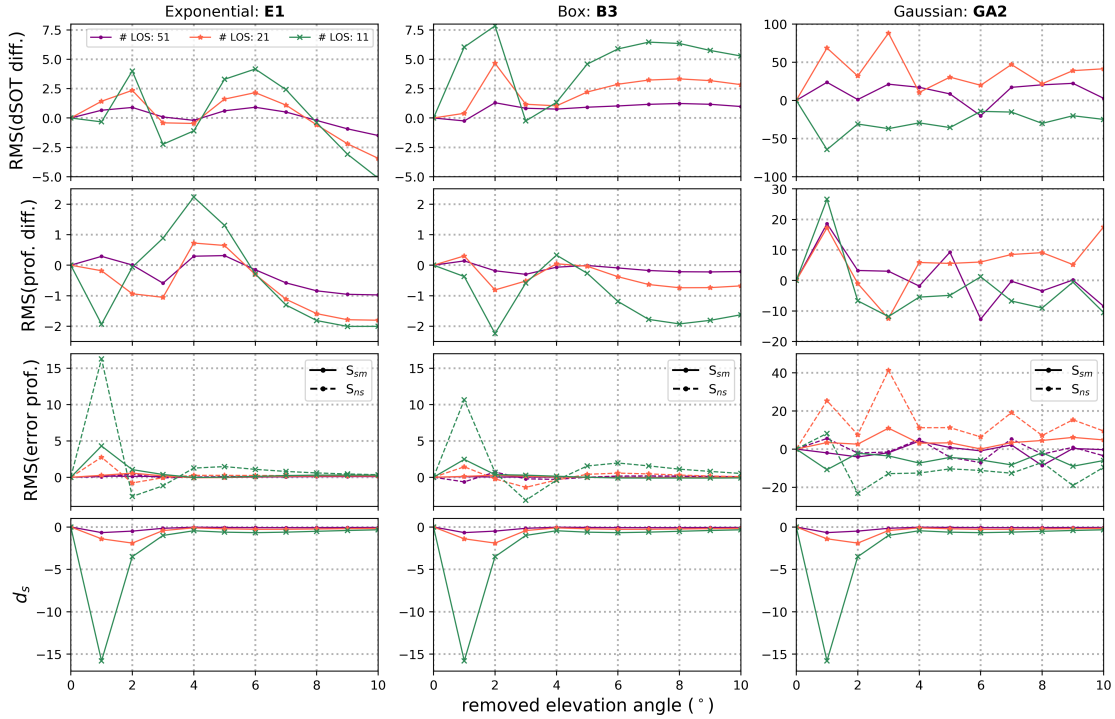


Fig. 4.7: *Relative differences of the specific quantity without* $0°$ *and a certain LOS angle (removed elevation angle) to the quantity with neglection of* $0°$ *only, for three scenarios and different sets of LOS angles, in percent. Relative difference (in %) of,* ***Top row****: RMS of the difference of retrieved and simulated dSOT;* ***2nd row****: RMS of the difference of the true and retrieved profile;* ***3rd row****: RMS of the error profiles;* ***Bottom row****: degrees of freedom.*

All sub-plots indicate that the discard of a certain LOS angle is less important when the total number of angles is larger. Starting with the exponential profile E1 (first column), all curves have the same maximum deviations for the angles $2°$ and $6°$ for the dSOT RMS. The neglections of both angles separately lead to positive deviations meaning that the retrieval results are not as good as with the angles included. In contrast to that, when removing the $3°$ angle of the LOS11 set, the relative deviation is negative, leading to a better RMS. For E1, the discard of the $10°$ LOS shows the best improvement of the RMS. In contrast to that, this RMS plot of the box profile B3 depicts the vice versa behaviour for $10°$. Here, the neglection leads to merely the same deterioration as the removal of the $2°$. For this profile, maxima can be found for $2°$ and $7°$ while an improvement of the RMS of dSOT differences is hardly possible. An exception can be found for the LOS11 and the LOS51 set, where a small negative RMS shows that the neglection of $3°$ and $1°$, respectively, would not deteriorate the result. This is in a clear contrast to the RMS of the profile differences which show an improvement for nearly all LOS angles for B3. However, the relative improvement is about $2\,\%$ (for LOS11) indicating that the effect is generally small (in comparison to the max. of $7.5\,\%$ for the RMS of dSOT differences). Nevertheless, for B3, the RMS of dSOT seems to be the more or less mirrored curve behaviour of the RMS of profile differences.

This can not be found for E1. Here, the deviations between both RMS plots are larger meaning that there are not correlated or anti-correlated changes in relative deviations for certain LOS angles. The RMS of the error profiles as well as d_s show that the neglection of the $1°$ LOS for LOS11 has the most dramatic impact, which is supported by the RMS deviation of the profile differences. Although, there is also a clear $1°$ peak for B3 for LOS11, the RMS of profiles has its maximum improvement at $2°$. An interesting feature is that the relative deviations of the RMS of the retrieval noise profile is always larger than the RMS of the smoothing error profile. This indicates that changes for this LOS might affect the retrieval more due to a high sensitivity of the retrieval to changes in the aerosol load rather than to changes in specific retrieval parameters (which have a larger impact on the smoothing error). This is supported by d_s, which show negative values, meaning that there was less information content for the neglection of the lowest LOS angles. The Gaussian shaped profile GA2 shows again chaotic behaviour. No clear conclusion can be found except for the $1°$ LOS neglection which leads to higher RMS profile values. This is interesting as it means that even for an elevated aerosol layer, the lowermost LOS angles are important.

From the test above, it can be concluded that, because of the importance of low elevation angles, a retrieval with additional LOS angles around $1°$ might also improve the results. In Figure AF.38 in the appendix, an additional LOS angle was used within the retrieval of the LOS11 set. No clear statement can be made for an additional elevation angle leading to an overall improvement of the retrieval. However, adding of LOS close to $0°$ shows larger deviations in the RMS of dSOT differences as for retrievals without this angle, indicating either the impact of the surface spectral reflectance or an inaccuracy of the RTM for near-horizontal pointing directions. For E1, an improvement of the RMS of profile differences can be found for all added elevation angles while the dSOT RMS shows a deterioration. B3 stays on a constant improvement level for the profile RMS and a more or less constant deterioration in the dSOT RMS. GA2 is again oscillating but a decrease with larger angles for the dSOT RMS and a negative relative difference of the profile RMS indicates that an improvement can be made by adding $1.8°$.

As a last test regarding elevation angles, Figure 4.8 shows a variation of the field of view (quantified as FWHM of a Gaussian distribution in $^\circ$). Synthetic datasets of scenario E1 were calculated for the LOS11 set and the retrieval was started with different FOV values (colour-coded) and with the settings from the test above. Both plots show the relative difference of the retrieved quantity and the retrieval of this scenario without FOV (in percent). As an example, the green line depicts the results for different FOV used in the forward calculation as a representation of the measuring instrument (x-axis) and a fix retrieval FOV of $0.25°$.

It can be seen that a measurement with a FOV $\neq 0$ always deteriorates the RMS of dSOT differences with a strange behaviour for a field of view of $1.0°$, $1.2°$ as well as $2.0°$. It was already stated that a line of sight pointing to the surface or to the horizon might lead to unwanted results. Figure 4.8 supports this statement as the RMS values for $1.0°$ and $1.2°$ are clearly outliers. It is interesting to see that the retrievals with FOV ≤ 1.0 do not differ much but the $1°$ retrieval shows a larger deviation. The relative difference of the RMS of profile differences shows that the profiles are slightly better. Since this problem is not as pronounced for different FOV within the retrieval it is considered as a simulation issue when calculating the synthetic datasets. Nevertheless, the remaining curve without

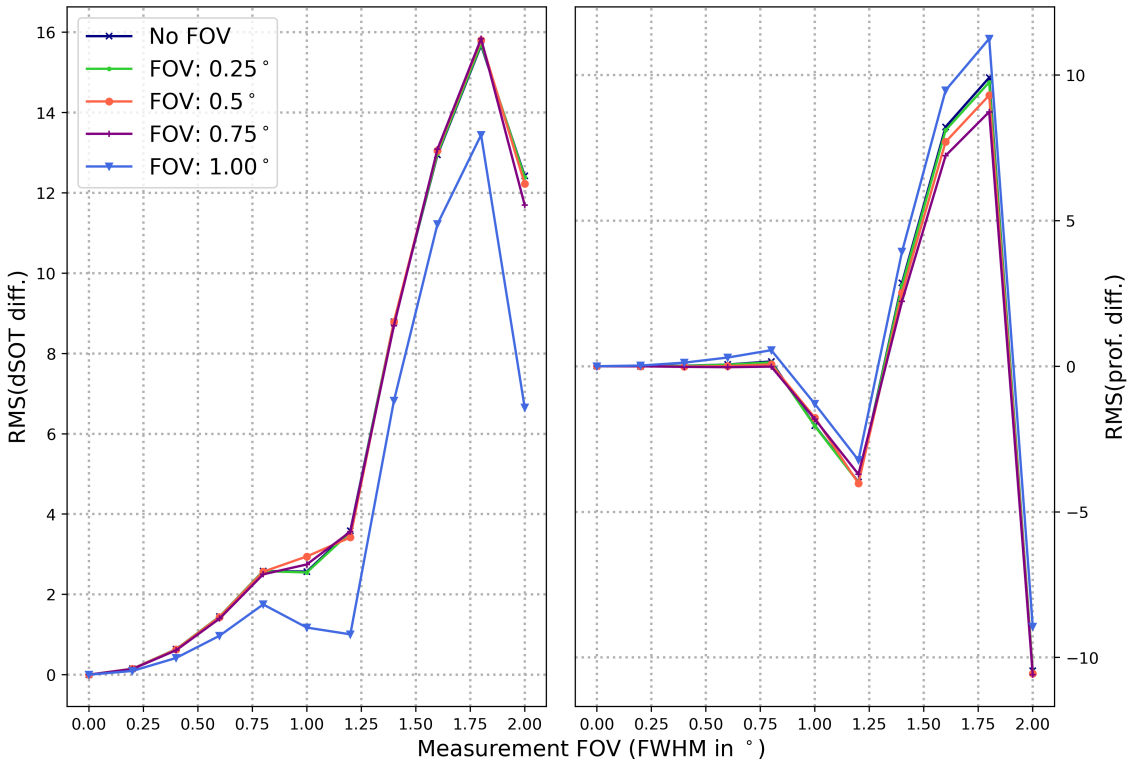


Fig. 4.8: *Retrieval test of synthetic ΔS datasets for scenario E1 calculated with various FOV angles (as FWHM in °) and with different FOV used within the profile retrievals (colour-coded). Shown are relative differences of a retrieval with a specific FOV and without i (in %),* ***Left****: for the RMS of the differences of real and simulated dSOT,* ***Right****: for the RMS of the differences of true and retrieved profile.*

these outliers show that a wider FOV of the instrument decreases the profiling results most likely due to the deterioration of the ill-posedness of the mathematical problem when nearby elevation angles cover similar air masses. This was also found by BRUNS [2004] for airborne limb measurements. The inclusion of a specific FOV within the retrieval has a smaller impact and can be neglected.

As a conclusion of the LOS discussions, the discard of certain LOS angles has a lower impact on the retrieval when the overall numbers of LOS is high. For the commonly used approach of measurements in 1° steps, the removal of a lower elevation angle might introduce large uncertainties depending on the true profile. Additional LOS for the lowest elevation angles can improve the profile when the aerosol load is close to the surface but pointing to the horizon or the surface might have a negative effect on the retrieval result. For the latter point, the effect of the measurement itself, the spectral surface reflectance and the accuracy of RTM are not fully analysed but should be considered in future studies.

Test 4: A priori mode

BOREAS can be started with different a priori modes having various effects on the resulting profiles and their quality criteria. The standard mode is referred to as *apri* here, meaning that the iteration of Equation 4.10 is done with one a priori profile only (see e.g. RODGERS, 2004). The a priori profile is then considered as a good representation of the atmosphere, for example when using an annual mean profile (with the assumption of a typically low temporal aerosol variability) or by using modelled profiles. When the true aerosol conditions vary only slightly around this a priori profile, the retrieval should easily calculate an accurate a posteriori solution. However, when the true aerosol profile shape or the total AOD is far away from the a priori profile, the assumption of the a priori as an accurate representation of the atmosphere is not valid. In these cases, the accuracy of the a posteriori solution depends on the capability of the retrieval algorithm in compensating for the larger gap between measured and simulated atmosphere.
The second retrieval mode uses the previous solution as a priori profile for the next iteration. This mode is referred to as *prev* in this thesis. In Equation 4.10, every a priori profile x_0 is replaced by x_i, meaning that the weighting function matrix $\mathbf{K}_i$, the simulated measurement $\mathbf{y}_i$ as well as the Tikhonov-matrix $\mathbf{S}_t$ are calculated around the previous profile (with $\Delta x = x - x_i$ instead of $\Delta x = x - x_0$, cf. Eq. 4.6). This mode was implemented in order to reach solutions which are far away from the original a priori profile. This approach usually needs more iterations for convergence, in comparison to the *apri*-mode. The solutions are often *smoother* than those retrieved from one a priori profile only but also oscillating solutions are possible. The actual outcome of the retrieval depends stronger on the Tikhonov parameter γ as the results of the *apri*-mode do. This is due to the calculation of the first derivative Tikhonov matrix $\mathbf{S}_t$ around each subsequent profile $\Delta x = x - x_i$ rather than one a priori profile $\Delta x = x - x_0$. For example, if γ is chosen to be small, oscillations are possible due to the propagation of noise into the retrieval which might be amplified when using *prev*. On the other hand, when γ is large, each subsequent solution will be smoothed again. This smoothness comes with the price of a possible out-smoothing of specific profile features like peaks or large near-surface values.

RODGERS [2004] described iterations from the previous profile as *a popular mistake* without giving a mathematical proof for this opinion. However, a common agreement can be found by assuming that the a priori profile, in the case of iterations with the prev-mode, are not considered as an *optimal estimator*. As already stated, a proper estimate of the true atmosphere is sometimes hard to find so that iterations in the *prev*-mode might converge where the *apri*-mode might have failed. Especially when the a priori profile is far away from the true atmosphere, a reasonable solution cannot always be found when using certain a priori constraints.
The third retrieval mode is referred to as *zero* within this thesis. In Equation 4.10, the two explicitly written x_0 variables are considered to be zero. However, $\mathbf{K}_i$, $\mathbf{S}_t$ and $\mathbf{y}_i$ are calculated around the initial a priori profile (unequal to zero). This mode was implemented in SCIATRAN's retrieval branch in order to retrieve inversion products for really small true retrieval quantities.

Figure 4.9 shows the retrieval results for the three before-mentioned modes for scenarios E2, E3 and GA2 (cf. Fig. 4.2). Similar settings as in the tests above were chosen and the different geometry results are again averaged (cf. settings in e.g. **Test 1**).
On the left-hand side, it can be seen that the results for E2 are better than those for E3, meaning that the retrieved bottom extinction coefficient is closer to the true one, for all retrieval modes. The difference between the individual retrieval modes can clearly be identified. The *zero*-mode enables the algorithm to retrieve a surface-near profile, where the main extinction is expected. The other altitudes are retrieved to be zero as the measurement does not provide sufficient informations for higher altitudes (cf. **Test 6**). For these higher altitudes, the other modes force the profile close to the a priori. Depending on the true scenario, this can be an advantage but it might also lead to problematic results, when the aerosol load is mainly located close to the surface. The *apri*-mode forces the results stronger to the a priori profile as the *prev*-mode. On the other hand, since the *prev*-mode leads to *smoother* solutions, the bottom extinction is not as well retrieved as for the *apri*-mode. The results for E3 are worse as the bottom aerosol load was not satisfactorily retrieved. The worst bottom extinction can be found for the *prev*-mode while *zero* and *apri* show similar values.
The right sub-plot shows the retrieval results for the Gaussian-shaped profile GA2. Here, the *zero*-mode has advantages as it is not forced to an a priori profile shape. Just the altitude was not properly found but the overall peak was covered the best. The other two modes lead to much smaller peaks and also to artefact extinction values close to the surface for the *apri*-mode. This indicates the problem with wrongly shaped a priori profiles which not only prevent an accurate retrieval but which might also find aerosol loads where are none are supposed to be.

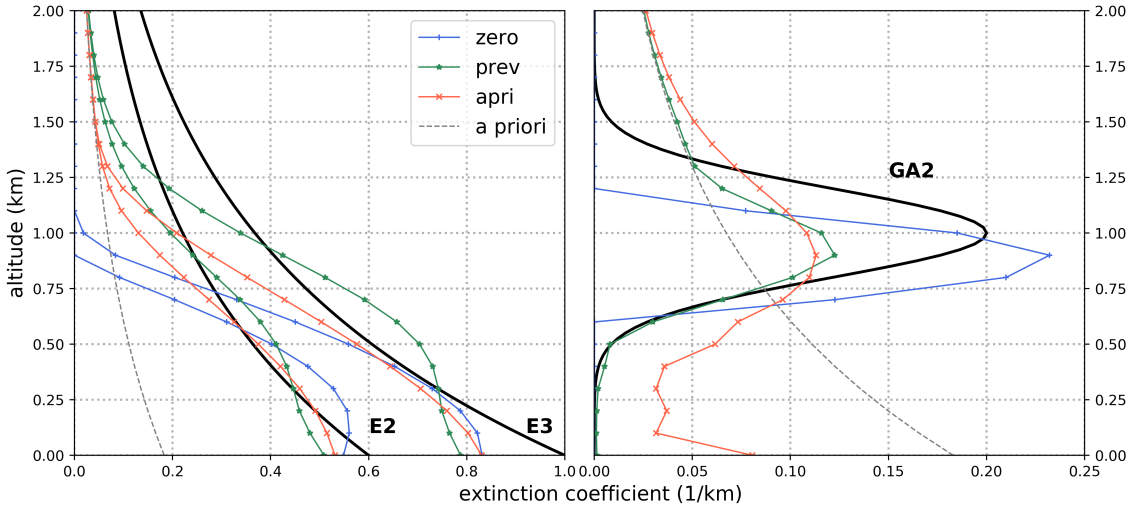


Fig. 4.9: *Retrieval results for scenarios E2, E3 (**left**) and GA2 (**right**) retrieved with different retrieval modes, colour-coded.*

Table 4.2 shows quality criteria for the profiles shown in the figure above. It can be seen that in the case of the exponential profiles, the true AOD is never reached (shown as relative deviation between retrieved and true AOD). The relative deviation is the smallest for the *prev*-mode indicating that the a priori profile is not a proper estimate of the atmosphere. For the other modes, the *apri*-mode leads to better results for the exponentially decreasing profiles while the *zero*-mode leads to a small relative deviation for GA2. These findings are supported by the RMS of the difference between simulated and measured dSOT's. The number of iterations is strongly correlated with this RMS and the AOD deviations. For GA2 and the *apri*-mode, no convergence was reached (a maximum number of 50 Iterations was used). For the exponentially decreasing profiles, *prev* and *apri*-mode show similar iteration numbers while the *zero*-mode needs more. Convergence for GA2 was reached similar fast for the *prev* and *zero*-mode.

	E2			**E3**			**GA2**		
	z	a	p	z	a	p	z	a	p
AOT (%)	-43.2	-27.9	-19.3	-49.4	-35.6	-22.0	-16.9	48.0	11.3
RMS(dSOT)	88.9	40.8	15.1	90.1	55.0	14.2	66.6	1213.9	48.2
RMS(prof.)	1.38	1.37	1.37	1.37	1.35	1.34	1.41	1.40	1.40
Iterations	30.3	10.0	10.3	31.7	12.0	16.3	17.0	50.0	21.0

Tab. 4.2: *Quality parameters for the profiles shown in Figure 4.9 averaged for all geometries. The abbreviations z, a, p describe the retrieval modes zero, apri and prev, respectively. The four rows show the relative deviation of the retrieved and true AOT (in percent), the RMS of the simulated and measured dSOT differences, the RMS of the simulated and measured profiles and the number of iterations.*

All modes and scenarios have in common that the a priori profile is the most crucial parameter for an optimal retrieval. Since the information content of a measurement is limited (cf. **Test 1, 2** and **Test 6**), a better a priori estimate is essential for an improvement of retrieval results. Furthermore, since the RMS of dSOT differences seems to be strongly correlated with the total AOT, *a priori pre-scaling* is introduced in order to find a better a priori estimate. The idea is to call the aerosol retrieval two times. Once, BOREAS is started with the *zero*-mode for finding the AOT of the prevalent aerosol profile. This AOT is used for a pre-scaling of the a priori profile which is then used within the actual aerosol retrieval. The idea of finding the AOT with the *zero*-mode is exactly opposite to the usage of an a priori profile in the second retrieval. In the first case, no profile shape constrains are allowed in order to let the retrieval be fully supported by the measurement rather than a priori parameters. In the second retrieval, the assumption of a proper a priori profile shape but an inaccurate total AOT is solved by applying the pre-scaling. Since the *zero*-mode is rather slow but an optimal solution, in the sense of an accurate profile, is not needed, the first run is done on reduced wavelength and vertical grid and with weaker convergence criteria.

Figure 4.10 and Table 4.3 show the results of a retrieval with applied pre-scaling, for similar settings and scenarios as above. It can be seen that the exponentially decreasing profiles are retrieved much better than before, when applying the *apri* or *prev*-mode. Even though, the *zero*-mode results seem similar to the results above, smaller deviations for the individual profiles were found. The Gaussian shaped profile seems to be slightly improved for *prev* but an improvement for the other modes can not be seen with the naked eye.

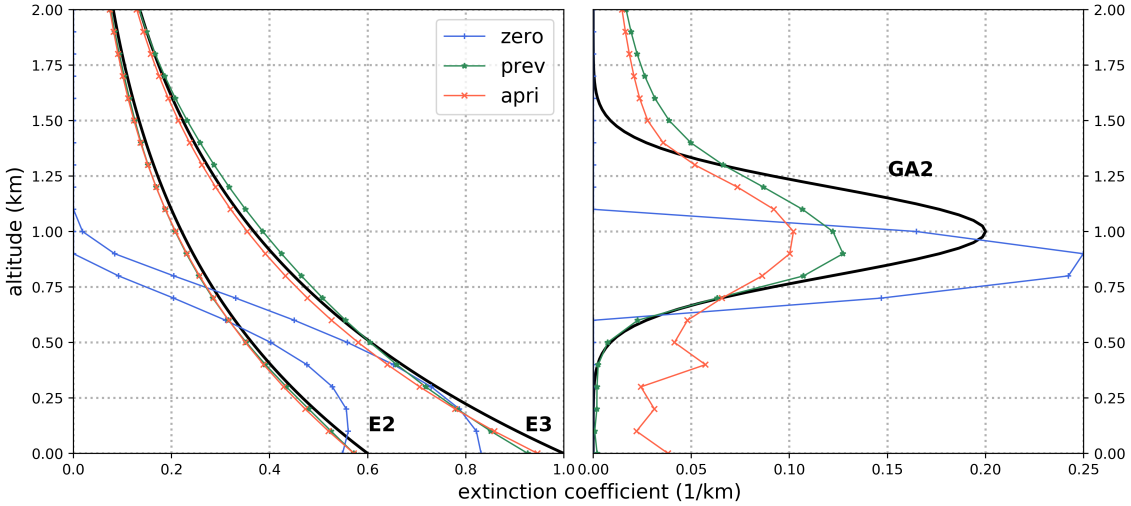


Fig. 4.10: *Retrieval results for scenarios E2, E3 (**left**) and GA2 (**right**) retrieved with different retrieval modes, colour-coded. A priori pre-scaling was applied.*

Table 4.3 summarizes again quality parameters for the retrieved profiles. In comparison to Table 4.2, all parameters are improved or stayed similar. Note that the relative deviation of the pre-scaled a priori AOT to the true AOT was added. It can be seen that the a priori AOT is already much closer to the true AOT but further improvements can be found for E2 and E3 after the main retrieval. Here, the number of iterations describes the main retrieval only. Since the pre-scaling was done on *coarse* parameters, the maximum number of iterations was set to 20 (instead of 50 for the main retrieval). This means that the total number of iterations is the sum of the pre-iterations (at maximum 20) and the table values. However, the total number of iterations is not reached often due to the weak convergence criteria so that the total retrieval time is not significantly larger than for a retrieval without pre-scaling.

	E2			**E3**			**GA2**		
	z	a	p	z	a	p	z	a	p
AOT (%)	-43.2	-7.1	-6.3	-49.4	-6.1	-2.2	-19.3	11.2	4.4
a priori AOT (%)	-10.6	-10.6	-10.6	-6.6	-6.6	-6.6	2.4	2.4	2.4
RMS(dSOT)	88.9	15.1	11.0	90.1	16.9	12.2	85.0	743.7	48.1
RMS(prof.)	1.38	1.36	1.36	1.37	1.33	1.32	1.41	1.40	1.41
Iterations	29.3	3.3	9.3	34.3	3.0	12.0	18.0	50.0	11.3

Tab. 4.3: *Quality parameters for the profiles shown in Figure 4.10 averaged for all geometries. The abbreviations z, a, p describe the retrieval modes zero, apri and prev, respectively. The five rows show the relative deviation of the retrieved and true AOT (in percent), the relative deviation of the a priori and true AOT (in percent), the RMS of the simulated and measured dSOT differences, the RMS of the simulated and measured profiles and the number of iterations of the main retrieval. A priori pre-scaling was applied.*

Note that a detailed test on all modi with and without pre-scaling for more scenarios is shown in **Test 7**.

Test 5: Convergence criteria

Three different convergence criteria are used as a stopping parameter for the iterations within the aerosol retrieval:

1. **Convergence in parameter space**: This criterion refers to the state vector (Eq. 4.8) as parameter and means a comparison of two subsequent iterations. The maximum difference of the components of the state vectors of two subsequent iterations is usually larger than a selected criterion C_p. The iteration stops as soon as one difference drops below the criterion.
2. **Convergence in the residual space**: If the RMS of the difference of measured and modelled O_4 dSOT is less than the criterion C_r, the iteration stops.

3. **Maximum iteration number**: In some cases, convergence cannot be reached for the specific scenario so that the retrieval stops after a maximum number of iterations.

The last bullet point is the most obvious one and its use needs no investigation. The maximum number of iterations is empirically chosen as a stop criterion when the both first mentioned criteria fail. It is often selected to stop after 50, 60 or 70 iterations. The exact number depends strongly on the retrieval mode (cf. **Test 4** and **Test 7**) because the distance between subsequent solutions in 1/km might vary for *prev*, *apri* and *zero* due to the individual a priori smoothing constraints and profiles so that the optimal solution is reached after less or more iterations.
The first two convergence criteria are shortly investigated within Figure 4.11. Here, two different exponential profiles (E2 and E3) were retrieved for different convergence criteria with one a priori profile only (the same as in the tests above). For each sub-plot, two x-axes are shown with one criteria changed by keeping the other criterion on a constant value of 0.001 ($0.1\,\%$). The first two sub-plots show the RMS values which were used in the other tests before while the third sub-plot represents the RMS of the total profile error instead of the individual smoothing and noise errors. The last sub-plot shows the numbers of iterations which were needed in order to achieve convergence. The variation of C_r is shown as red, dashed line whereas the results for different C_p is plotted as blue, solid lines. Note that an offset was used for the RMS sub-plots in order to show more details for the specific lines, since the actual RMS values do not matter here. In general, all RMS quantities of E3 are higher than those for E2.

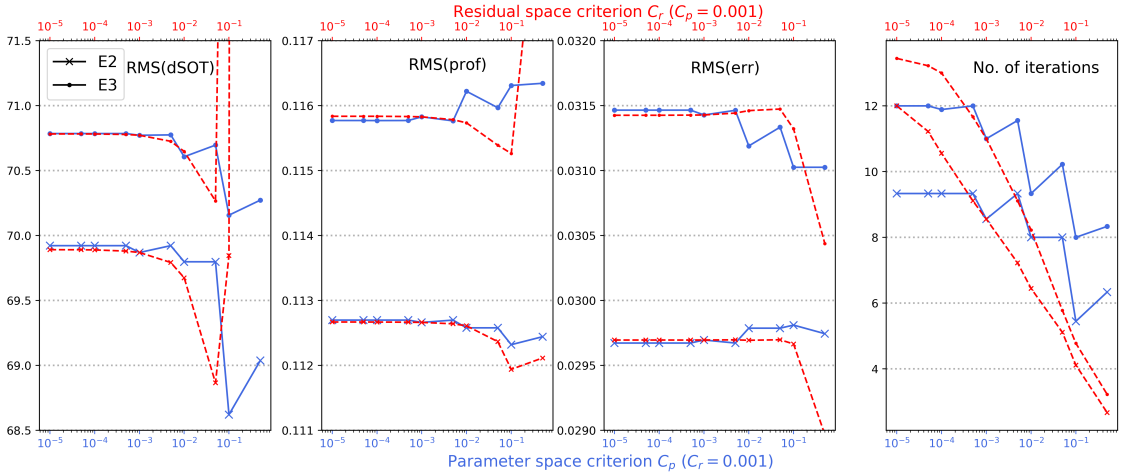


Fig. 4.11: *RMS values and number of iterations in dependence of both convergence criteria for the scenarios E2 and E3. On each x-axis, one criteria is varied by keeping the other constant.* ***1st sub-plot****: RMS of the differences between measured and simulated dSOT.* ***2nd sub-plot****: RMS of the profile differences of true and retrieved profile.* ***3rd sub-plot****: RMS of the total profile error.* ***4th sub-plot****: Number of iterations.*

It can be seen that all RMS curves converge against a certain RMS value for smaller convergence criteria. The curve behaviour for large criteria shows oscillating features. For the first RMS sub-plot, both criteria show an anti-correlation as the RMS is high for the largest C_r values while it is low for the largest C_p. For C_r, clear minima can be identified in each RMS plot. However, the minima do not coincide for the individual sub-plots indicating variable and unstable solutions. Furthermore, both scenarios show similar trends for the first RMS sub-plot but a different behaviour for the second. The large variations in all RMS sub-plots for both scenarios decrease for criteria ≤ 0.01 and enter constant values at criteria ≤ 0.001. The last sub-plot shows the number of iterations which is higher for smaller convergence criteria. Since no improvement in RMS can be found for criteria ≤ 0.001 but the number of iterations increases for larger C_r, the residual space criterion is selected to be $C_r = 0.001$. With this constant C_r, changes of C_p do not lead to more iterations indicating that the residual stopped due to C_r and not C_p. However, large RMS(prof) values for E3 at $C_p \geq 0.05$ indicate that the parameter space criterion should also be small enough. Note that these curves will differ strongly for other scenarios so that constant RMS values for smaller criteria cannot be considered as a general result. In specific cases, the convergence criteria must be varied in an empirical way to retrieve a converging aerosol profile.

Test 6: Vertical resolution

In this thesis, the vertical resolution of the retrieval is analysed by calculating the FWHM of the averaging kernels. Since $\mathbf{A} = \mathbf{GK} = \frac{\partial x}{\partial x_{true}}$ (cf. **Test 1**), the averaging kernels can be understood as the sensitivity of the retrieval to the true state x_{true} (RODGERS, 2004). The width of each kernel can also be calculated by using the spread function defined by BACKUS and GILBERT [1970] instead of the FWHM. However, this function is not defined for the lowermost kernel where the sensitivity is expected to be the highest. Here, the FWHM is calculated by finding both values which are half the maximum value for each kernel and calculating the vertical distance of both half-values. For the lowest Kernels, this value gives only half of the FWHM since the other side of the peak would be below the surface.
Figure 4.12 shows the averaging kernels for two different cases as well as the area under the kernels (calculated as the sum of each kernel at a specific altitude), and the FWHM at the nominal altitudes (colours as in the area plot) and the altitude where the actual peak has its maximum (green and purple curves, referred to as $\text{Var}_{1,2}$(act)). The first case shows the kernels for the Var: $100\,\text{M}\%$ case (here Var_1) without Tikhonov parameter ($\gamma = 0$) discussed in **Test 1**. This case was introduced as a retrieval with the maximum amount of information. Var_2 shows the kernels for a retrieval with a more realistic regularization. The true atmosphere consisted of the aerosol profile E1 (cf. Fig. 4.2) with the fix exponentially decreasing a priori profile used in the tests above. The retrieval was started

with the **zero** mode (cf. **Test 4**) for Var_1 to reduce the effect of the a priori profile shape.

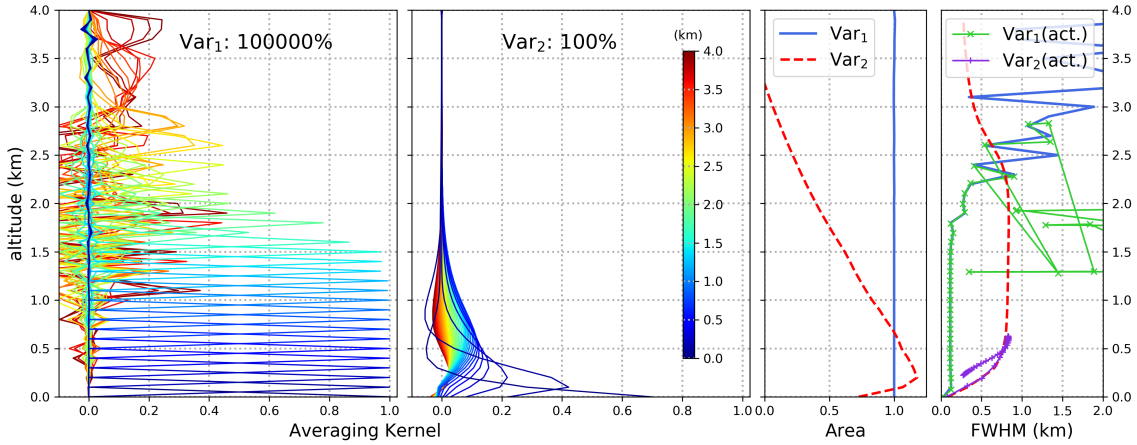


Fig. 4.12: *Averaging Kernels for two different sets of parameters (2 Figures left):* ***1st sub-plot****: Maximum information content (Var_1);* ***2nd sub-plot****: Optimal regularization ratio (Var_2).* ***3rd sub-plot****: Area of* ***A*** *for both sets of parameters.* ***4th sub-plot****: FWHM of* ***A*** *at their nominal altitudes (blue and red lines without markers) and at their actual altitudes (green and purple lines with x and + markers).*

It can be seen that the kernels for VAR_1 are close to δ-peaks for the lowermost kilometre with a strong decrease in the maximum amplitude for the next kilometre. From $2\,\mathrm{km}$ to $4\,\mathrm{km}$, noise like features begin to grow and covering the actual kernels. The uppermost altitude point is forced to zero within BOREAS in order to reduce problematic boundary oscillations in altitudes which are poorly constrained by the measurement. It is interesting to see that the ideal case of a unity matrix as **A** is achieved only for altitudes lower than $\sim 1.5\,\mathrm{km}$. Even though the kernels can not be distinguished from the noise for higher altitudes, the area under the curves is always 1 for Var_1. RODGERS [2004] explained that the area can be thought of as "the fraction of the retrieval that comes from the data". In this sense, the retrieval with Var_1 is taking everything from the data and nothing from the a priori information. On the other hand, the FWHM is the smallest for the lowermost $\sim 1.5\,\mathrm{km}$ with a high variability for higher altitudes. For Var_1, the nominal altitudes of the averaging kernels equal the actual altitudes up to $\sim 1.7\,\mathrm{km}$. For higher altitudes, the noise prevents a proper calculation of the kernel maxima. In contrast to these results, Var_2 shows much smaller peak amplitudes and finite kernel widths. Only the lowermost peak has still a large amplitude indicating that the sensitivity is the highest for the lowest retrieval altitude. The area shows values close to 1 for altitudes smaller than $1\,\mathrm{km}$, indicating that the retrieval uses mostly the measurement rather than the a priori information close to the surface. The higher the altitude, the smaller the area meaning an increased fraction of a priori informations was used for the calculation of the a posteriori solution. The FWHM (at the nominal heights, red) is again the smallest closest to the surface but does not immediately stay constant for medium altitudes as for Var_1. The vertical resolution

decreases for the first $500\,\mathrm{m}$ from $80\,\mathrm{m}$ to $\sim$790 m and is then more or less constant up to $\sim 2.3\,\mathrm{km}$. The actual altitudes (purple) for the FWHM values show a different behaviour. Here, the curve is similar to the nominal altitude FWHM curve up to $500\,\mathrm{m}$ but the actual peak height decreases after that showing a discrepancy between nominal and actual altitude. This is due to the negative part of the kernel curves for higher nominal altitudes indicating that e.g. an increase of the aerosol extinction coefficient profile in this altitude would decrease the retrieved profile close to the surface. Note that the FWHM values for higher altitudes are getting smaller not due to an increased sensitivity but because of two distinct peaks in the kernel curve (one positive and one negative) affecting the FWHM calculation. By using a Tikhonov value $\gamma \neq 0$, another a priori smoothing parameter is introduced which decreases the vertical resolution by broadening **A** in the vertical dimension. Furthermore, the maximum amplitude of each kernel is strongly reduced implying that the sensitivity is lower even though the retrieved profile might be more accurate. However, the best looking averaging kernel do not necessarily lead to an optimal profiling result. Highly oscillating profiles can be expected without an additional smoother so that quantities like vertical resolution and degrees of freedom should be treated carefully.

Test 7: Regularization constraints

Within this test, the proper regularization ratio between a priori smoothing and measurement constraints is investigated by changing the Tikhonov parameter γ and the signal to noise ratio, gradually. The parameters of the synthetic datasets are noted in Table 4.1 and the grid was set to $100\,\mathrm{m}$ steps from the surface up to $4\,\mathrm{km}$. γ was varied in the following interval with 15 values in total: $0.5, 1, 2, ..., 5, 10, 15, ..., 50$. The SNR interval was set from 500 to 5000 in 500 steps (10 values). In general, the SNR ratio is fully quantified by the measurement i.e. one SNR of the CCD for a specific exposure (integration time). However, due to varying intensities throughout the day (e.g. when pointing towards a bright cloud or to the sun), dynamic ranges of exposure and integration times are used in order to measure in an ideal saturation regime of the CCD. Furthermore, because of signal averaging of the second dimension of the CCD, high SNR are possible showing the need for an investigation of a large SNR range. However, the SNR, as input of the measurement covariance matrix, should be assumed as a representing error quantity of the dSOT input rather than the exact SNR value of the specific intensity spectrum.
In Table 4.4, the sets of parameters changed within this test are noted, together with the abbreviations used for an easier denotation in the following analysis. The first column represents the a priori mode which was already described in **Test 4**. Here, only the modes *apri* and *prev* are used for the main retrieval while the zero mode is applied for the prescaling of the a priori profile. The second column describes whether a constant a priori variance of $100\,\%$ was used or a height depending variance, with $150\,\%$ at the surface, $100\,\%$

at $1\,\mathrm{km}$, $50\,\%$ at $2\,\mathrm{km}$ and $1\,\%$ at $3\,\mathrm{km}$ to $4\,\mathrm{km}$. The variances in-between are linearly interpolated on the matching retrieval grid points. The question behind the altitude depending variance is, if there is a general improvement of the bottom extinction values possible, when giving the retrieval more freedom for the lowest altitudes (by increasing the error covariances of the a priori profile). The last column in Table 4.4 denotes the before mentioned a priori pre-scaling option (p) or the usage of a fixed exponential profile (f, with SH $= 1\,\mathrm{km}$, AOD $= 0.18$). As an example, "acp" means that the retrieval mode was set to *apri*, the variance was constant and a priori pre-scaling was applied.

mode	**variance**	**a priori prof.**
apri, a	constant, c	fixed, f
prev, p	alt. depending, z	pre-scaled, p

Tab. 4.4: *Abbreviation table showing the specific parameters and the matching abbreviations used within this section.*

In Figure 4.13, results are presented for the most simple set of parameters (acf: apri mode with a constant variance and one fixed a priori profile) for the profiles E1 - E5 and B1 - B3. The represented colour-coded parameters are mean values of all geometries (cf. Tab.4.1) and are from the left to the right:

- **BOT**: The relative difference of the retrieved and true bottom extinction coefficient value (in %).
- **AOD**: The relative difference of the retrieved and true aerosol optical depth (in %).
- **RMS(dSOT diff.)**: The absolute RMS values of the difference of retrieved and measured dSOT.
- **RMS(prof. diff.)**: The absolute RMS values of the difference of retrieved and true profile (in 1/km).
- **DOF**: The degrees of freedom for the specific SNR/γ ratio.

Note that the profile E2 is depicted in row 5 and 8 to show the transition from E1 to E3 (same scale heights) and E4 to E5 (different scale heights), respectively (cf. Fig. 4.2). Starting the discussion with the first column, it can be seen that for the profile B3 nearly all SNR/γ pairs show a positive value range (red) while all other scenarios find an underestimation of the bottom extinction (blue). This underestimation is the smallest for the scenario E1 which is close to the a priori profile and it is getting larger when the true profile is farer away (E2 and E3). Furthermore, when considering the true profiles with the same bottom extinctions but different scale heights (E4, E2 and E5), varying relative differences show that some profile bottom values are harder to retrieve due to the different profile shapes in higher altitudes. The worst retrieval result was achieved for B1 which is the thinnest Box profile with the aerosol load concentrated close to the surface. This can

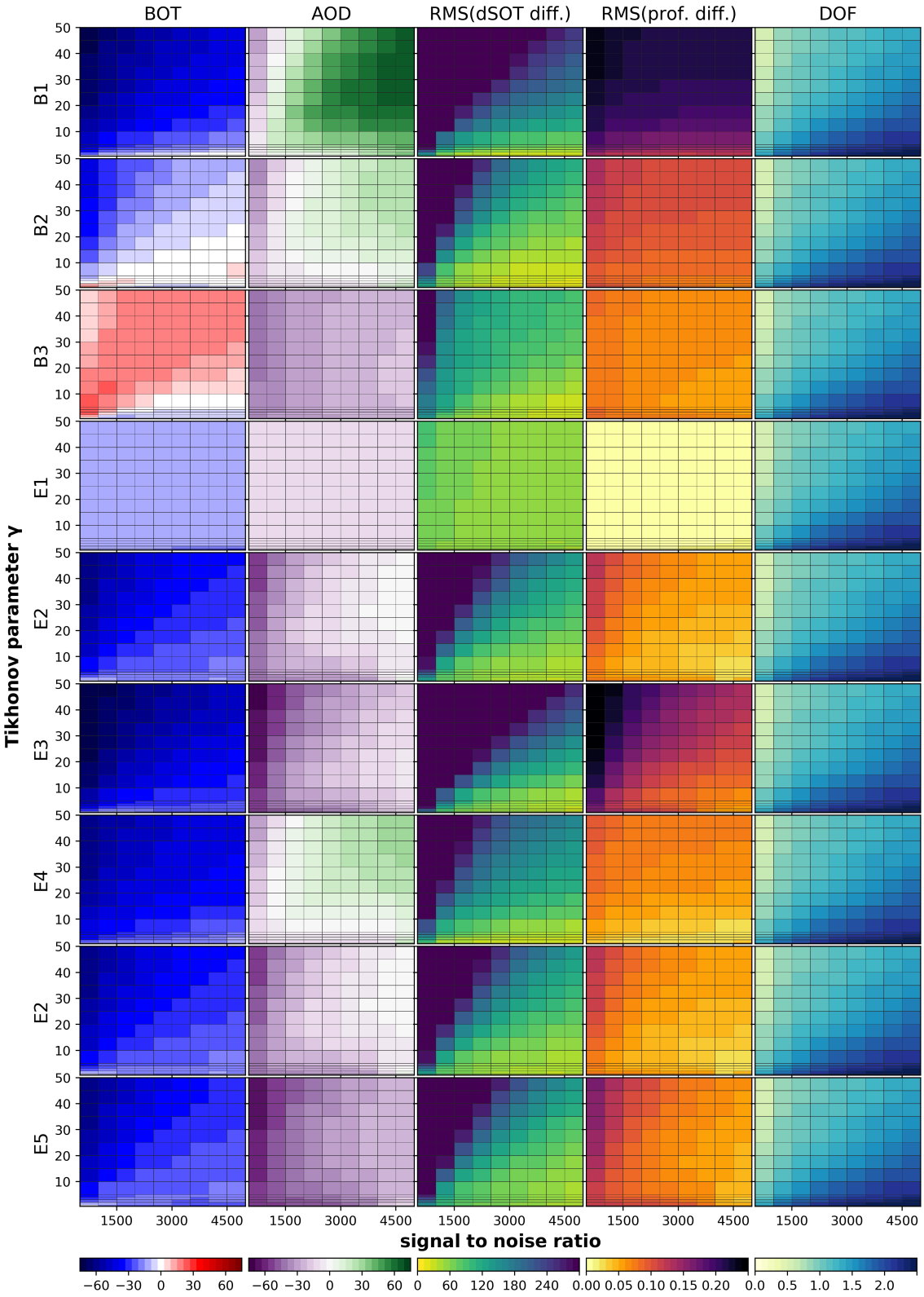


Fig. 4.13: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOD (in %).* ***3rd, 4th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: degrees of freedom.* ***Retrieval mode****: apri, const., fixed a priori profile (see Tab. 4.4)*

be explained by the a priori profile shape, which is completely wrong for this scenario. The AOD column supports some of the findings of the bottom extinction discussion but contradicts other. Some ranges of SNR/γ (e.g. SNR = 1000, $\gamma = 35$) lead to a low deviation to the true AOD for the profile B1 where the relative difference of the bottom extinction was found to be large. On the other hand for E3, low Tikhonov values and small SNR (e.g. SNR = 1000, $\gamma \leq 1$) show relative differences of $\sim 40\,\%$ for the AOD while the bottom extinction coefficient is good ($\leq 10\%$). These findings already imply an important conclusion which can be found for many profiles and parameter settings: Depending on the true profile shape, there might exist a range of SNR and γ values which improve the AOD of the retrieved profile and there might be a range which improves the bottom extinction coefficient value. But these SNR/γ ranges do not necessarily have to be similar. When comparing the RMS columns it can be seen that for some profiles (E3, E5) there are SNR/γ pairs which deteriorate both RMS quantities. Note that small SNR and large γ values force the retrieval solution into the direction of the a priori profile. This can be seen as the upper left corner of each RMS subplot shows the largest RMS values, except for E1 which is already close to the a priori. For most profiles, low Tikhonov values lead to the best adaptation to the bottom extinction which is accompanied with both RMS quantities being small. Note that d_s is similar for all profiles with minimum values close to 0 and maximum values up to 3. However, as shown in **Test 1**, the largest d_s does not necessarily lead to the best result. For example B3, E1 or E5 show larger BOT deviations for SNR ≥ 6000 and low γ values, where d_s is the highest.

Figure 4.14 shows the same settings as above but the retrieval mode *prev* was used. It can be seen that the bottom extinction value deviations are usually smaller than compared to the *apri* results. For the *apri*-mode, the AOD colour-coding was mainly dominated by shades of purple. In contrast, the *prev*-mode shows also greenish areas. The transition between purple and green show a good agreement between retrieved and true AOD which is more often found for iterations from the previous profile. Both RMS quantities are either smaller or similar to the values of the *apri*-mode showing that the agreement between retrieved and true profiles is in average better for the *prev*-mode. However, this does not say anything about the best as possible solution of each retrieval mode. Furthermore, there are still many scenarios for which either the AOD or the bottom extinction is good. A common good SNR/γ range can not be found. Notably, the problem of a deterioration for profiles which are far away from the a priori profile is still present.

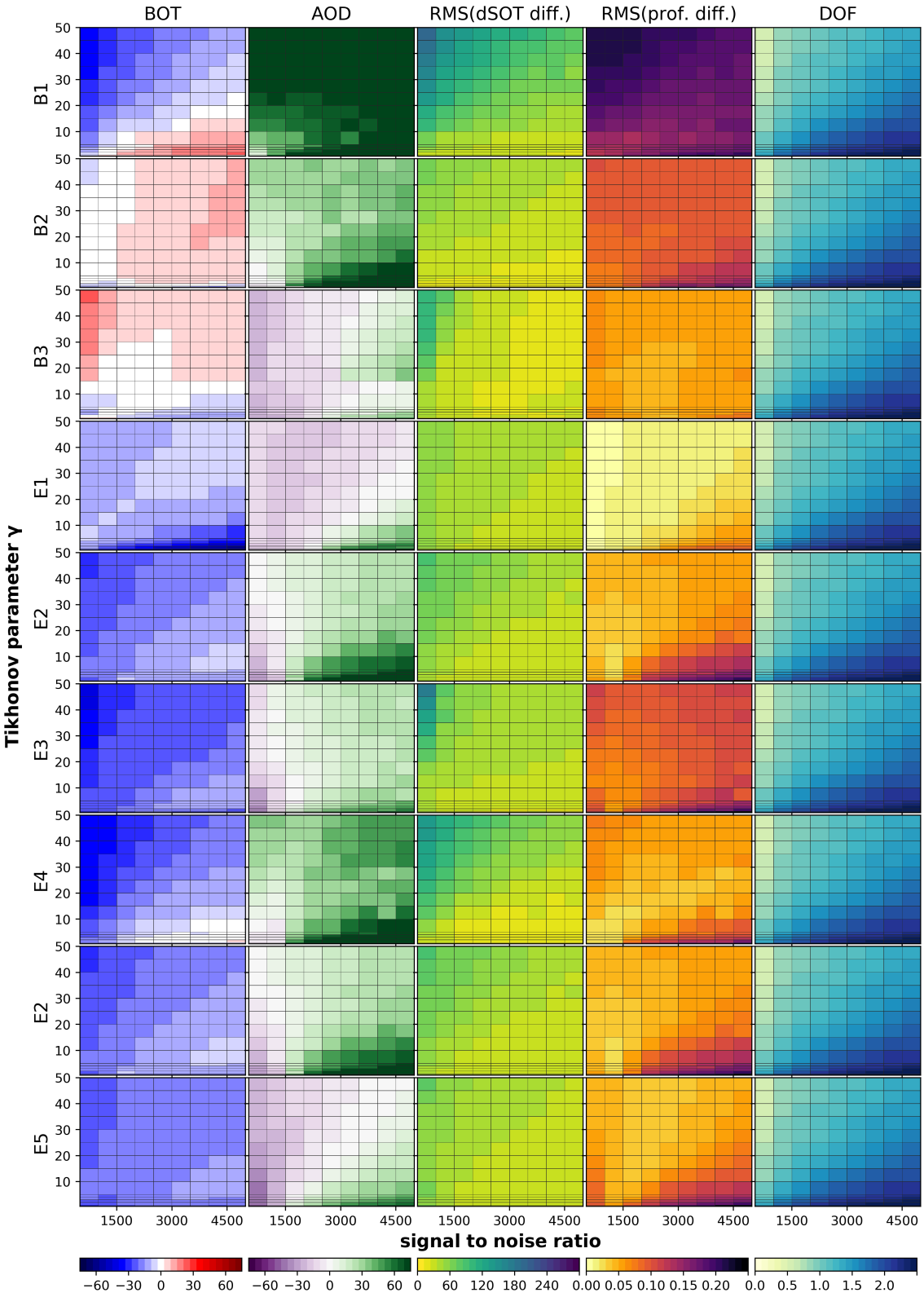


Fig. 4.14: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOD (in %).* ***3rd, 4th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: degrees of freedom.* ***Retrieval mode****: prev., const., fixed a priori profile (see Tab. 4.4)*

The Figures 4.15 and 4.16 show again the retrieval results of the *apri* and *prev*-mode, respectively, but with an additional a priori pre-scaling applied (introduced in **Test 4**). It can be seen that for both modes, the pre-scaling option improves all retrieval results for the scenarios E2 - E5. E1 shows also large bright areas which indicate a range of regularization parameters with good retrieval results but also more darker parts can be identified. The results for the box profiles for both modes, with and without pre-scaling, differ more strongly. When comparing results for the *apri*-mode it can be seen that B1 was improved while B2 shows varying results, depending on the regularization ratio and the specific quality criteria. With pre-scaling, larger positive deviations are found for the bottom extinction while the AOD loses the transition from purple to green, indicating that the profiles overestimate the near-surface extinctions with an in-turn average deterioration of the AOD. However, the RMS quantities are either smaller or similar, showing the difficulty of retrieving box profiles from MAX-DOAS data. The differences for B3 are smaller but the red to blue transition of the bottom extinction is lost by applying a priori pre-scaling. In contrast to that, the *prev*-mode shows more variable results with a slight deterioration for B2 in the BOT sub-splot but smaller deviations in the AOD sub-splot for B3 with pre-scaling.
Since the comparison of both modes with and without pre-scaling is hard with the naked eye, Figures AF.41 to AF.46 in the appendix show the ratios of the shown quantities for two different retrieval modes (e.g. $\text{RMS}_{acf}/\text{RMS}_{acp}$). Note that for BOT and AOD the ratio of the absolute values is plotted (e.g. $|\text{AOD}_{acf}|/|\text{AOD}_{acp}| = |\tau_{acf} - \tau_{true}|/|\tau_{acp} - \tau_{true}|$). These Figures enable an easier comparison as bright spots mean a similar performance for both modes while the different dark colours indicate that either the first retrieval performs worse (ratio smaller than 1) or the second retrieval shows larger deviations or RMS (ratio larger than 1).

For finding out which retrieval mode is the best, Table 4.5 and 4.6 summarize the comparison-plots in the appendix for the box profiles and the exponential profiles as follows: For all box or exponential profiles, the median values of the ratio for all SNR/γ combinations were calculated and averaged for all profiles. The median was chosen because the mean is affected by large values, when one retrieval performs much better than the other. The first row or column uses the mode abbreviations listed in Table 4.4 and the ratios are calculated by dividing the results for the specific mode in the first column by the matching mode listed in the first row.

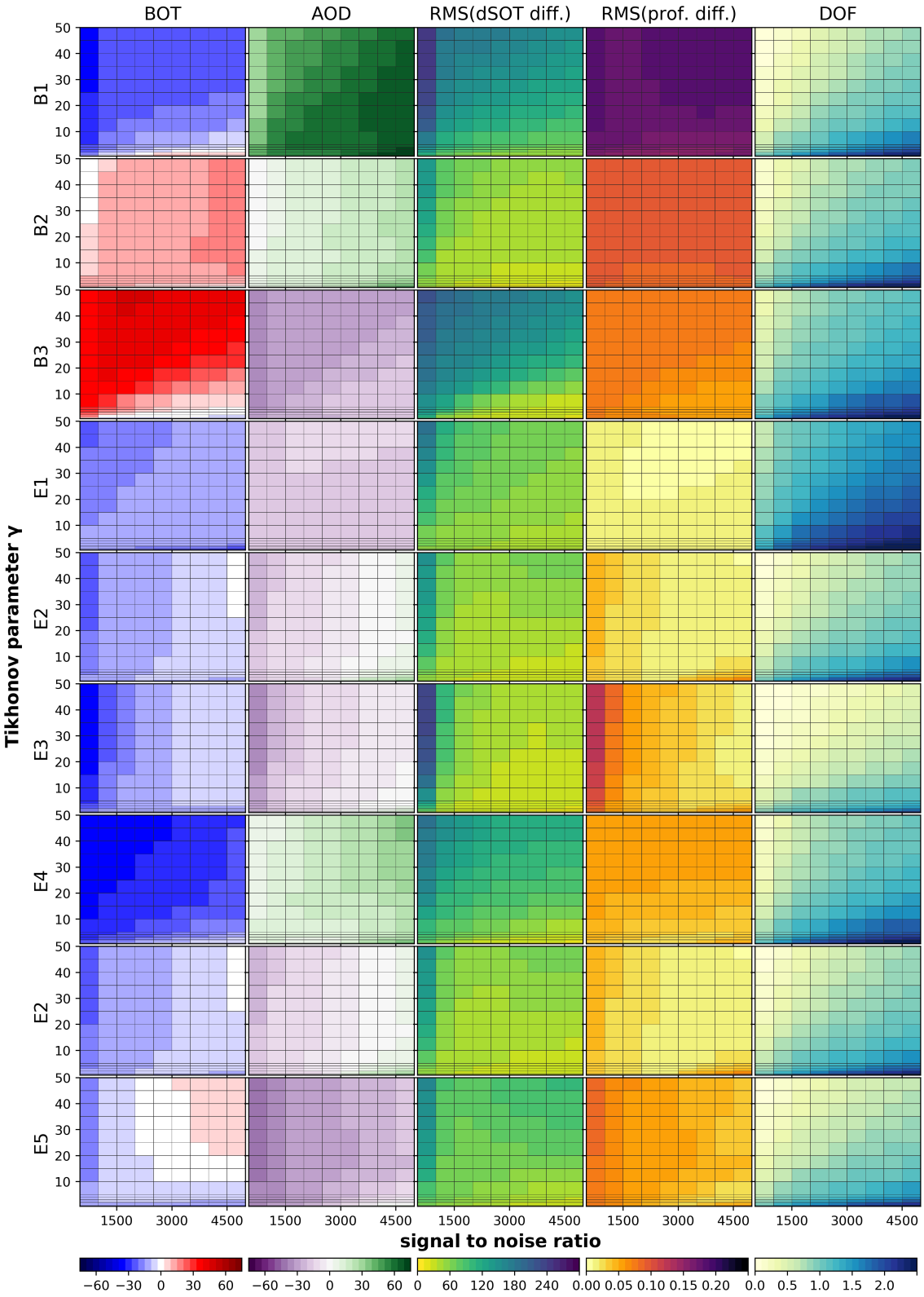


Fig. 4.15: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOD (in %).* ***3rd, 4th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: degrees of freedom.* ***Retrieval mode****: apri, const., pre scaling (see Tab. 4.4)*

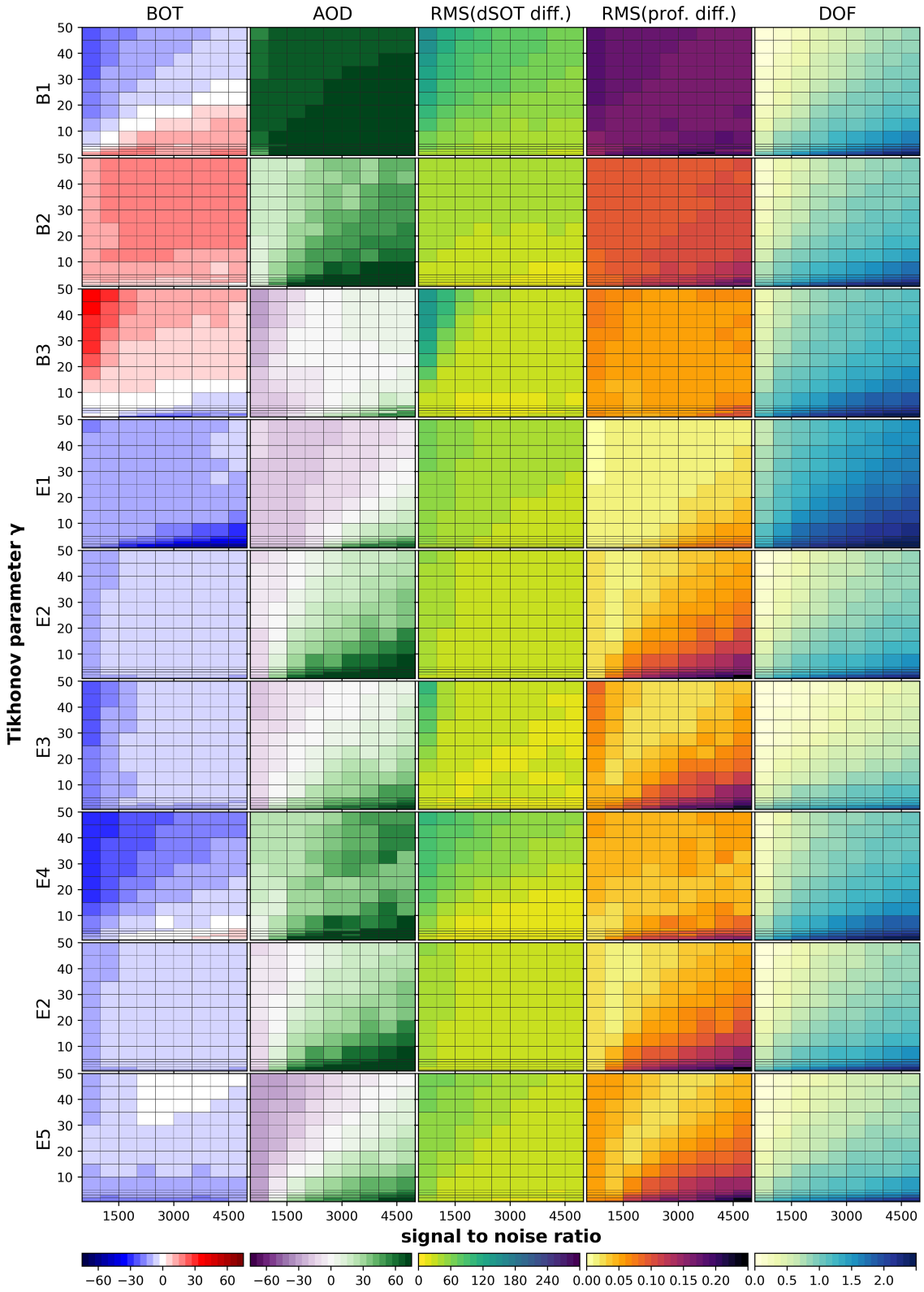


Fig. 4.16: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOD (in %).* ***3rd, 4th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: degrees of freedom.* ***Retrieval mode****: prev., const., pre scaling (see Tab. 4.4)*

Box profiles

	BOT						AOD					
	acn	acp	avp	pcn	pcp	pvp	acn	acp	avp	pcn	pcp	pvp
acn												
acp	0.96						0.84					
avp	0.95	1.02					0.87	1.06				
pcn	2.24	3.18	3.15				1.04	1.09	1.10			
pcp	2.13	2.38	2.36	0.83			1.37	1.51	1.59	0.99		
pvp	2.57	3.20	3.18	0.97	1.19		0.95	1.03	1.07	0.97	0.97	

	RMS(dSOT)						RMS(prof)					
	acn	acp	avp	pcn	pcp	pvp	acn	acp	avp	pcn	pcp	pvp
acn												
acp	1.43						1.04					
avp	1.44	0.99					1.05	1.01				
pcn	3.04	2.31	2.34				1.14	1.08	1.08			
pcp	2.85	2.24	2.28	0.96			1.10	1.05	1.05	0.97		
pvp	2.88	2.13	2.15	0.98	1.00		1.20	1.13	1.13	1.05	1.08	

Tab. 4.5: *Comparison of all modes evaluated as ratio of median values for all SNR/γ combinations and box profiles, for the specific quality parameter. Results of the modes in the first column are divided by the results of the modes in the first row. The colour-coding shows if the mode of the matching column or row led to better results.*

The results for the box profiles indicate a slight deterioration of the bottom extinction and a larger one for the AOD, for *apri* and *prev* when pre-scaling is used. A comparison of the *apri* and *prev* mode shows always a ratio larger than 1 indicating that the *prev*-mode leads to better results, with and without pre-scaling. When using an altitude depending variance instead of a constant one, smaller changes show a minor importance for the retrieval results. The large values of the RMS(dSOT) for the comparison of *apri* and *prev* indicate that the previous mode is more suitable for profile shapes which are far away from the a priori profile shape.

The comparison for exponential profiles indicate varying results depending on the specific quality parameter. While the AOD values reveals that pre-scaling improves the results when using the *apri*-mode, a slight deterioration can be found for the *prev*-mode. The bottom extinction is always improved by applying pre-scaling. The comparison of *apri* and *prev* shows that the bottom values are better for *prev* but the AOD is improved when using *apri*. However, the RMS of profile differences indicates a clear winner as the *apri*-mode with pre-scaling while the RMS of dSOT differences feature that the *prev*-mode with pre-scaling leads to the best profiles. These findings imply that pre-scaling leads always to better results when the true atmosphere is similar shaped as the a priori profile. However, the results for the box profiles with pre-scaling are only slightly worse than without so that a priori pre-scaling is considered as an optimal way for improving the a priori profile. The

Exp profiles

	BOT						AOD					
	acn	acp	avp	pcn	pcp	pvp	acn	acp	avp	pcn	pcp	pvp
acn												
acp	2.66						1.63					
avp	2.63	0.98					1.48	0.99				
pcn	1.84	0.85	0.86				0.99	0.76	0.78			
pcp	3.68	1.34	1.40	1.90			0.92	0.65	0.69	0.88		
pvp	3.07	1.19	1.22	1.60	0.87		1.41	0.90	0.97	1.26	1.34	

	RMS(dSOT)						RMS(prof)					
	acn	acp	avp	pcn	pcp	pvp	acn	acp	avp	pcn	pcp	pvp
acn												
acp	2.01						2.00					
avp	1.95	0.98					1.94	0.98				
pcn	2.99	1.41	1.43				1.03	0.53	0.54			
pcp	3.81	1.66	1.70	1.17			1.19	0.58	0.60	1.06		
pvp	3.35	1.45	1.48	1.05	0.89		1.72	0.78	0.80	1.42	1.21	

Tab. 4.6: *Comparison of all modes evaluated as ratio of median values for all SNR/γ combinations and exponential profiles, for the specific quality parameter. Results of the modes in the first column are divided by the results of the modes in the first row. The colour-coding shows if the mode of the matching column or row led to better results.*

retrieval mode comparison shows clearly better results when iterating from the previous profile, for the box profiles, and slightly better profiles for the exponential scenarios.

Unfortunately, the before-discussed comparison does not say anything about the best profile and its specific SNR/γ combination. However, it shows that smaller variations of SNR or γ are not problematic for varying aerosol profiles, on average. For one specific scenario, the impact of an optimal range of regularization parameters is strong when no pre-scaling is applied, independent from the retrieval mode. The impact was found to be smaller with a priori pre-scaling. The best profiles will be discussed in the next sub-section.

4.3.1.2 Retrieval of aerosol profiles with exponential and box shapes

The optimal SNR/γ pairs for the retrieval modes introduced above are hard to find and the results vary depending on the choice of quality parameter used for finding the optimal parameters. This is due to the fact that often either BOT, AOD or one of the RMS values shows a good agreement with the true value (or small values for the RMS). Unfortunately, all modes, with and without pre-scaling, have in common that the choice of the best parameters depends on the scenario. However, since a small deviation in the AOD column

does not indicate if this profile is reasonable or oscillating, the more reliable quantities are supposed to be BOT and the RMS of the profile differences. On the other hand, the AOD is a frequent validation quantity so that it should not be neglected.
Here, the optimal settings are found as follows:

For each retrieval mode and scenario, all SNR/γ results of the relative difference of retrieved to true quantity for BOT and AOD are summarized by calculating the RMS of both ($q = ((AOD^2 + BOT^2)/2)^{0.5}$). This RMS quantity q should be small when either AOD or BOT are close to the true value and it should be the smallest if both relative differences are small. It is then multiplied with the RMS of the profile differences, which also should be small, so that $q_{min} = q * RMS(prof)$ is expected to have a minimum for an optimal SNR/γ pair. Again, the retrieval is supposed to retrieve optimal profiles for a wide range of scenarios so that all SNR/γ pairs are averaged for all box and exponential profiles (except E1) for finding the overall optimal SNR/γ pair. E1 is not considered for the averaging as the results are good for nearly all modes and parameter settings.

The following Table 4.7 shows these optimal values for all retrieval modes and scenarios. In addition, the mean values for all scenarios (except E1) and the used values are noted in the last rows. Since these tests were done on discrete SNR/γ values, the *used*-row indicates the applied optimal values closest to the calculated mean.

	acf		**acp**		**avp**		**pcf**		**pcp**		**pvp**	
	SNR	γ	SNR	γ	SNR	γ	SNR	γ	SNR	γ	SNR	γ
B1	1000	0.5	500	0.5	500	0.5	500	1	500	1	500	1
B2	3500	3	500	45	500	40	500	3	500	0.5	500	0.5
B3	5000	10	5000	5	5000	5	3500	10	2500	5	4500	5
(E1)	500	50	4500	50	4500	50	500	50	500	50	500	50
E2	4500	3	4500	50	5000	50	1000	4	1500	50	3000	40
E3	5000	1	5000	25	5000	25	1500	3	3500	50	4500	50
E4	4500	2	2500	0.5	2500	0.5	1000	2	500	0.5	1500	3
E5	5000	3	5000	4	5000	2	2000	20	4500	50	5000	30
mean	4071	3.2	3286	18.5	3357	17.6	1429	6.1	1929	22.4	2786	18.5
used	4000	3	3500	20	3500	20	1500	5	2000	20	3000	20

Tab. 4.7: *Optimal SNR/γ pairs for each profile and as a mean value of all profiles, except E1 (in brackets). The used row indicates the closest values to the mean.*

From these numbers it can be seen that the exponential profiles can be well retrieved with more or less similar settings, depending on the retrieval mode. When pre-scaling is applied, the Tikhonov-parameter should be selected larger so that more smoothing is necessary. This is reasonable as scenarios far away from the a priori need less smoothing

constrains in order to be retrieved well. The box profiles show varying results, the parameters for B1 and B2 are close to each other while B3 seems to be closer to the exponential mean values. The scenarios B1 was found to be problematic most likely because the aerosol load is concentrated close to the surface, in contrast to the a priori profile.
As a conclusion, depending on the retrieval mode, the range of SNR and γ values can be wide for different scenarios indicating that not all scenarios can be retrieved well when using a common mean SNR/γ pair.

Figures 4.17 and 4.18 show retrieved profiles for the retrieval modes *apri* and *prev* with applied pre-scaling. Profiles retrieved without pre-scaling are shown in the appendix, in the Figures AF.47 and AF.48. In all plots, the dashed lines represent the profiles retrieved for the optimal SNR/γ pair of this specific scenario (individual rows in Tab. 4.7) while the pointed lines show the average optimal SNR/γ pair result (last row in upper Tab. 4.7).

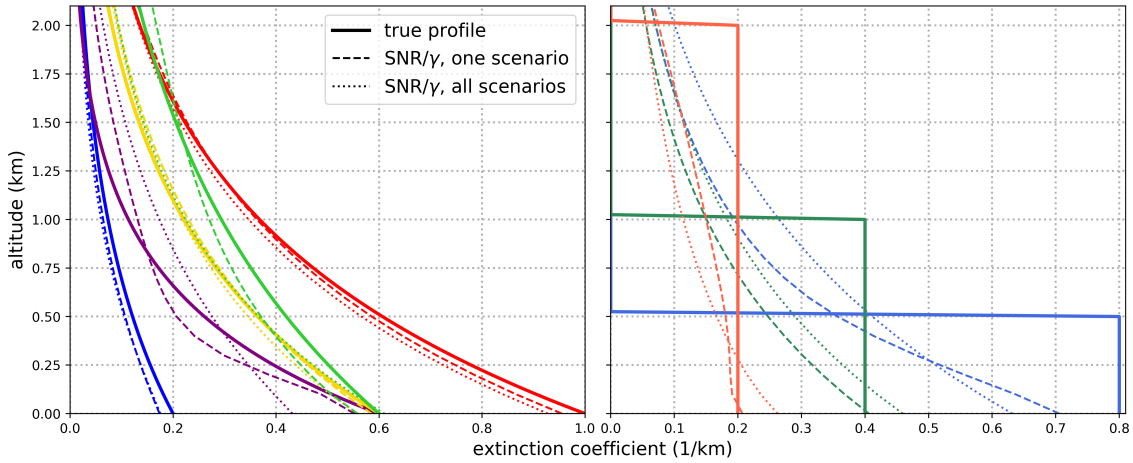


Fig. 4.17: *Profiles retrieved with the optimal regularization parameters for each individual scenario (dashed lines) and for the mean optimal SNR and γ values of all scenarios (pointed lines).* ***Retrieval mode****: apri, const., pre scaling (see Tab. 4.4)*

It can be seen that both retrieval modes are able to reproduce the exponential profiles similarly good for the individual optimal scenario settings and the mean values. However, the individual optimal settings are slightly closer to the true profile than the mean values. Larger differences between both modes can be seen for the scenarios E4 and E5 (purple and green), which represent the true profiles with a varying scale height, compared to the a priori profile. Here, the *apri*-mode shows larger deviations for surface-near altitudes for the overall mean. The *prev*-mode shows a better agreement for the bottom extinctions even though the profile scale height was not retrieved well.
The box profiles indicate that the mean values tend to over-smooth the solutions for both modes as the pointed lines seems to be exponentially decreasing. On the other hand, also the individual optimal settings are close to exponential functions for B1 and B2 (blue and

green) retrieved with the *apri*-mode. Only B3 (orange) depicts a more box like profile. The results of the *prev*-mode seem to be more adapted to the true profile for the individual scenarios for B2 and B3. The profile shape of B1 was never retrieved well. This might indicate that large near-surface concentrations are harder to retrieve even though the sensitivity is the highest in these altitudes. This problem can be attributed to two circumstances: 1. The a priori profile shape is far away and forces the retrieval to non-zero values for higher altitudes, where the measurement is not sensitive enough. This forcing has a large impact on the near-surface values as the correct AOT can only be created by compensating to high values at higher altitudes by lower values close to the ground. 2. The retrieval solution is always a smoothed response to the measurement due to the a priori smoothing via covariance matrix and Tikhonov-term. Because of this smoothing, edgy profiles are always hard to retrieve.

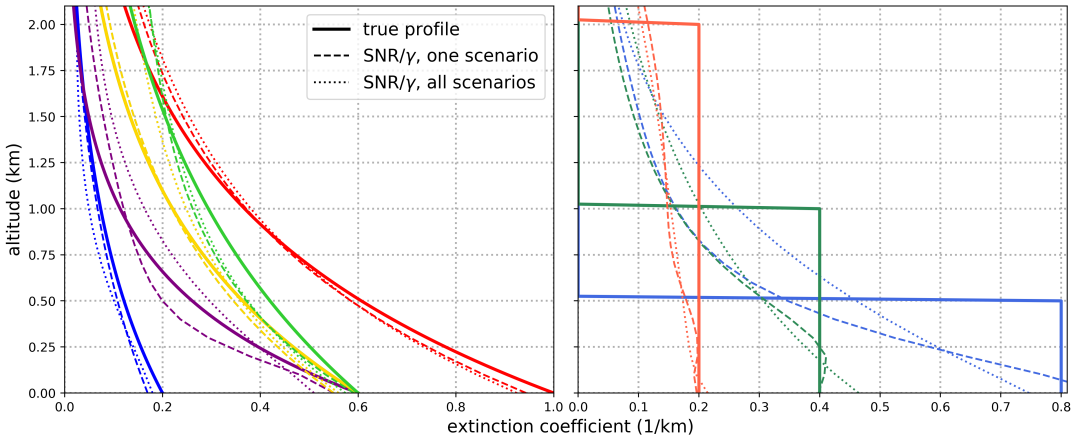


Fig. 4.18: *Profiles retrieved with the optimal regularization parameters for each individual scenario (dashed lines) and for the mean optimal SNR and γ values of all scenarios (pointed lines).* ***Retrieval mode****: prev, const., pre scaling (see Tab. 4.4)*

In contrast to the results with pre-scaling, the results for the exponential profiles are not retrieved well for both modes but the box profiles look surprisingly good. This is due to the stronger smoothing if the a priori is closer to the true profile when applying pre-scaling. Still, it can be concluded that the *prev*-mode with applied a priori pre-scaling led to the best results. The *apri*-mode performed satisfying but differing profile shapes might be problematic as it is not possible for the retrieval to reach extinction values which are far away from the a priori profile.

4.3.1.3 Retrieval of aerosol profiles with Gaussian shapes

The previous sub-section dealt with profile shapes close to the a priori profile (exponential profiles) and slightly farer away (box profiles). Here, the retrieval responds to the Gaussian-shaped profiles is presented (cf. Fig. 4.2).

The same analysis as in **Test 7** and the previous sub-section led to the mode comparison Table 4.9. The results of the variation of SNR and γ for the *apri* and *prev*-mode, with and without pre-scaling can be found in the appendix (Fig. AF.49 - AF.52, without mode-difference-plots). Since the bottom extinction has no meaning for these scenarios, it was replaced by the maximum altitude and maximum extinction (d_s was removed).
The table shows again that the *prev*-mode results in better profiles than the *apri*-mode. Furthermore, pre-scaling applied to the *prev*-mode leads to a slightly better adaptation to the maximum peak value while the maximum altitude is not found so well.

Gauss profiles

	Max. alt.						**Max. val.**					
	acn	acp	avp	pcn	pcp	pvp	acn	acp	avp	pcn	pcp	pvp
acn												
acp	1.22						0.88					
avp	0.94	0.80					0.87	0.99				
pcn	8.97	5.89	8.39				1.39	1.62	1.64			
pcp	6.30	4.09	6.48	0.85			1.39	1.62	1.64	1.04		
pvp	1.87	1.32	1.82	0.57	0.65		1.40	1.61	1.63	1.03	0.99	

	AOD						**RMS(prof)**					
	acn	acp	avp	pcn	pcp	pvp	acn	acp	avp	pcn	pcp	pvp
acn												
acp	1.03						1.02					
avp	0.92	0.84					0.98	0.97				
pcn	1.57	1.54	2.07				1.57	1.5	1.58			
pcp	3.40	2.44	2.91	1.18			1.49	1.43	1.50	0.96		
pvp	2.89	1.77	2.08	1.14	1.01		1.32	1.26	1.29	0.94	0.97	

Tab. 4.8: *Comparison of all modes evaluated as ratio of median values for all SNR/γ combinations and Gaussian profiles, for the specific quality parameter. Results of the modes in the first column are divided by the results of the modes in the first row. The colour-coding shows if the mode of the matching column or row led to better results.*

Interestingly, it seems vice versa for the *apri*-mode. Here, the maximum altitude was found well with pre-scaling while the maximum value is closer to the true extinction when using a fix a priori profile. The AOD is better with the *prev*-mode and with pre-scaling. The RMS of profile differences is slightly better without pre-scaling and with the *prev*-mode.
The optimal parameters were found in a similar way as in the previous sub-section. The AOD was replaced with both, the maximum peak height and maximum extinction value. The RMS of these essential parameters was again multiplied with the RMS of the profile differences (RMS(prof.)).
Table 4.9 shows the optimal parameters for the individual profiles as well as the mean value without GC00. This profile can be considered as close to an exponential profile re-

garding optimal retrieval parameters so that it was neglected here (cf. Fig. 4.2).
In contrast to the results of the exponential and box profiles, the differences when applying pre-scaling are small for the *apri*-mode and slightly larger for the *prev*-mode. The comparison of both modes reveal that iteration from the previous profile need a stronger Tikhonov parameter (similar to the box and exp discussion). In general, the spread of values around the mean SNR/γ pairs is not as large as found in the previous section but outliers are possible, indicating problematic scenario and parameter combinations.

	acf		**acp**		**avp**		**pcf**		**pcp**		**pvp**	
	SNR	γ	SNR	γ	SNR	γ	SNR	γ	SNR	γ	SNR	γ
GC05	5000	2	5000	2	5000	3	1500	4	1500	5	1500	4
GC10	2500	1	4500	0.5	2500	3	500	5	2000	1	4500	25
GC15	5000	1	4500	1	4500	4	2500	2	2500	2	4500	4
GC20	4000	2	5000	2	5000	3	3500	4	3500	3	4500	20
GC25	4500	2	5000	2	5000	2	4500	5	5000	4	5000	25
GC30	5000	2	5000	3	5000	2	5000	15	5000	25	5000	15
GC35	5000	2	4500	2	4500	1	5000	20	4500	15	4500	1
GC40	4500	2	5000	2	5000	1	4000	5	5000	5	5000	1
mean	4438	1.8	3481	1.8	4563	2.4	3313	7.5	3625	7.5	4313	11.9
used	4500	2	3500	2	4500	2	3500	10	4000	10	4500	10

Tab. 4.9: *Optimal SNR/γ pairs for each profile and as a mean value of all profiles, except GC00. The used row indicates the closest values to the mean.*

The optimal profile results of Table 4.9 for the *apri*-mode without and with pre-scaling can be found in Figures 4.19 and 4.20, respectively. The comparison of both plots show that pre-scaling leads to a decreased maximum extinction and maximum peak altitude. As already shown in **Test 6**, the vertical resolution decreases strongly with increasing altitude so that neither the proper peak height nor the true maximum extinction can be found for the higher altitude Gaussian profiles. Furthermore, noise-like, oscillating features are introduced for the altitudes below higher Gaussian profiles due to the lack of flexibility of the algorithm when applying a priori smoothing and the a priori profile shape is completely wrong. For the scenarios GC05 - GC20, the main profile features were retrieved satisfactorily for the *apri*-mode and both pre-scaling options but a fixed a priori profile leads to slightly more accurate peak heights. For the higher altitudes scenarios (GC25 - GC40) the proper peak maximum altitude and extinction were not retrieved well, especially with pre-scaling.

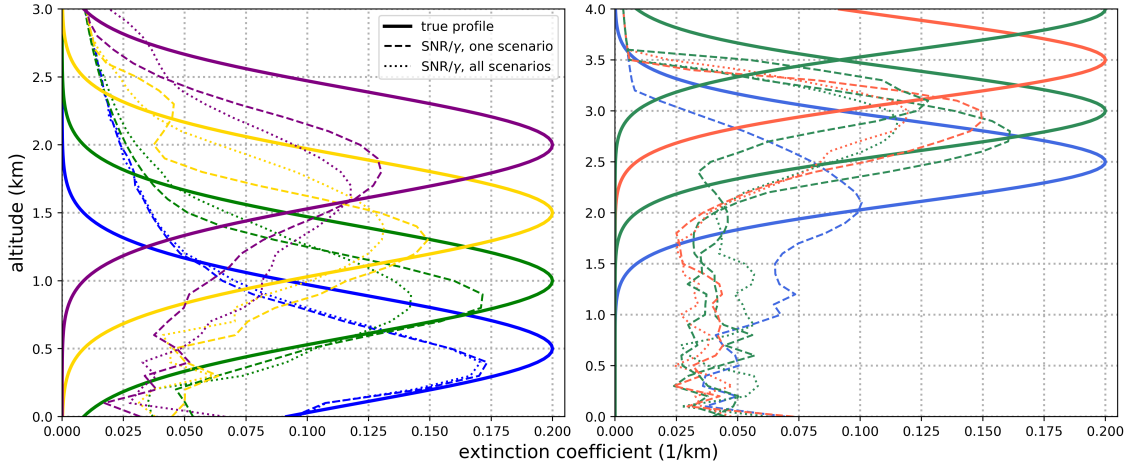


Fig. 4.19: *Profiles retrieved with the optimal regularization parameters for each individual scenario (dashed lines) and for the mean optimal SNR and γ values of all scenarios (pointed lines).* ***Retrieval mode****: apri, const., fixed a priori profile (see Tab. 4.4)*

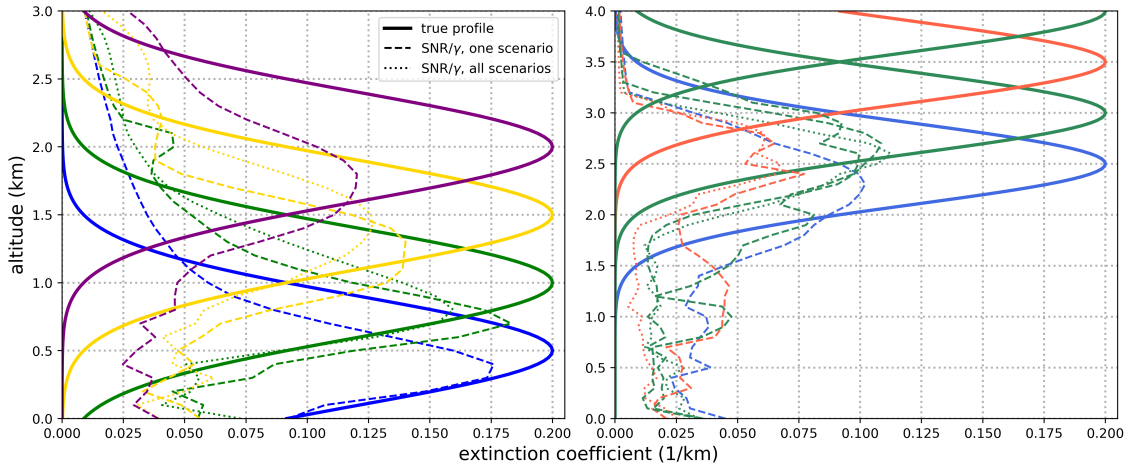


Fig. 4.20: *Profiles retrieved with the optimal regularization parameters for each individual scenario (dashed lines) and for the mean optimal SNR and γ values of all scenarios (pointed lines).* ***Retrieval mode****: apri, const., pre scaling (see Tab. 4.4)*

Figures 4.21 and 4.22 show the optimal results retrieved within the *prev*-mode, without and with pre-scaling, respectively. A comparison of both retrieval modes without pre-scaling (Fig. 4.19 and 4.21) shows that the previous mode leads to a higher accuracy regarding the true peak height for the higher altitude scenarios. Also the maximum peak extinction was found to be closer to the true values for GC25 - GC40. However, the lower altitude scenarios are not retrieved as good as for the *apri*-mode. On the other hand, the *prev*-mode decreases the impact of noise so that near-surface oscillations are strongly decreased. The comparison with pre-scaling shows that for the *prev*-mode, again, the higher altitude profiles are improved while the lower scenarios are retrieved with a slight decrease in maximum altitude and extinction.

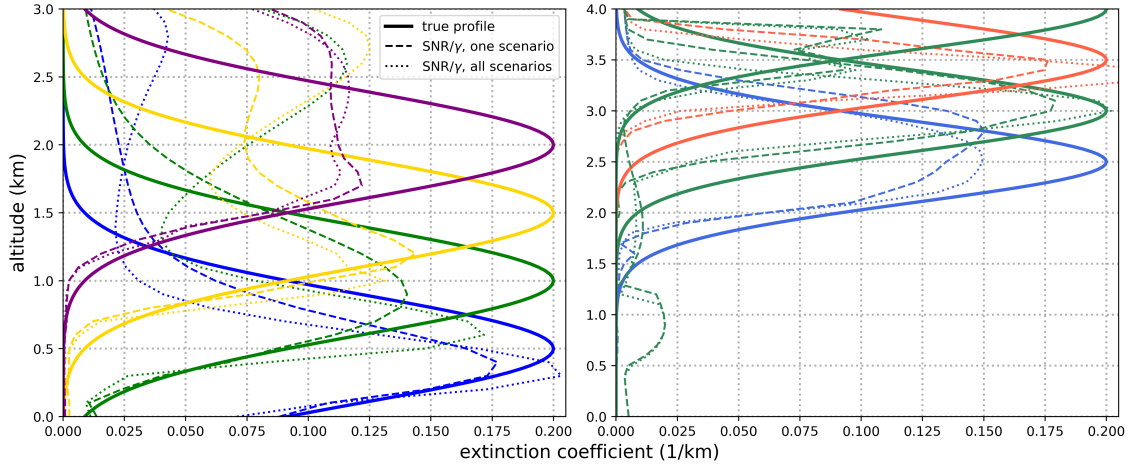


Fig. 4.21: *Profiles retrieved with the optimal regularization parameters for each individual scenario (dashed lines) and for the mean optimal SNR and γ values of all scenarios (pointed lines).* ***Retrieval mode****: prev, const., fixed a priori profile (see Tab. 4.4)*

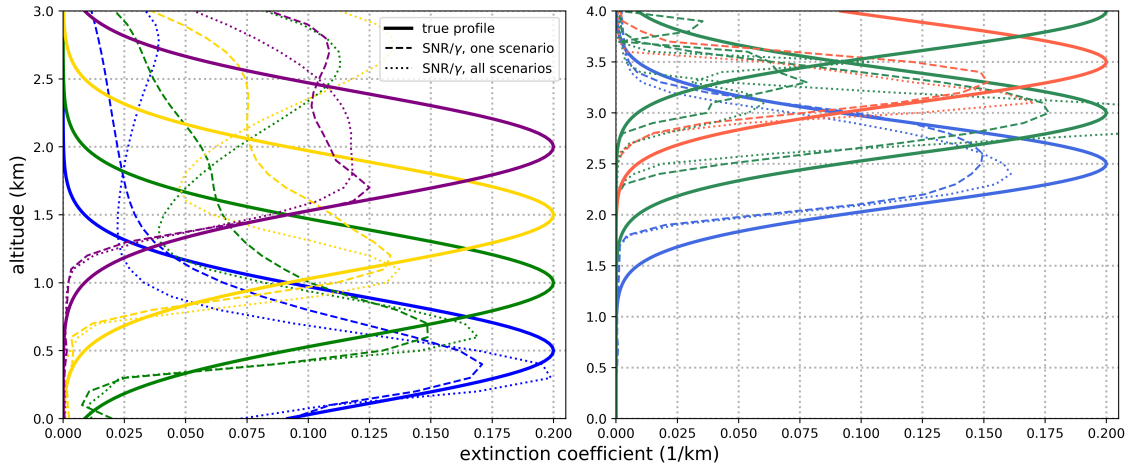


Fig. 4.22: *Profiles retrieved with the optimal regularization parameters for each individual scenario (dashed lines) and for the mean optimal SNR and γ values of all scenarios (pointed lines).* ***Retrieval mode****: prev, const., pre scaling (see Tab. 4.4)*

The discussion of the Gaussian profiles supported some findings from the analysis above but revealed also interesting new features. First, it is necessary to mention that the retrieval parameters for surface-near aerosol loads differ from those needed for elevated layers. Still, the retrieval with non-optimized parameters has a chance to be successful when the retrieval results indicate that an elevated layer was found so that the user might adapt the signal to noise or Tikhonov-parameter. It is possible to automatize these changes in the retrieval parameters but this has to be done due to the results of empirical studies as e.g. shown in this sub-section rather than analytical relationships. Similar to the box-profiles, a priori pre-scaling might deteriorate some profiles (e.g. GC05 - GC15) but leads to an improvement for others (e.g. GC25 - GC40), depending on the retrieval mode. The com-

parison of *prev* and *apri* results show again differing results, depending on the scenario. The near-surface profiles are retrieved slightly better with the *apri*-mode while the higher altitude scenarios are not good and oscillations for lower heights are possible. On the other hand, the *prev*-mode shows a small deterioration of near-surface profiles while the higher altitude scenarios are strongly improved.
Due to the versatility when retrieving different profile shapes, the *prev*-mode is considered to be slightly better and pre-scaling is applied on a regular basis within this thesis. For specific scenarios and case studies, a change of these parameters should be considered. Note that even though the sensitivity of the measurement and the algorithm is the highest for near-surface aerosol loads, successful retrievals of elevated aerosol layers are possible. However, the accuracy of the actual layer altitude as well as the maximum extinction is usually smaller for aerosol layers at higher altitudes.

One important fact should be noted. The analysis of synthetic scenarios does not mean an in-turn usability of these optimized parameters for real data. As the actual signal to noise of the measurement might change drastically throughout a day and in comparison to synthetic data, retrieval parameters must be optimized for real datasets as well. Furthermore, different fitting windows might also introduce the need for further optimization tests. However, the relative changes in SNR and γ will be similar giving the user a better feeling for the parameters itself and for the possibilities of the algorithm.

4.3.2 Sensitivity of the trace gas retrieval

The profiles for the NO_2 sensitivity study are chosen similar to the scenarios from the aerosol study (cf. Sec. 4.3.1). Figure 4.23 shows 5 different exponential profiles, 3 box profiles as well as Gaussian profiles with different shapes and in different altitudes. These scenarios were again used for the calculation of ΔS values from which NO_2 concentration profiles were retrieved. No noise was added to the simulations and the underlying atmosphere was assumed to be aerosol free in order to prevent falsifying of results due to possibly inaccurate aerosol retrievals.

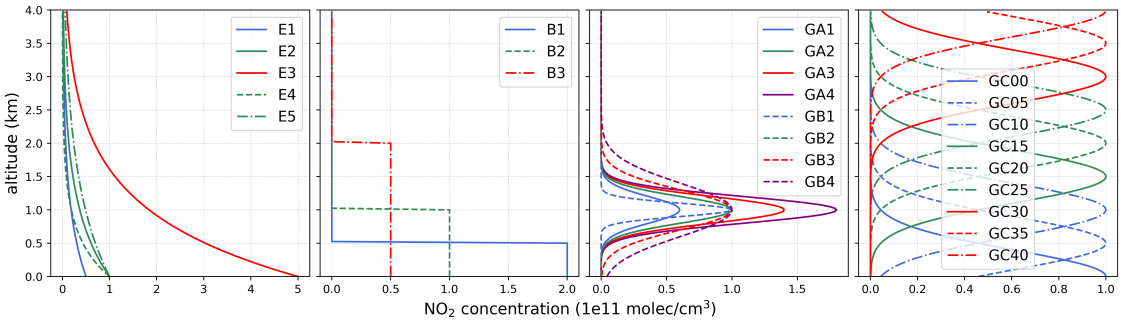


Fig. 4.23: *NO_2 concentration profiles used for the creation of synthetic ΔS datasets.*

4.3.2.1 Optimal trace gas retrieval settings

Before the optimal profiles of the scenarios above are presented, it should be noted that the aerosol and the trace gas retrieval have some dependencies in common:

- a priori trace gas concentration profile $\mathbf{x}_0$
- a priori covariance-matrix $\mathbf{S}_0$
- measurement covariance-matrix $\mathbf{S}_y$
- vertical grid steps and maximum retrieval height
- number of LOS for one profile and actual values of LOS angles
- atmospheric climatology profiles

Note that no further smoothing via Tikhonov regularization should be necessary since the mathematical problem is mostly linear (cf. Sec. 4.2). Furthermore, the three already introduced retrieval modes cannot be applied within an (non-iterative) optimal estimation based algorithm. However, the usage of the state vector in the linear or logarithmic space is often discussed in order to prevent oscillations (see e.g. FRIESS et al., 2006; CLÉMER et al., 2010). Additionally, since the problem is not ideally linear when the a priori is far away from the true state of the atmosphere, iterative approaches (e.g. by usage of the *Newton-Gauß* or *Levenberg-Marquardt* algorithm) are already introduced and in use (see the MMF algorithm by M. M. Friedrich (BIRA), shortly explained in the upcoming paper by Frieß et al. 2018).
In comparison to the aerosol synthetic study above, the results for **Test 1** (Maximum pieces of information and retrieval errors) and **Test 3** (Number and spacing of LOS angles) should be similar (but not necessarily equal) so that no further analysis is presented here. However, the averaging kernels and therefore, the degrees of freedom will be different and shortly discussed in this section. In addition, the sensitivity on the vertical grid (**Test 2**) is expected to be similar. On the other hand, the altitude depending sensitivity varies as the averaging kernels are usually different, in contrast to the aerosol retrieval. This point is discussed together with the vertical resolution in **Test 8**.
The first three bullet points in the above listing can be again summarized as regularization problem between a priori and measurement weighting. For MAX-DOAS profiling algorithms, the measurement covariance matrix is usually calculated by using the ΔS fitting error (Eq. 3.4) as variances within a diagonal shaped covariance matrix (e.g. FRIESS et al., 2006; CLÉMER et al., 2010; HENDRICK et al., 2014). Since this quantity is fixed, the definition of the a priori covariance matrix is a crucial factor for changing the regularization ratio (see Sec. 4.2.2). In Equation 4.13, the regularization factor g was introduced in order to change this weighting ratio. **Test 9** will be an introductory test on different ways of how to retrieve trace gas profiles on the example of one profile only but the later sub-sections 4.3.2.2 and 4.3.2.3 show results for the scenarios introduced in Figure 4.23.

Test 8: Vertical resolution

Figure 4.24 depicts the averaging kernels **A**, the area of **A** and the FWHM of the kernel peaks at their nominal and actual altitudes (height where the maximum value was found). Similar to the aerosol retrieval, the largest kernel values are found close to the surface with a strong decrease for higher altitudes.
The area is close to 1 for altitudes up to $2.5\,\mathrm{km}$ to $3\,\mathrm{km}$ with larger deviations close to the surface and for altitudes $\geq 3\,\mathrm{km}$. The smaller sensitivity for higher altitudes is similar to the aerosol retrieval but the decrease is not as strong as in Figure 4.12 due to the missing Tikhonov smoothing.
The FWHM supports that the highest resolution is close to the surface with a strong deterioration of resolution for higher altitudes. The FWHM at the actual altitudes is shifted to lower heights and shows distinctive peaks for $\leq 1.35\,\mathrm{km}$ but a strong decrease in FWHM due to the dominant negative part of the averaging kernels for nominal altitudes $\leq 2\,\mathrm{km}$ is prevalent as well. The larger near-surface kernel values for the high altitude averaging kernels (reddish colour) indicate a higher sensitivity which is not to be expected. This is rather a numerical smoothing response due to the a priori profile to equalize the negative values around $0.5\,\mathrm{km}$.

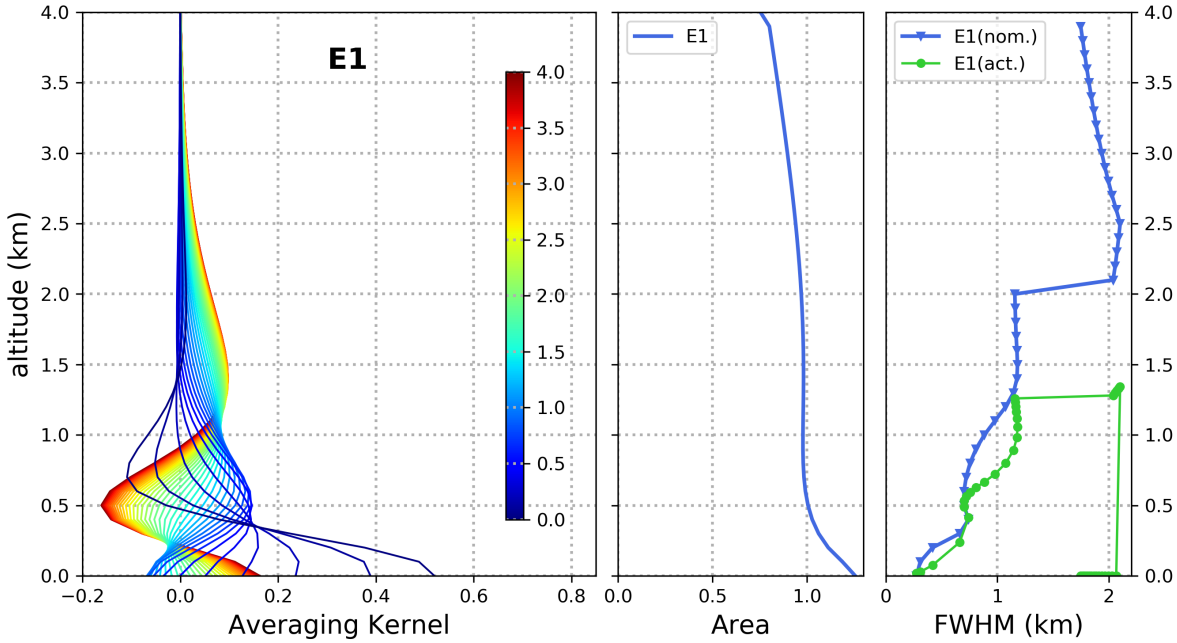


Fig. 4.24: *Averaging Kernels for scenario E1 (**left**): **Mid sub-plot**: Area of **A**. **Right sub-plot**: FWHM of **A** at the nominal altitude (blue line with v shaped markers) and at their actual altitudes (green line with dotted markers).*

Test 9: Regularization and different retrieval methods

The optimal estimation method (Eq. 4.13) requires an appropriate a priori profile and matching covariances in order to retrieve accurate trace gas profiles. Unfortunately, the true atmospheric conditions can be highly variable throughout a day so that these conditions are often not satisfied.

Figure 4.25 shows this case by applying an a priori profile (grey and dashed, exponential profile with a scale height of $1\,\mathrm{km}$ and a bottom concentration of $9.13 \times 10^{10}\,\mathrm{molec/cm^3}$) to the retrieval of E3 which is far away from the true NO_2 concentration. On the left-hand side of this figure, the blue dotted curve represents the standard regularization ($g = 1$). The bottom concentration was poorly retrieved and the overall profile shapes show smaller near-oscillating features. The change of g by several orders of magnitude results in different retrieval responses with individual and unique solutions. With g-factors smaller than 1, the bottom concentration is larger but it is still far away from the true value. An increase of g leads to elevated profiles for ranges of $1 \leq g \leq 1000$ which undergo a transition to the best bottom concentration for the largest g. However, the coinciding vertical column density is clearly underestimated for $g \geq 1000$.

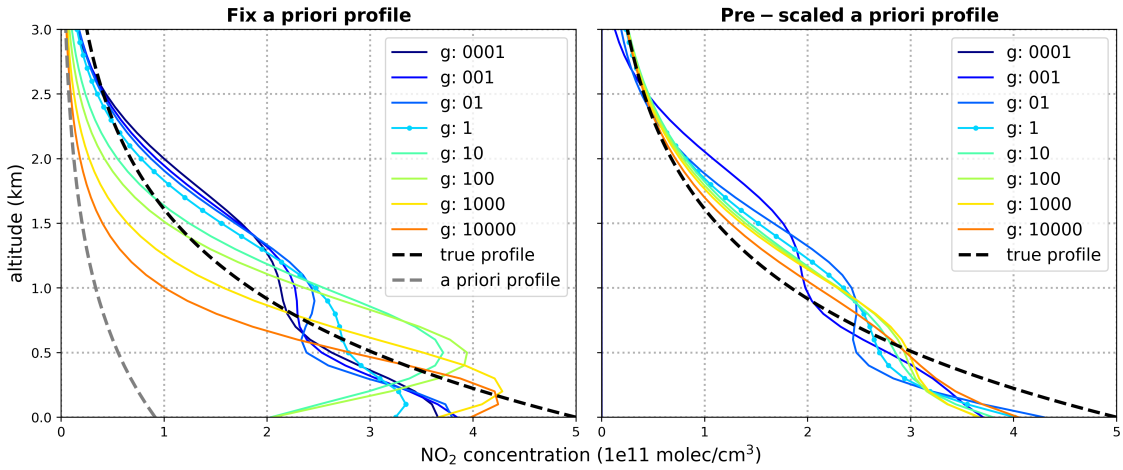


Fig. 4.25: *Profiling results for different g factors multiplied with the measurement covariance matrix without (**left**) and with pre-scaling (**right**).*

Similar problems were also found for the aerosol retrieval and were slightly improved by applying a pre-scaling option for the a priori profile (cf. **Test 4** in Sec. 4.3.1.1). For the trace gas retrieval, pre-scaling is much easier as the 30° LOS leads to a geometric differential air mass factor of 1 meaning that $\Delta S = V$ for this elevation angle. In other words, the a priori profile can be scaled with the 30° ΔS measurement in order to find a better estimate of the true atmospheric conditions. The retrieval results for profiles with an additional application of a priori pre-scaling is shown on the right-hand side of Figure 4.25.

Even though a large range of different g-factors is applied, the profiling solutions are quite similar. Note that the closest bottom concentration can be seen for $g = 0.1$ rather than 1 indicating again that the standard regularization ratio might be poorly chosen and should be considered as variable, depending on the true conditions and the a priori profile. Additionally, all profiles show a better agreement for higher altitudes which means that the vertical column density is closer to the true scenario as well. On the other hand, a perfect profile cannot be found with respect to the bottom concentration even though pre-scaling improved the profiles strongly.

Figure 4.25 leads also to another conclusion. Different solutions exist, depending on the g-factor, for which some might improve the vertical column density while others lead to a better bottom concentration. Since these solutions seem to be more or less stable for smaller changes in g, minima in the solution space can be expected. Figure 4.26 shows different quality quantities for the results of the E3-retrieval with a finer variation of g-factors. Here, a fix a priori profile was used. The similar pre-scaling figure can be found in the appendix (Fig. AF.53). The top plot shows again the retrieved profiles and in the lower four subplots, different quantities are depicted. In the 2nd row, the relative deviation of retrieved and true bottom concentration (blue) and vertical column density (green) is shown. It can be seen that VC is retrieved the best for $g \leq 1$. In this range, also the bottom concentration shows small deviations but the smallest values are found for $g \sim 6 \times 10^3$. Here, VC has already large deviations indicating the

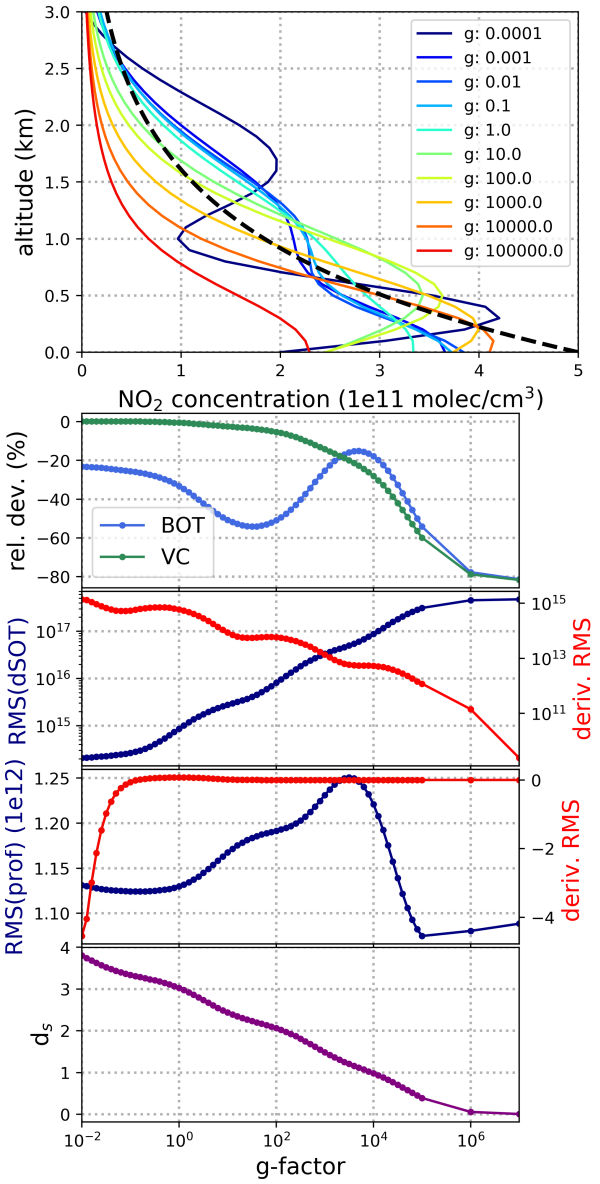


Fig. 4.26: *Different quantities retrieved with varying g-factors.* ***Top****: profiles retrieved with a fix a priori profile.* ***2nd row****: relative deviations of retrieved to true bottom concentration (blue) and vertical column (green), in %.* ***3rd row****: RMS of retrieved and true dSOT differences (blue) and its derivative (red).* ***4th row****: RMS of retrieved and true profile (blue) and its derivative (red).* ***Bottom row****: degrees of freedom.*

before mentioned trade-off between optimizing different quality quantities. The 3rd row shows the RMS of the retrieved and true dSCD difference (blue) as well as the derivative of this curve (red). The general curve behaviour shows small RMS values for the smallest g and a convergence to constant values for large g. The latter point is due to the strong a priori forcing when increasing the measurement covariance matrix strongly (by increasing g). The low RMS values for $g \leq 0.01$ show an unfortunate circumstance. Highly oscillating solutions tend to describe the measurement well. This is a problem as it complicates the finding of the best/true profile based on the analysis of the RMS alone. However, smaller fluctuations in the RMS curve already indicate that turning points and local minima exist which can be understood as stable sampling points in the measurement (or residual) space where slight changes of g do not affect the solution strongly. This is highlighted by the derivative of this curve which shows clear local minima, maxima and saddle points. One interesting feature is that $g = 1$ is not an exact maximum indicating again the insufficient a priori regularization when using the standard settings. The smallest relative difference of the bottom concentration is accompanied with either a local minimum or saddle point (not clearly distinguishable) in the RMS derivative curve. The 4th row shows the RMS of the profile differences with the largest value for the smallest relative deviation of BOT. In this RMS curve, a local minimum can be identified for $g \sim 0.13$ which is close to a local maximum in the derivative of RMS(dSOT). The degrees of freedom are the largest for the smallest g and are decreasing linearly (in this semi-logarithmic representation) for a large g range. Small oscillating features can be identified as well but the assigning of curve features to the other curves above is mostly not possible.
Nevertheless, local minima/maxima and saddle points in the derivative of the RMS indicate that an optimal regularization parameter g can be found by retrieving profiles with varying g (a test on this will be shown below). Unfortunately, these optimal g-factor will not improve all profile features.

Another way of retrieving more accurate profiles is the usage of an iterative scheme such as the *Levenberg-Marquardt* (LM) algorithm (see RODGERS, 2004):

$$\mathbf{x}_{i+1} = \mathbf{x}_i + (\mathbf{K}_i^T \mathbf{S}_y^{-1} \mathbf{K}_i + (1+\xi)\mathbf{S}_0^{-1})^{-1} \left[\mathbf{K}_i^T \mathbf{S}_y^{-1}(\mathbf{y} - \mathbf{K}_i x_i) - \mathbf{S}_0^{-1}(\mathbf{x}_i - \mathbf{x}_0)\right] \tag{4.18}$$

Here, $\mathbf{S_y}$ and $\mathbf{S}_0$ are again the covariance matrices of the measurement and the a priori profile, respectively. $\mathbf{K}_i$ represents the weighting function matrix (BAMF) around profile $\mathbf{x}_i$. $\mathbf{y}$ is the vector of measurements ΔS, of the trace gas of interest. The factor ξ is a damping factor which leads to the *Newton-Gauss-algorithm* for $\xi \longrightarrow 0$. The algorithm works as follows: First, $\mathbf{x}_{i+1}$ is calculated with $\mathbf{x}_i = \mathbf{x}_0$ with an initial factor $\xi > 0$ (factors suggested by FLETCHER, 1987). Then, if the residual χ^2 has increased, ξ will be increased as well in order to loosen the a priori constrains (with the assumption that a larger χ^2 was due to an inaccurate a priori profile). The profile $\mathbf{x}_i$ does not change. Increasing factors

have to be found empirically (here it is 8). On the other hand, if the residual has decreased, ξ is slightly decreased as well (here by a factor of 2) and the profile $\mathbf{x}_i$ is updated by the previous solution. In this way, the previous profile is only updated if the residual was minimized. The algorithm has the advantage that a wrong a priori/measurement weighting will be corrected for by changing ξ based on changes of the residual. In contrast to the *Newton-Gauß* algorithm, the factor ξ prevents the amplifying oscillations. However, wrong solutions are still possible when the convergence criteria are chosen to be too fine. Then, retrieval noise can introduce oscillations. On the other hand, if the criteria are too coarse, convergence is reached after too few iterations and the solution is often close to the optimal estimation result (which is not necessarily good).
As an additional option for both, the LM and the OE algorithm, calculations can be done in the logarithmic state rather than the linear one. This means that the logarithm of the state vector $\mathbf{x}$ and a priori vector $\mathbf{x}_0$ is used together with the matching converted weighting function matrix $\mathbf{K}_z$ and the converted a priori covariance matrix $\mathbf{S}_z$[7].

In Figure 4.27, the retrieval results for all introduced retrieval options are depicted for the scenario E3. The abbreviations are chosen as follow:

- **g1_ln**: g was set to 1 which means the usage of the standard OE. The linear representation was used.
- **g1_lg**: Same as above but the logarithmic representation was used.
- **gVar_ln**: g was found empirically. The linear representation was used.
- **LM_ln**: The LM algorithm was applied in the linear representation.
- **LM_lg**: Same as above but the logarithmic representation was used.

For all options in the listing, a fix a priori profile as well as a priori pre-scaling was applied. The gVar option was implemented for test purposes within the linear representation of the OE algorithm only. With one a priori profile, the OE creates profiling results for $0.01 \leq g \leq 10000$ with an equidistant spacing in the linear space ($10^{-2+N\cdot 0.2}$ with N $= 1, 2, ..., 30$). Then, the minimum in the residual space with the smallest g is searched for (mostly several minima exist). This restriction assumes that the regularization is not too bad so that a factor close to $g = 1$ can be expected. If no minimum can be found in the residual space, a minimum in the differential vertical column density between the previous and current g-result is taken ($|VC(g_i) - VC(g_{i+1})| \longrightarrow$ min). With this minimum g_{min}, a finer g-range calculation is applied in order to improve the results ($g_{min,-1}, ..., g_{min,+1}$ with 20 equidistant sampling points in-between). The coarse and fine g-searching-ranges are somehow arbitrarily chosen but smaller changes do not end in completely different solutions. This

[7] The conversion is done by applying the chain rule: $\mathbf{K}_z = \frac{\partial F}{\partial z} = \frac{\partial x}{\partial z}\frac{\partial F}{\partial x} \Rightarrow \mathbf{K}_z = \mathbf{MK}$, with $\mathbf{K}_z$, the weighting function in the logarithmic space. Then, $\mathbf{S}_z = \mathbf{M}_z^T \mathbf{S}_0 \mathbf{M}_z$.

two-step approach was applied to reduce the computation time. So far, it has never happened that neither the residual nor the VC difference minimum searching failed.

Figure 4.27 shows that the LM and the OE algorithms in the linear space end in more or less the same profiling results (blue and yellow dotted lines). On the other hand, both retrieval options find strongly overestimated profiles in the logarithmic space, when no a priori pre-scaling was applied (green and purple curves). The gVar profiling result found the best bottom concentration but beginning oscillation can be identified. The VC values of the specific profiles are given in brackets in the legend. Here, all retrieval options in the linear representation lead to relative differences smaller than $3\,\%$ to the true VC. The RMS of dSOT differences is the smallest for the gVar implementation with still reasonable values for g1 and LM. A comparison with the pre-scaled results (right-hand side) shows that all retrieval options found strongly improved profiles. The zoom-in in the lower left corner of the sub-plot depicts that the bottom concentration is the best for g1_lg, LM_ln and g1_ln. The gVar results shows the poorest surface value but the VC is closest to the true value. Also LM_lg finds a similar VC showing again that the different options seem to end in an optimized VC or bottom concentration. The smallest RMS(dSCD) was retrieved for g1_ln and LM_ln.

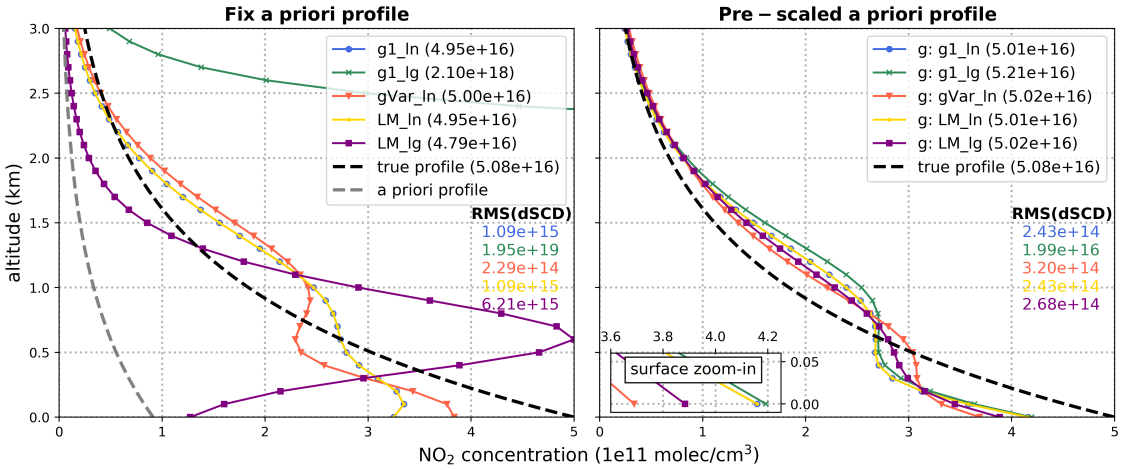


Fig. 4.27: *Profiling results for different retrieval options without (**left**) and with pre-scaling (**right**). The numbers in brackets give the vertical column density. The RMS of dSOT differences is given as text below the legend.*

As a short summary it can be stated that pre-scaling seems to be mandatory for this specific scenario. However, this conclusion might differ for other true profile shapes (see Sec. 4.3.2.2 and 4.3.2.3). The different retrieval options show similar profiles with only slight differences when pre-scaling is applied but larger deviations without. Especially the OE and LM options in the logarithmic space without pre-scaling ended in insufficient profiles which seems to be a matter of regularization. A clear winner of all options cannot be chosen so far. An analysis of more scenarios is necessary.

4.3.2.2 Retrieval of trace gas profiles with exponential and box shapes

Here, the profiling results for the scenarios B1 - B3 and E1 - E5 are shown, which were calculated with the retrieval options LM_lg, g1_ln, g1_lg and gVar_ln. The results for LM_ln are not depicted as they are similar to g1_ln. All profiles are again retrieved with and without a priori pre-scaling. The mean profiles of all geometries can be seen in Figure 4.28. With the aid of this plot, only a qualitative discussion is given. Quantitative findings will be explained with the help of Figure 4.29 below.

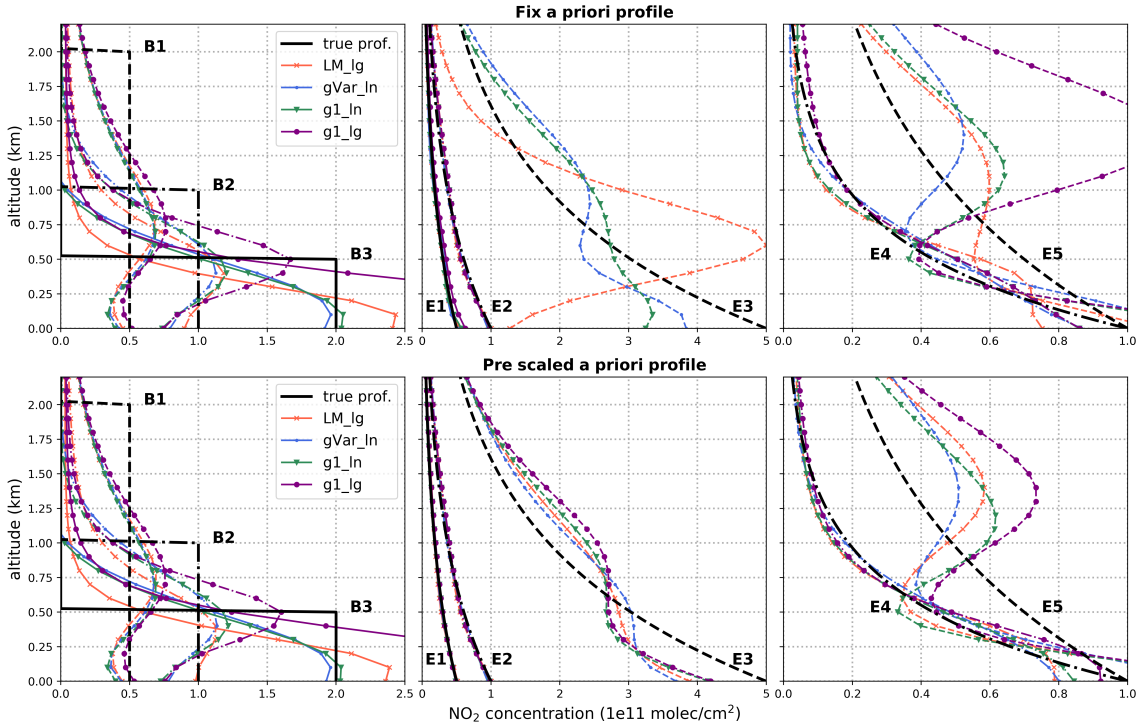


Fig. 4.28: *Profiling results for different modes with (**bottom row**) and without (**top row**) a priori pre scaling. Different colours and markers depict various modes. Different line styles indicate individual scenario results. The true profiles are shown in black.*

Starting with the box profiles, it can be seen that edgy boxes are not retrievable (similar to the aerosol retrieval, Sec. 4.3.1.2). The retrieval response is rather a smoothed but oscillating representation of the underlying scenario. Depending on the specific box scenario, g1_lg seems to overestimate the actual box concentration while the linear representation shows good profile shapes. The LM algorithm finds an overestimation for B3 only. The results of gVar_ln are close to g1_ln. Interestingly, the results with and without pre-scaling do not differ strongly so that the recommendation for a scaled a priori is still valid, even though the profile shapes of the true scenario and the a priori vary. The scenarios E1 and E2 are retrieved well for all retrieval options. However, the agreement seems to be slightly better with pre-scaling since the curves cover the true line more or less exactly for E1 and

quite accurate for E2. The result for E3 was already shown in Figure 4.28 and discussed in the previous sub-section. The profiles for E4 and E5 differ strongly in comparison to E2. This is interesting as the bottom concentration is the same, just the exponential scale height differs. While the general profile shape of E3 seems to be retrieved well, E5 shows large oscillating features. On the other hand, the bottom concentration depicts severe underestimations for all retrieval options and E4 whereas it is overestimated for E5. Again, the pre-scaling results seem to be better, especially for E5.

In Figure 4.29, different quality criteria are shown for all retrieval options with and without pre-scaling for the box and exp scenarios. For the BOT and VC rows, the relative deviation between retrieved and true value (BOT: bottom concentration, VC: vertical column density) are shown, in percent. The RMS rows are total quantities which were already introduced (see Sec. 4.3.1.1). In contrast to the aerosol discussion, the first RMS rows represent the RMS of the difference between true and simulated differential slant column density (dSCD) rather than dSOT. The error column shows again the RMS of the total error vector (cf. Eq. 4.17). The last rows depict the degrees of freedom. Note that the RMS rows and the error rows use a logarithmic y-axis. In each row, the black crosses indicate the mean values while the green lines show the median of all geometry results. The error bars depict the standard deviation. The results for E2 are shown twice in order to enable an easier comparison of the transition from E1 to E3 and E4 to E5.

In Figure 4.29, the box profiles show similar retrieval responses depending on the used algorithm and quality criterion. The bottom concentration was retrieved with the smallest relative difference for B1 by gVar and g1 with and without pre-scaling but only in the linear space. g1_lg shows the worst surface value. On the other hand, B2 is well retrieved with LM and equally bad with the other algorithms and retrieval options. B3 was retrieved the best regarding the bottom concentration with g1_lg_fix and LM.
The vertical column differences are always good except for both g1_lg retrievals. While the RMS of the profile differences is on a more or less similar level for all retrieval options, the RMS(dSCD) row shows larger deviations. Here, the smallest RMS was found for gVar and g1_ln for B1 and B2, respectively. However, the RMS values for the other options are still small, except for g1_lg. The RMS for B3 is on a similar level for the first 6 options while g1_lg is again large. The error row is interesting as it shows the largest total errors for gVar, where the smallest RMS(dSCD) was found (B1). The values are also high for B2 where the RMS(dSCD) was not large in particular but the standard deviation indicates strong deviations for the individual geometry results. The degrees of freedom are $d_s \geq 3$ with the highest values for gVar. Here, the standard deviation is again large as different g-factors for the individual geometries are possible. For the box profiles, it can be summarized that the logarithmic representation of g1 shows the worst retrieval result. The large

range of possible g-factors for gVar are accompanied with varying results. Sometimes, the optimal value seems to be found well (B1) which can not be supported for other scenarios (e.g. B2). The LM and g1 options show stable results with only medium RMS values but a good adaptation regarding profile shape criteria to the individual scenarios.

For the exponential profiles, different results can be expected as the pre-scaling option should improve the profile quality for E1 and E3 strongly. This can be confirmed by comparing BOT and VC for E1 - E3. The retrieval with a fix a priori are almost always worse than the pre-scaled options. For E3, the BOT row shows that LM and g_1 are even not in the plotting range without pre-scaling and the logarithmic representation. For all profiles, BOT is retrieved best with LM and gVar with pre-scaling (also without for gVar). The VC values are on a similar level except for g1_lg (fix a priori). The RMS(dSCD) row depicts the smallest value for gVar and LM with applied pre-scaling. The other retrieval options are at best on a similar level but mostly worse. The error row is more variable than for the box profiles and shows the smallest values for all options with pre-scaling (E1 and E3). The values for E3 are on a more or less similar level due to the fact that the a priori profile was already close to the true profile, here. The results for E4 and E5 show a contrary behaviour for BOT. While all options ended in an underestimation for E4, E5 is overestimated by the results of all retrieval options. VC is again always good, except for g1_lg. The RMS(dSCD) is the smallest for gVar with a fix a priori while the pre-scaling version is slightly larger (E4). Here, also the LM algorithm finds smaller RMS without pre-scaling. This is opposing to the results of E5 where LM_lg (fix a priori) is larger than LM_lg (pre-scaled). The results for gVar are independent of the a priori treatment. The smallest error values are again found by the pre-scaled retrievals.

As a conclusion, pre-scaling led mostly to the best results, except for E4. In total, LM and gVar show the best profiles even though there might be scenarios where other retrieval options are more reliable. When comparing the logarithmic space with the linear one, both options are good when the a priori is close to the true profile (E2 and in general with pre-scaling). If not, than the logarithm tends to end in overestimations (B1, B3, E5 in general and E1, E3 with a fix a priori). A comparison of the best two options (LM_lg_pre and gVar_ln_pre) leads to no clear winner. However, the RMS(dSCD) and the error rows depict that the gVar algorithm tends to deviate larger due to the dynamic selection of g-factors. Here, the box profiles were mostly retrieved with $g \sim 0.02$ but for specific geometries, values larger than 10 can be observed. This variability in the g-range was also found for the exponentially decreasing scenarios with g-ranges between 0.016 and 1585. Even though most profiles were calculated with $0.01 \leq g \leq 1$, the dynamic selection based on the derivative of the residual shows instabilities so that the LM in the logarithmic space is recommended together with the application of a priori pre-scaling.

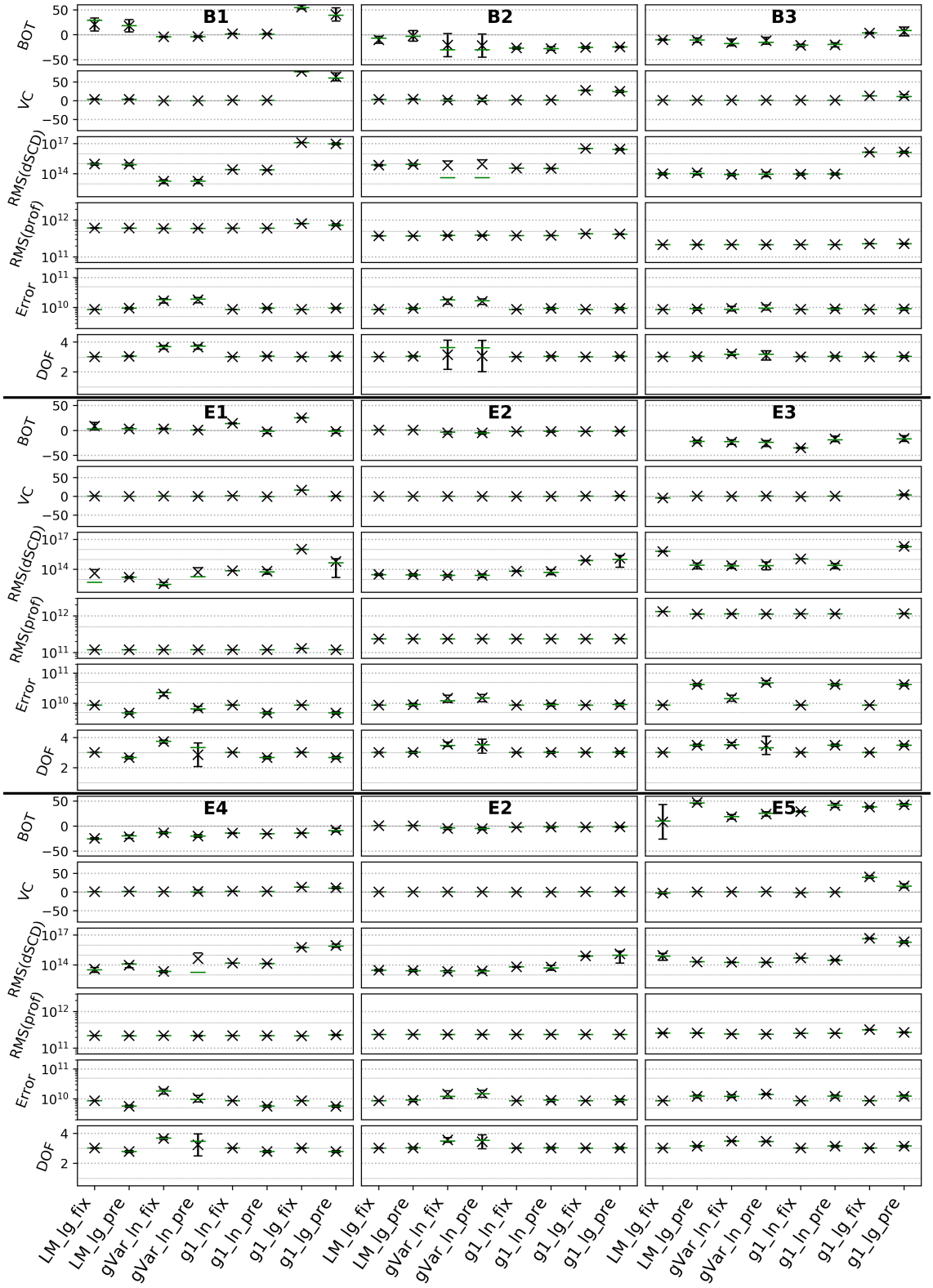


Fig. 4.29: *Results of different retrieval modes for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st and 2nd rows****: Relative difference of, retrieved to true bottom concentration (in %), retrieved to true vertical column density (in %).* ***3rd, 4th rows****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in molec/cm^3).* ***5th rows****: mean RMS of the total error.* ***6th rows****: degrees of freedom. See text for explanation of the x-axis abbreviations.*

4.3.2.3 Retrieval of trace gas profiles with Gaussian shapes

In contrast to the scenarios in the previous section, the Gaussian profiles show how the retrieval options deal with elevated layers. Since the sensitivity depends strongly on the altitude (cf. Fig. 4.24), poorly retrieved profiles can be expected for layer heights above 2 km.

Figure 4.30 shows the mean profiles for the retrieval options LM_lg, gVar and g1 (ln and lg) with and without pre-scaling. Only the profiles GC00 - GC25 are shown as the retrieval results for the higher layers are strongly underestimated in the peak altitude and maximum concentration (cf. Fig. 4.31).

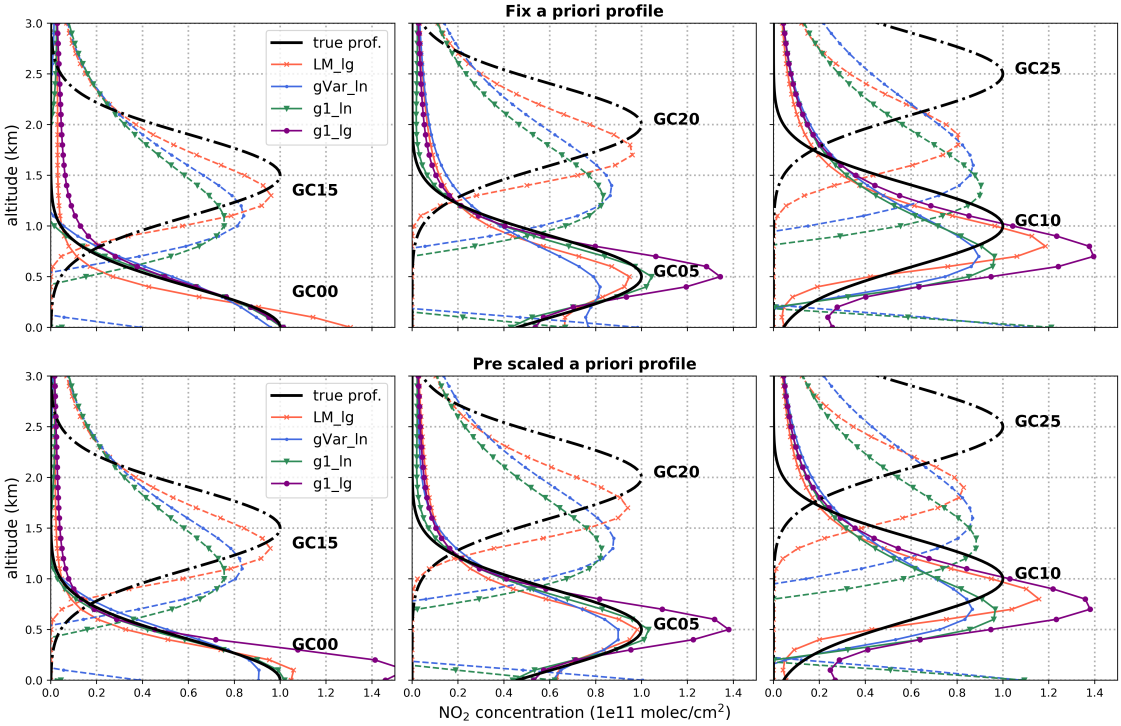


Fig. 4.30: *Profiling results for different modes with (**bottom row**) and without (**top row**) a priori pre scaling. Different colours and markers depict various modes. Different line styles indicate individual scenario results. The true profiles are shown in black.*

It can be seen that the peak altitudes of the maximum concentration seems to be well retrieved for GC05 and GC10 with a larger deviation for GC15. The peak altitude for GC20 was not found correctly for any of the retrieval options. Note that the results of g1_lg for GC15 or higher were suppressed as they suffer from larger overestimations. A qualitative comparison of the remaining options reveal that the best peak maximum altitudes where found for the LM algorithm followed by gVar. The standard OE option (g1_ln) shows the largest underestimations in comparison to the true maximum altitudes. It should be noted that the accuracy of the altitude and maximum concentration for the LM algorithm

depends strongly on the maximum number of iterations and the specific convergence criteria. Better adaptations are possible but the iteration numbers, and therefore, the retrieval time is strongly increased. The mean number of iterations was around 1000 for GC05 but larger than 20000 for GC20. Here, the computation time was decreased by permitting the update of the weighting function matrix $\mathbf{K}_i$. In general, block air mass factors do not depend strongly on the underlying profile so that this approximation is valid. However, when the a priori profile was far away from the true profile or the i-th profile $\mathbf{x}_i$, larger uncertainties are possible.

Figure 4.31 shows different quality criteria for the profiles GC00 - GC40. Compared to the similar figure in the previous sub-section, d_s and BOT have been replaced by the relative difference of the maximum altitude and the maximum concentration to the true values. Starting with the maximum altitude row, all Gaussian profiles higher than GC05 were retrieved best for the LM and g1_lg. The results for LM and GC05 are not as good as for g1_lg and g1_ln. The maximum concentration deviates stronger from the true values for higher Gaussian profiles. Besides GC00 and GC05, LM shows the best concentrations, especially for higher altitudes. gVar retrieved the maximum concentrations still satisfyingly while g1_ln deviates stronger for GC30 or higher. g1_lg is not in the range any more for the Gaussian profiles GC20 or higher. The vertical column density was well retrieved for all retrieval options except the standard OE in the logarithmic state (g1_lg). However, for the higher Gaussian's ($\geq$ GC25) the deviations for LM and g1_ln are getting larger while gVar stays on a constant good level. The RMS(dSCD) row depicts the smallest RMS for gVar and the largest for g1_lg. The values for LM and g1_ln are still small for GC05 - GC15 but getting larger for the higher profiles. The RMS of the profile differences is similarly good for all retrieval options up to GC15. For higher peaks, g1_lg starts to increase drastically. Also gVar and g1_ln are significantly larger than LM for GC25 or higher. The error row shows again that the pre-scaled options have the lower total error values. The highest errors are always found for gVar while they are on a similar level for the other retrieval options.

The conclusion in the previous section is still valid. The LM algorithm with a priori prescaling seems to be the best retrieval option even though there are scenarios for which other options lead to more accurate profiles.

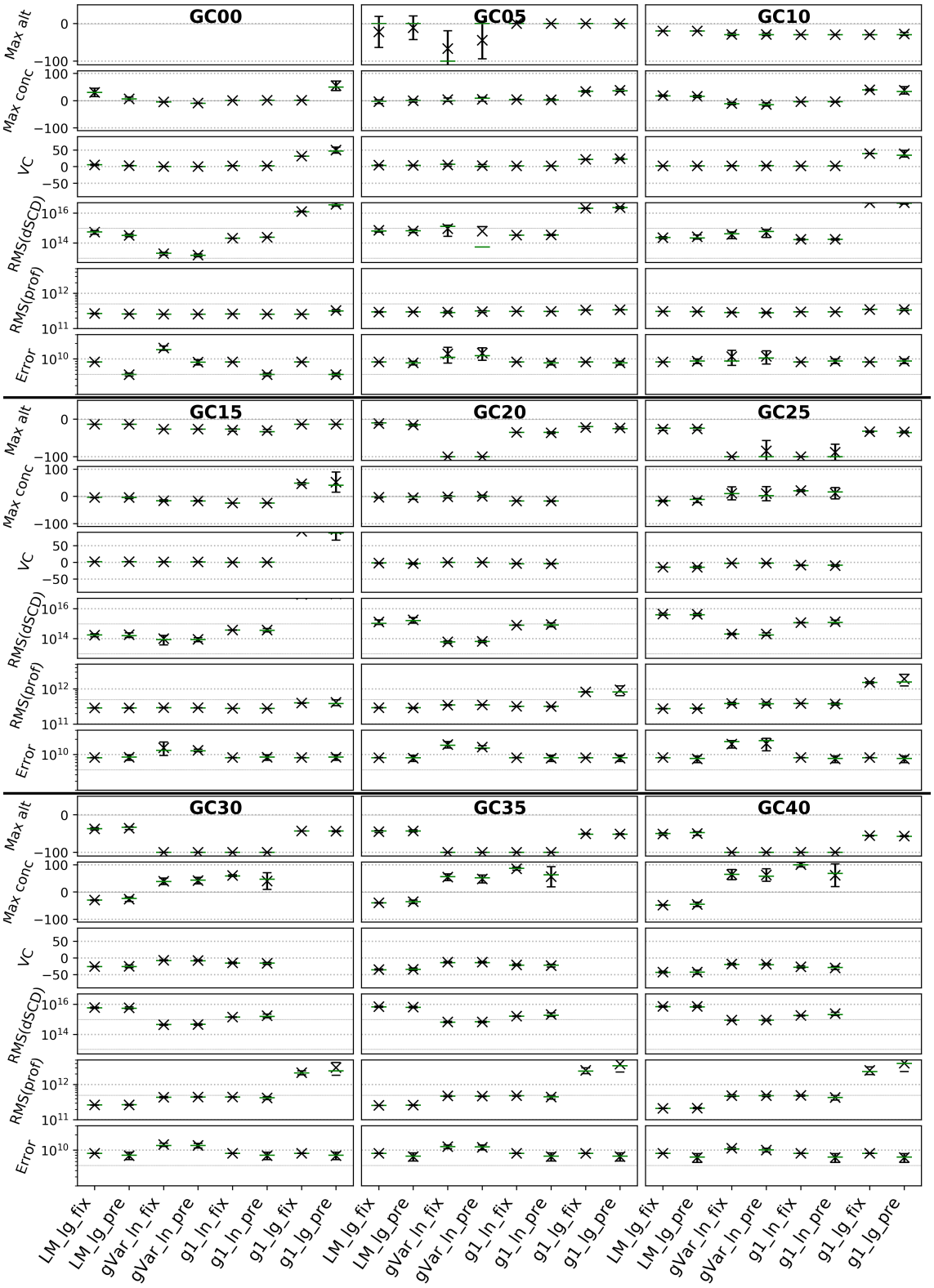


Fig. 4.31: *Results of different retrieval modes for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st and 2nd rows****: Relative difference of, retrieved to true bottom concentration (in %), retrieved to true vertical column density (in %).* ***3rd, 4th rows****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in molec/cm^3).* ***5th rows****: mean RMS of the total error.* ***6th rows****: degrees of freedom. See text for explanation of the x-axis abbreviations.*

4.4 Profile retrieval of real data

Within this section, profile retrieval results of real data measured in Bremen and during the CINDI-2 campaign in Cabauw (the Netherlands, in 2016) are presented. The CINDI-2 measurements are used to validate the profiling results with ancillary data from simultaneously measuring instruments covering similar or completely different air masses. The first sub-sections show the validation results of BOREAS profiles with these additional measurements. Note that the findings shown in the first sub-section 4.4.1 were already published in BÖSCH et al. [2018]. Subsequently to the CINDI-2 data introduction, additional analyses on relevant quantities are presented in order to extend on the published results.
The Bremen results are used, together with the discussion in Section III, to bring more light into the spatial and temporal distribution of trace gases in the area of Bremen.
For this purpose, a long time series of Bremen profiling results is presented and discussed together with results from Section 4.5.

4.4.1 Retrieval of profiles during the CINDI-2 campaign*

The 2nd Cabauw Intercomparison of Nitrogen Dioxide Measuring Instruments campaign (CINDI-2) took place in Cabauw (the Netherlands) from 25.08 - 07.10.2016 and was funded by the European Space Agency (ESA) and the research groups participating. The campaign goals were the characterization of differences between NO_2 measurements by various instruments and approaches, and the evaluation of the resulting datasets for the validation of the Copernicus satellite, Sentinel 5 Precursor (S5P). Over 40 different instruments operated by 30 groups from all over the world provided an outstanding ensemble of datasets for this task. The campaign was a successor of CINDI which was held in Cabauw from 16.06. - 24.06.2009 (see e.g. PITERS et al. [2012]; PINARDI et al. [2013]; ZIEGER et al. [2011]).
The measurement site Cabauw is located in a rural region dominated by agriculture but is surrounded by four of the largest cities of the Netherlands (Rotterdam, Amsterdam, Den Haag and Utrecht). Thus, depending on the wind direction, long range transport from highly industrialized areas is likely which results in high pollution events. Here, three of the five investigated days show a more or less steady wind from south-easterly directions with a change in wind direction in the evening of the 15th of September (see Fig. 4.32). On the 23rd, wind came from the west whereas one day later the wind came mostly from southerly directions. In this study, the instrumental set-up for the validation of aerosol and trace gas profiles consists of in situ and remote sensing instruments. Near-surface concentration/extinction values are provided by ceilometer, NO_2-LIDAR, long path DOAS

* This section has already been published in BÖSCH et al. [2018].

(LP DOAS) and in situ samplers (ICAD/CAPS, NAQMN). Integrated values are provided by Pandora and AERONET direct sun instruments. Besides the MAX-DOAS measurements, the only routinely retrieved profiling information comes from ceilometer and NO_2-LIDAR data. In addition, three NO_2-sondes were launched on one of the days investigated here (15.09.2016).

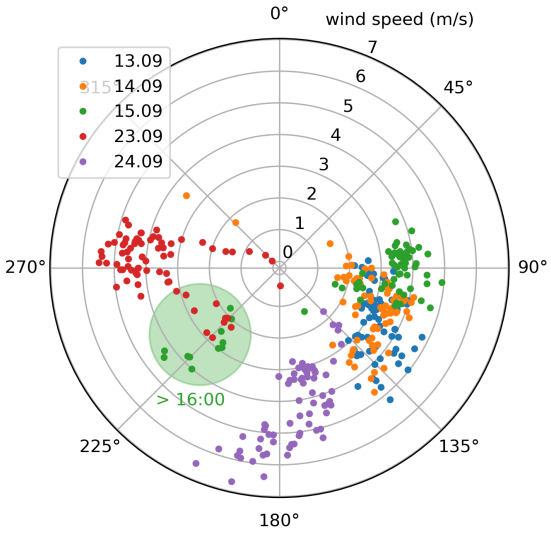


Fig. 4.32: *Wind speed and direction for the investigated days (CESAR [2018]) resampled to 5 minute values from 06:00 - 17:00 UTC.*

The profile validation is realized on two cloud free days (13.09. - 14.09.2016) and three days with broken clouds (15.09. and 23.09. - 24.09.2016). The operators of MAX-DOAS instruments were asked to perform elevation scans on the beginning of each hour and at 11:15 and 11:45 UTC every day. The ancillary measurements introduced in the following subsections were resampled or averaged on this time interval to prevent discrepancies due to time lags between two observations. Each scan took around 8 minutes and consisted of the following elevation angles: 1°, 2°, 3°, 4°, 5°, 6°, 8°, 15°, 30°. The telescope pointed into the western direction with a viewing azimuth angle of 107° from the south.

In this section, BOREAS retrievals were performed on data from the IUPB MAX-DOAS instrument on a 100 m step width vertical grid from the surface up to 4 km. The aerosol retrieval uses the a priori pre-scaling option with an exponential a priori profile described by a 1 km scale height and an aerosol optical thickness of 0.18. Asymmetry factor and single scattering albedo were chosen as 0.68 and 0.92, respectively. The SNR was set to 2500 with a Tikhonov-parameter of 2. The a priori variance was decreasing with altitude from 1.5 at the surface to 0.01 at 4 km. The trace gas retrieval uses again a priori pre-scaling with an exponentially decreasing profile with a scale height of 1 km and a vertical column density of $9 \times 10^{15}\,\mathrm{molecules/cm^2}$. The a priori covariance matrix uses the same variances as the aerosol retrieval but includes Gaussian distributed side diagonal elements to account for correlations between individual layers (BARRET et al. [2002]). The measurement covariance matrix for NO_2 consists of the total differential slant column error on the diagonal elements only.

4.4.1.1 Validation of aerosol retrievals

Instrumentation for the aerosol validation

Aerosol profiles were validated using three different instruments. The retrieved AOT was compared with values from an AERONET station (AErosol RObotic NETwork, DUBOVIK et al. [2000]; HOLBEN et al. [1998]). The Level 2 AOT at 440 nm was scaled with the Angström Exponent from the ratio 440 nm/675 nm to calculate the AOT at 477 nm (see Fig. 4.33). Due to measurement intervals varying between 4 and 30 minutes, we decided to resample the AERONET signal. The errors from AERONET instruments are usually in the range of 0.01 (VIS, IR) to 0.02 (UV) (SAYER et al. [2013]). Here, we used a constant error of 0.01.
The bottom extinction coefficient of the retrieved aerosol profiles was compared with in situ PM10 concentrations from the National Air Quality Monitoring Network (NAQMN) operated by the Dutch National Institute for Public Health and the Environment (RIVM). The NAQMN measurement site Wielsekade is located at a distance of around 900 m from the MAX-DOAS station. The site is listed as a regional background station and the PM10 measurement principle is beta attenuation (Thermo Fisher Scientific FH62I-R). The instrument provides data on one minute intervals but only the hourly values are validated by RIVM. These hourly values were resampled on the MAX-DOAS times and the minute values were used to calculate the errors as standard deviation for the data points shown (see Fig. 4.34).
Furthermore, we used AERONET scaled ceilometer near surface extinction as a further validation dataset. The ceilometer (CHM15k Nimbus) was operated by the Royal Netherlands Meteorological Institute (KNMI) and sampled backscattering signals every hour at a wavelength of 1064 nm. The integrated backscattering signal was divided by the 1020 nm AERONET AOD to get a conversion factor which was applied to the backscattering signal for the conversion into extinction coefficients. This new ceilometer profile was again scaled with AERONETs AOT at 477 nm. The error is expressed as the standard deviation calculated from the temporal and vertical averaging.

Comparison of aerosol retrieval parameters

Figure 4.33 shows the time series and the scatter plot of AERONET and BOREAS derived AOT. The BOREAS data was filtered for profiles which reached the iteration limit during the retrieval. Grey areas indicate clouds within the specified time period. In the scatter plot, BOREAS errorbars were calculated by integrating the total error (Eq. 4.17) vertically.

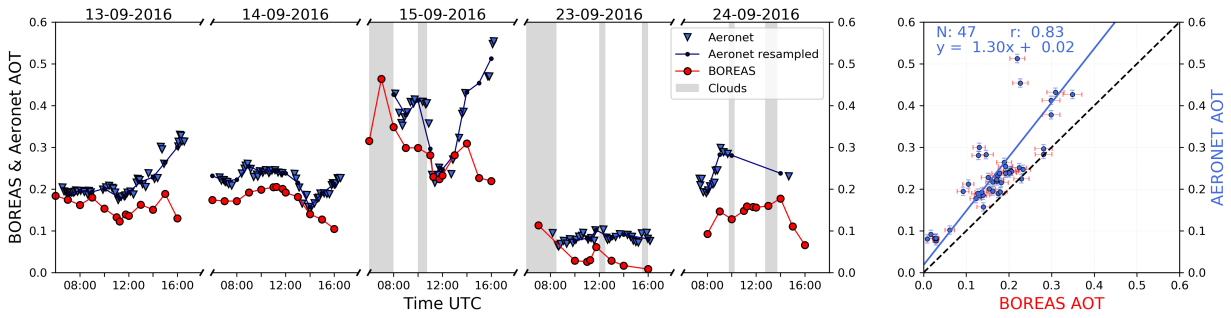


***Fig. 4.33: Left**: time series of AERONET (blue) and BOREAS AOT (red). Small triangles show the original AERONET measurement, small dots with connecting lines depict the resampled data. Grey areas indicate clouds. **Right**: scatter plot of both datasets including parameters of the orthogonal regression and Pearsons correlation coefficient.*

On the first two cloud free days, the temporal variability shows a similar pattern with a slight offset between both instruments most likely introduced by an underestimation of the AOT by BOREAS. Due to the limited sensitivity of MAX-DOAS profiling for higher altitudes (see Fig. 4.12), elevated aerosols do not contribute to BOREAS AOT but can be measured by direct sun instruments like AERONET sunphotometers.

In the evening, results of both instruments vary more. Different reasons for this finding are possible. First, a developing planetary boundary layer (PBL) might exceed the vertical extension where BOREAS has sufficient sensitivity. Second, when pointing towards the sun, saturation of the CCD becomes a problem leading to low integration times which decrease SNR and fitting quality. In addition, RTM calculations might introduce uncertainties at high SZA, when the aerosol load is high and the light path is strongly increased. Furthermore, the aerosol phase function used leads to uncertainties as the forward scattering peak might be underestimated by the Henyey-Greenstein parametrization. The other days show a higher variability due to an increase in cloudy scenes. Note that especially in the morning and around noon, clouds may influence the instruments in various ways because of their different azimuthal viewing directions. BOREAS profile retrieval cannot be considered as reliable when one or several elevation angles pointed into clouds during the measurements, as intensity and light path vary in unpredictable ways. This might also lead to large iteration numbers or to no convergence at all. On the 15th of September, the temporal patterns of the two instruments differ strongly. Although the extinction values around noon agree well, AERONETs decrease before and the increase afterwards cannot be found this strongly by BOREAS. Especially in the evening, both curves deviate, indicating highly variable atmospheric conditions which might also be introduced by a change in wind direction (see Fig. 4.32). The aerosol load on the last two days is much lower because of rainfalls in the time between 16.09 and 22.09. In the morning of the 23rd, several profiles were discarded due to exceeded iteration limits.

The correlation coefficient of 0.81 is high and shows the general good agreement at three

of five days. The regression line indicates the before mentioned small underestimation of the AOT by BOREAS.

Figure 4.34 shows the comparison of near-surface aerosol values. Extinction coefficients from BOREAS and ceilometer are depicted together with PM10 concentrations of the NAQMN in situ instrument. Note that the y-axis for PM10 concentrations was chosen to match the results of BOREAS on the 14th of September. The NAQMN in situ sampler and BOREAS show a good temporal agreement within the first three days. Only in the morning hours at the 13th and 15th, data from the two instruments show a larger spread. This deviation is most likely due to the insufficient vertical resolution of MAX-DOAS profile retrievals when the PBL has not yet evolved and near surface aerosol loads are dominant (see Fig. 4.12 and 4.35 for further details). On the 15th, early morning clouds might also have an impact on the deviation. Again, a larger variability for BOREAS results can be found for the last two days, when the retrieval was influenced by broken clouds. The in situ instruments show a smoother daily variation here. The ceilometer bottom extinction shows a good agreement with BOREAS and NAQMN on the first day and an underestimation on the second day. On the 15th of September, bottom values are influenced by thick clouds which might interfere with the backscattering signal of altitude ranges below the clouds (see Fig. 4.36). The 23rd shows more variability than the first two days but this can be found for the other instruments as well.
The correlation of BOREAS bottom extinctions with data from both validation instruments is high, indicating the good agreement on cloud free days. The correlation with ceilometer near-surface values is high as well but the regression line shows a general underestimation of ceilometer near surface values in comparison to BOREAS.

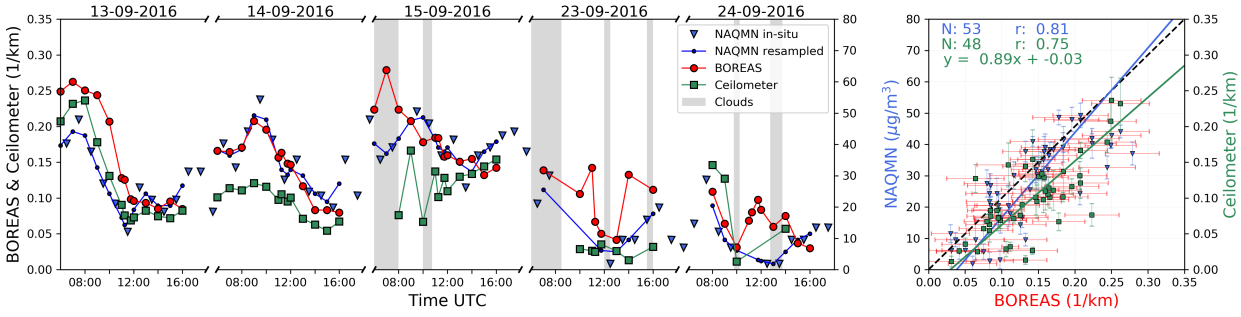


***Fig. 4.34: Left**: time series of BOREAS (red), ceilometer (green) and NAQMN in situ (blue) and near surface aerosol parameters. Small triangles show the original hourly NAQMN measurements, small dots with connecting lines depict the resampled data. Green squares show ceilometer near-surface extinction values evaluated by averaging the 10 - 50 m coefficients. Grey areas indicate clouds. **Right**: scatter plot including parameters of the orthogonal regression and Pearsons correlation coefficients.*

Figure 4.35 shows the aerosol layer height found by the ceilometer for all days. Especially during the first three days, an upward extending PBL can be identified from noon to the late afternoon. Only on the 15th, a more or less stable residual PBL seems to exist in the morning hours. The other days show a high variability in the morning with layer heights from 200 m to 600 m with more or less individual high signals which might be produced by smaller clouds at higher altitudes. The 24th of September shows a stable layer around 200 m and a diffuse developing PBL in the afternoon.

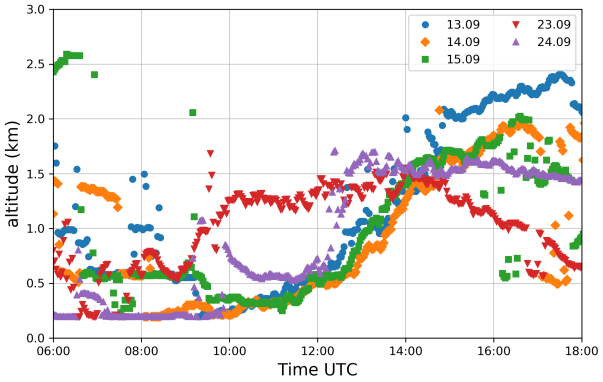


Fig. 4.35: *Ceilometer aerosol layer height within the planetary boundary layer.*

The underlying extinction coefficient profiles for the above discussed bottom extinction and AOT values can be seen in Figure 4.36. In the top row, temporal and vertically averaged ceilometer profiles are depicted. In the mid row, this data was smoothed to the MAX-DOAS vertical resolution by the application of BOREAS averaging kernels ($\mathbf{x}_{new} = \mathbf{x}_{apri} + \mathbf{A}(\mathbf{x}_{ceilo} - \mathbf{x}_{apri})$ see RODGERS and CONNOR, 2003). The first day shows a good agreement of BOREAS and ceilometer near surface extinction with a small offset between both curves. In the morning, both instruments find the aerosol load mainly located in the lowermost layers. Beginning around noon, the PBL starts developing with a maximum height of 2 km in the afternoon found by the ceilometer. BOREAS can not resolve this increasing PBL as well as the averaged ceilometer but the AK smoothed data indicate that the reason might be a limited sensitivity for the top boundary of the PBL.

On the 15th of September, the averaged ceilometer data show high and thick clouds in the morning and a rising PBL in the afternoon. In the AK smoothed data, these clouds can not be identified anymore. BOREAS introduces elevated aerosol layers which can be understood as a retrieval artefact due to these cloudy scenes. In the afternoon, BOREAS finds an upward expansion in the PBL similar to the ceilometer with the exception of the last profile which was influenced by the telescope pointing towards the sun. BOREAS extinction values are smaller in the PBL which is a consequence of the already explained underestimation of BOREAS AOT to AERONETs AOT which was used for the backscatter signal scaling.

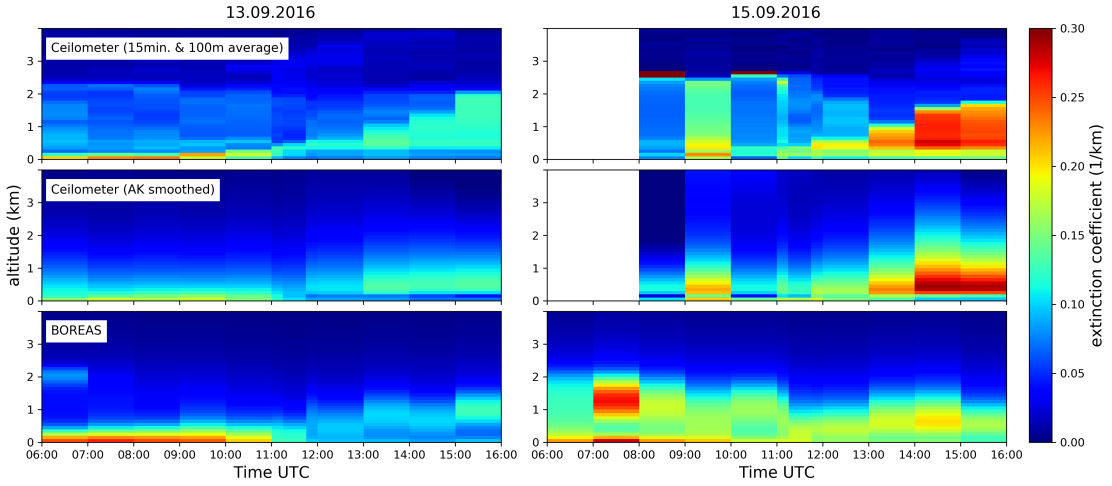


Fig. 4.36: *Comparison of aerosol extinction coefficient profiles from BOREAS and ceilometer for the 13.09.2016 (**left**) and 15.09.2016 (**right**).*

4.4.1.2 Validation of nitrogen dioxide profiles

Instrumentation for the NO_2 validation

Several instruments were used for the validation of BOREAS nitrogen dioxide profiling results. The VCD results were validated with the help of a Pandora instrument (#128) operated by LuftBlick (HERMAN et al. [2009]) and a NO_2 LIDAR from RIVM. In addition, three NO_2 sondes launched on the 15th of September by KNMI were also used.
The Pandora instrument was used with its direct sun capability to retrieve total nitrogen dioxide columns on hourly timesteps (see Fig. 4.37). Since the measurement time did not match to BOREAS scans, Pandora values were resampled. The stratospheric column was subtracted based on the approach of KNEPP et al. [2015] by using OMI stratospheric columns during the Cabauw overpass (L2 OVP NASA, 2018). The data was quality filtered and the errors were provided by LuftBlick and are based on HERMAN et al. [2009].
The NO_2 LIDAR (BRINKSMA et al. [2008]) measured not as frequently as Pandora and MAX-DOAS instruments but performed several measurements on the 15th of September. One LIDAR scan took approximately 30 minutes with the temporal midpoints of the scans on a non-regular temporal grid which necessitated resampling. The scan was done on different elevation angles which enabled the retrieval of vertical profiles with altitude points up to 2.5 km. For the VCD calculation, these profiles were vertically integrated. Errors were provided by the RIVM team and Gaussian error propagation was applied for the calculation of VCD errors. In addition to the remote sensing instrument, NO_2 sondes were launched on the 15th around 5:15, 8:04 and 10:25 UTC. The sondes provide NO_2 VMR as well as pressure, temperature and altitude data for the conversion to concentrations which were vertically integrated up to 4 km.

Near surface concentration validation was done with the help of in situ (NAQMN, ICAD, CAPS) and remote sensing instruments (LIDAR and LP DOAS). The in situ samplers NAQMN, CAPS and ICAD were operated by RIVM, the Royal Belgian Institute for Space Aeronomy (BIRA-IASB) and the institute for environmental physics in Heidelberg (IUPH), respectively. In addition to the aerosol bottom concentration, NAQMN provided also NO_2 concentrations (Teledyne API 200E, see Section 4.4.1.1 for details on data handling). IUPH and BIRA-IASB created a common NO_2 in situ dataset from IUPHs ICAD (Iterative CAvity DOAS, (PÖHLER et al. [2017])) and BIRAs CAPS (Environment SA, AS32M- CAPS, Cavity Attenuated Phase Shift, (KEBABIAN et al. [2008]). This joint dataset filled gaps in the individual measurement series and fixed a scaling issue for CAPS. BIRA-IASB operated two different CAPS instruments which were installed at the 27 m and 200 m level of the Cabauw measurement tower. The joint ICAD/CAPS dataset as well as the 200 m CAPS data were averaged over 15 minutes time steps. For both instruments (ICAD/CAPS) an error of 1 ppb was assumed.
In addition to the in situ instruments, a long path DOAS (LP DOAS) provided by IUPH (PÖHLER et al., 2010) was used for the validation. Four reflectors attached on different altitude levels of the measurement tower (12.7 m, 47 m, 107 m and 207 m) provided unique and well-defined light paths which enabled the calculation of concentration on several altitudes. The instruments elevation measurements lead to profiles every 30 minutes. Errors were calculated by the IUPH team.
Nitrogen dioxide profiles were provided only by the NO_2 LIDAR and sonde launches. The LP DOAS did not cover enough altitude levels for a comparison. The conversion from trace gas VMR to concentrations for all datasets was done with the meteorological data from the CESAR observatory (CESAR [2018]).

Comparison of NO_2 retrieval parameters

Figure 4.37 depicts the comparison of VCD for different instruments. The agreement with Pandora is good on all days with larger deviations in the morning and on the 15th of September in general with the exception around noon. Deviations in the morning might be (similar to the AERONET comparison) due to spatial differences in the NO_2 distribution, when Pandora points to the east and MAX-DOAS to the west. This argument is supported by the LIDAR measurement on the 15th that agreed well with BOREAS and shared the same azimuthal viewing direction. In the afternoon of that day, BOREAS VCDs increase faster than Pandora and LIDAR with a strong overestimation for the last profile at 16:00 when the sun was low. Here, we can exclude horizontal gradients as a reason for differences, because of the same viewing direction of LIDAR and MAX-DOAS. On the same day, NO_2 sondes proved the BOREAS VCDs only during the ascend (triangle with edge to the top) of the second launch correct. The matching descend (triangle with edge to

the bottom) for the first flight was approximately 38 km into the northwestern direction between Amsterdam and Rotterdam and agrees more with Pandora columns indicating differences due to azimuthal viewing directions. The strong differences in VCD for ascend and descend of the first and second sonde launches show a high temporal and spatial variability of the NO_2 concentration which is supported by all remote sensing instruments. The correlation with Pandora is slightly better than with LIDAR but the regression parameter shows, that Pandora VCDs are higher than those from BOREAS whereas the LIDAR regression is closer to one.

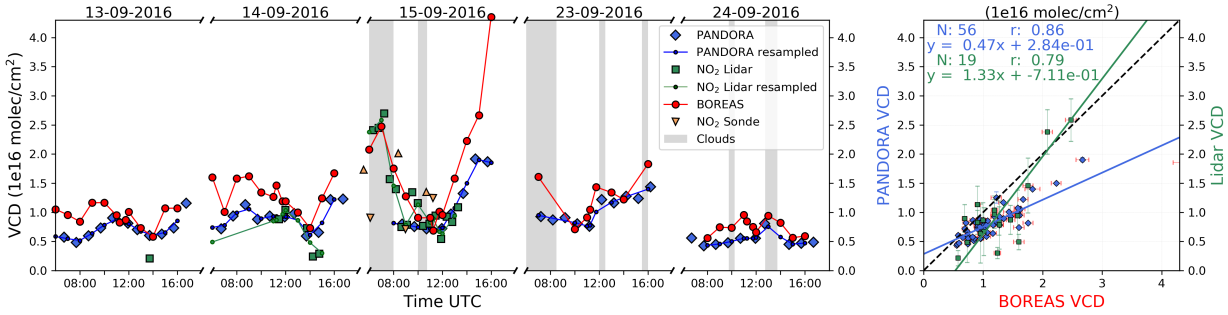


Fig. 4.37: ***Left****: time series of Pandora, LIDAR and BOREAS NO_2-VCD. Blue diamonds show the original Pandora measurements, green squares the LIDAR datapoints. Small dots with connecting lines depict the resampled datasets of Pandora and LIDAR. BOREAS is shown as red circles. Grey areas indicate clouds. The plot includes orange triangles as the representation of the integrated NO_2 sondes measurements with the ascend and descend separated in different triangles with the edge to the top or to the bottom, respectively.* ***Right****: scatter plot of the datasets including parameters of the orthogonal regression and Pearsons correlation coefficients.*

In Figure 4.38, near surface NO_2 concentrations are depicted. Both in situ datasets (NAQMN and ICAD/CAPS) agree well with each other and show larger differences only in the morning of the 14th of September. The LP DOAS profiles at 12.7 m and 47 m show similar concentrations as the in situ instruments. Differences in the morning of the first two days between the lower and the higher LP DOAS values indicate a strong vertical inhomogeneity which was also found for the aerosols (see Section 4.4.1.1).
BOREAS surface concentrations agree very well with all datasets except during the morning hours. On the 13th of September, in addition to the ICAD/CAPS data at the 27 m level (blue), the 200 m points are shown (green). The anti-correlated behaviour in the morning hours confirms the strong inhomogeneity and proves that nitrogen dioxide was mainly concentrated close to the surface at that time. The mean values of both CAPS instruments (yellow line) show a better agreement with BOREAS than the 27 m and 200 m data. It seems that BOREAS cannot fully resolve this thin near-surface layer and retrieves rather smooth profiles instead of sharp concentration peaks (see also the discussion about vertical sensitivity in Sec. 4.3.1.1). This is again in agreement with LP DOAS results on the 24th of September, where the 200 m level LP DOAS finds similar concentrations as

BOREAS. Note, that the NO_2 LIDAR has a finer resolution within the lowest 100 m which might be the reason, why it depicts similar concentration numbers as the other instruments in the morning hours of the 15th of September. As a rough estimation of the thickness of the individual thin layers in the morning, we divided BOREAS VCD by the ICAD/CAPS concentration and found, that the layer height lies around 200 m (06:00 - 08:00), 300 m (09:00) and 400 m (10:00) on the 13th of September. The mean AK FWHM for the surface layer on that day is 214 m indicating that the lowermost layer can be resolved only from around 09:00 UTC where the curves approach each other. A good agreement is found at 10:00 UTC when the concentration is focused in a layer twice as thick as BOREAS surface resolution.

The correlations are high for all instruments with the highest value for NAQMN and the NO_2 LIDAR. The in situ instruments show similar correlations but a slight underestimation of BOREAS can be found with slopes between 1.24 and 1.45.

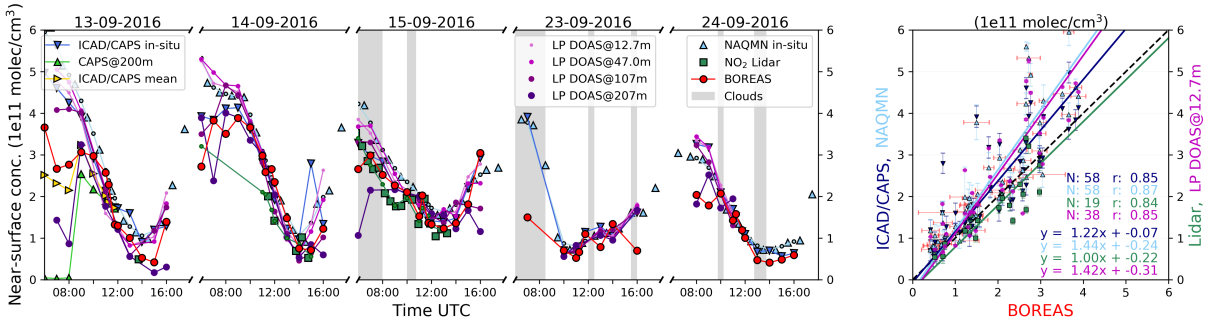


***Fig. 4.38: Left**: time series of in situ and remote sensing NO_2 near surface concentrations. Triangles show in situ instruments for NAQMN (light blue, triangles pointing upwards) and for ICAD/CAPS (dark blue, triangles pointing downwards). In addition, on the 13.09.2016, CAPS at the 200 m level is depicted as green triangles with the edge up and the mean values of both CAPS with the edge to the side (yellow). The NO_2 LIDAR is plotted as green squares and the LP DOAS as different sized circles in shades of magenta (lowest altitude as smallest circle). BOREAS is shown as large red circles. **Right**: scatter plot of the datasets including parameters of the orthogonal regression and Pearsons correlation coefficients.*

Figure 4.39 depicts NO_2 profiles for BOREAS, LIDAR and sondes on the 15th of September. Similar to the aerosol profile comparison, averaging kernels were applied to LIDAR and sondes measurements. The vertical profiles of the unsmoothed LIDAR show the previously discussed high vertical inhomogeneity. We can clearly identify a high concentration close to the surface and another elevated NO_2 layer with altitudes varying between 100 m and 500 m. The unsmoothed sondes agree well with the LIDAR measurements only at the ascend of the second launch. After the application of averaging kernels, the two distinct layers with high concentrations are smoothed to one layer with NO_2 concentrated at the surface indicating again a lower vertical sensitivity of MAX-DOAS profiling. In addition, smoothed and unsmoothed LIDAR find larger concentrations for altitudes over 500 m before 8:00 UTC which was not found by BOREAS and the sonde profiles.

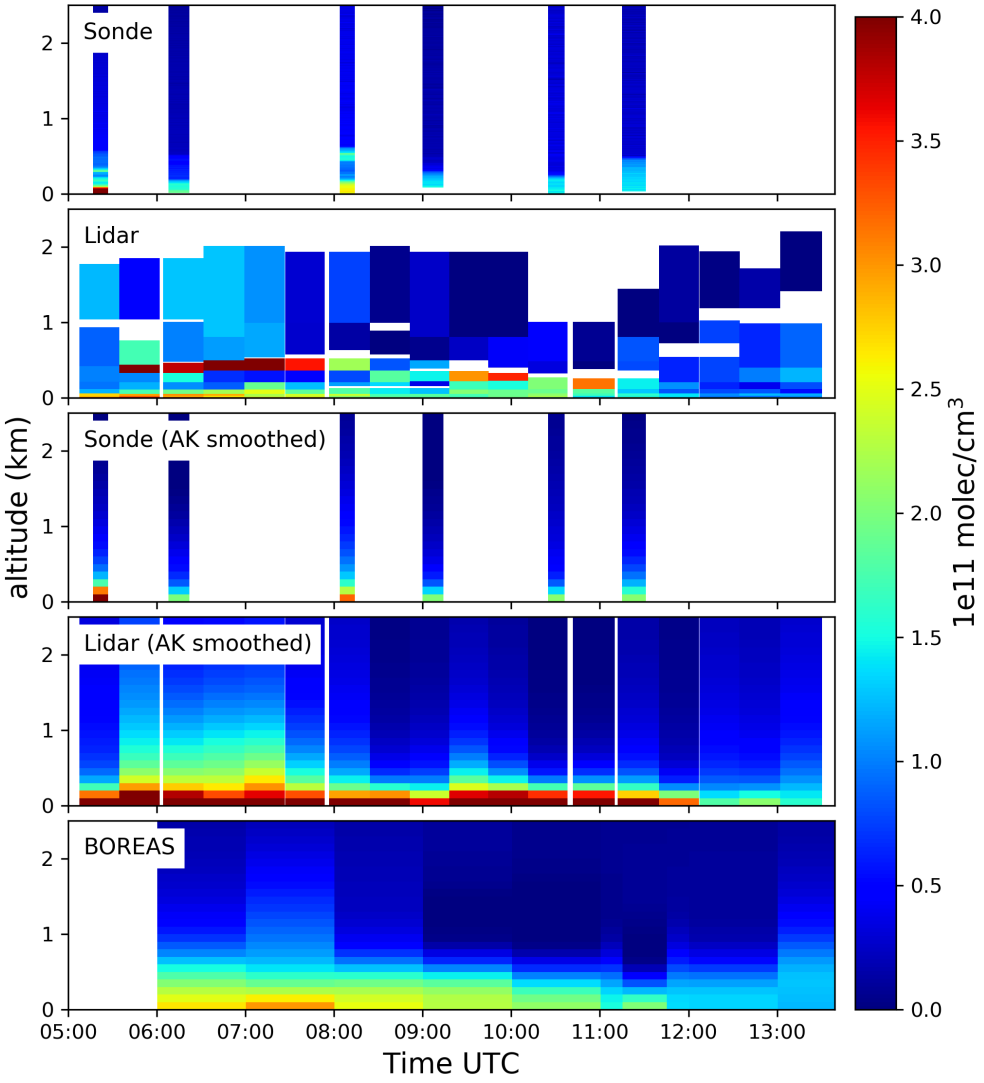


Fig. 4.39: *Comparison of NO_2 concentration profiles from BOREAS and LIDAR and sonde measurements for the 15.09.2016.*

4.4.2 Extended discussions of the CINDI-2 profiling results

The profiling results in the previous sub-sections were retrieved by applying the Henyey-Greenstein phase function due to the missing compatibility of early BOREAS versions with measured AERONET phase functions. Here, some further analyses are presented in order to show now the profiling responds to measured phase functions (**Test 1**), different climatological scenarios (**Test 2**), and further improvements of the BOREAS algorithm (**Test 3**) since the first paper submission (see BÖSCH et al., 2018). Furthermore, retrieval results for NO_2 retrieved in the UV spectral range are presented and discussed (**Test 4**).

Test 1: Aerosol profiles retrieved with measured AERONET information

To evaluate the impact of measured AERONET information on the retrieval results, AERONET inversion products were averaged for September 2016 in Cabauw. Table 4.10 shows the minimum, maximum and mean values for the asymmetry factor and the single scattering albedo calculated from the Level 2.0 Almucantar Inversion product. In the previous sections, $g = 0.68$ and $SSA = 0.92$ were used which are both closer to the minimum than to the monthly mean values. However, it should be noted that only 7 data points were available for SSA but 47 for g from this inversion product. Nevertheless, the used aerosol quantities show a large spread and further deviations due to the inaccurate treatment during the previous sub-sections can be expected. Note that the $440\,\mathrm{nm}$ AERONET data was used for the following test. An interpolation on $477\,\mathrm{nm}$ with the $675\,\mathrm{nm}$ AERONET data was not applied as for the latter wavelength less data points were provided.

Figure 4.40 shows the comparison of AERONET and Henyey-Greenstein phase functions. Mean, minimum and maximum asymmetry factors g and measured AERONET phase functions are shown as solid lines and data points, respectively. Thin solid lines depict the matching phase functions after the Legendre series expansion for the RTM calculation within SCIATRAN. The minimum and maximum AERONET phase functions were selected as min and max for the $0°$ scattering angle only in order to evaluate the results for a wider range of possible phase functions.

Asymmetry factor	
Mean	0.725
Min	0.659
Max	0.796
Single scattering albedo	
Mean	0.944
Min	0.902
Max	0.984

Tab. 4.10: *Mean, minimum and maximum value of AERONET asymmetry factor and single scattering albedo for September 2016 in Cabauw.*

It can be seen that both, forward and backward scattering is highly underestimated when using the Henyey-Greenstein approximation. Furthermore, P_{HG} with $g = 0.68$ is again closer to the minimum phase function but it highly underestimates forward scattering compared to $P_{HG,max}$ while the backward scattering geometry is overestimated here. The AERONET phase functions show more oscillating features for scattering angles $> 140°$ and a gap between $102°$ to $142°$. Nevertheless, the Legendre series expansion covers all angular features well showing that this gap can be considered as not problematic as long as no stronger oscillations lie within this gap. The comparison of the measured phase functions shows different angular behaviours. While the mean curve lies between the minimum and maximum phase function for forward and backward scattering geometries, it shows a larger probability for scattering perpendicular to the incident photon direction as the min and max curves.

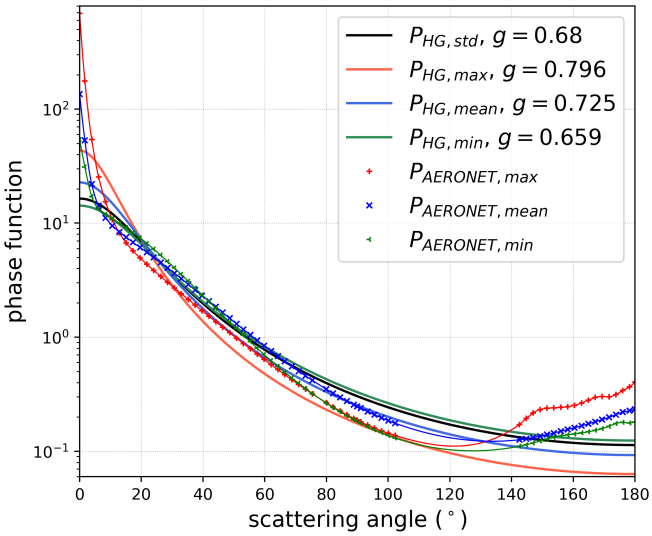


Fig. 4.40: *Comparison of AERONET and Henyey-Greenstein phase functions. Mean, min, and max asymmetry factors g and measured AERONET phase functions are shown as solid lines and data points, respectively. Thin solid lines depict the Legendre series expansion of the measured phase functions.*

Test results with varying aerosol parameters can be seen in Figure 4.41. Here, the relative difference between the results from BÖSCH et al. [2018] and the tested parameter results[8] are shown in percent for the AOT (top) and the surface extinction (bottom) for the five discussed days. The right-hand-side sub-plots show the relative difference between the profiles at 09:00 and 16:00 UTC. Only one parameter was changed. The other quantities were chosen to be as those from the previous results.

The red lines show the results for a change of the single scattering albedo. The line represents the mean SSA while the shaded area shows the minimum and maximum SSA results (see Tab. 4.10). It can be seen that especially in forward scattering directions the differences can be large for the AOT ($\pm$ $\sim$12 %) and are usually smaller for surface extinctions ($\pm$ $\sim$6 %), under cloud free conditions. Here, depending on the scattering geometry, positive and negative deviations compared to the fixed values of SSA$=0.92$ are possible for min and max. For the AOT, an U-shaped relative difference can be observed, indicating that the geometry, and therefore, the phase function describes the general shape while the SSA defines the amplitude and scaling of this shape. For the surface extinction, and increasing trend can be observed with negative differences in the morning and positive values in the evening.

The blue lines show the mean asymmetry factor results while the blue shaded area depicts quantities for the minimum and maximum asymmetry factor (see Tab. 4.10). Here, the differences are significantly larger showing the importance for a proper phase function representation. Relative differences up to -41 % are reached for the AOT and -50 % for the surface extinction, meaning that for the fixed value of $g = 0.68$, the profiles are highly underestimated. On the other hand, since the fixed value was closer to the minimum asymmetry factor, positive relative differences are much smaller in magnitude showing that the retrieved profiles are only in rare situations smaller than the results found in BÖSCH et al. [2018]. However, a clear daily, and therefore, scattering angle dependency can only be observed for the AOT with larger differences in the afternoon, when forward scattering

[8] Relative difference of quantity x: $(x_{paper} - x_{new})/x_{new} \cdot 100$.

is dominant. The surface extinction does not show a clear tendency but rather variable results when the mean or maximum asymmetry factor is used. This might be due to two reasons. First, the aerosol retrieval uses dSOT differences as state vector which can be considered as the main minimization quantity and are assumed to be more trust-worthy than specific profile values. This leads to the second reason: the bottom extinction is just one value which has to be adapted with respect to a certain regularization ratio and is therefore correlated to a certain amount with extinction values in other altitudes. Nevertheless, peaks in the blue lines or shaded areas indicate either problematic scenarios of the standard parameter result or of the new profiles since temporal changes should affect both parameter settings in a similar way.

The green lines show the results with applied mean, minimum and maximum phase functions (see Fig. 4.40). Since these phase functions were selected based on the $0°$ angle, the min and max results can not be considered as minimum and maximum uncertainty range for all geometries. Because of that, pointed and dashed line styles were used as a representation of the min and max phase function, respectively. The SSA value was kept at $SSA = 0.92$ for this test (similar to the published results). AOT as well as surface extinction results are mostly underestimated when using P_{HG} instead of the AERONET phase functions, indicated by the negative relative differences. However, a temporal dependency can be observed for all green curves. From a qualitative perspective, all curves show three local minima for the AOT, in the morning, around noon and in the evening. The minima are generally in the negative difference range but, for the mean phase function, also positive relative differences can be observed. The largest negative relative differences can be identified around noon ($\sim -25\,\%$) indicating that not only forward and backward scattering might be inaccurate when using Henyey-Greenstein but also the in-between scattering angles might have a strong impact on the results. The morning differences are usually $< \pm 5\,\%$ for min and mean but up to $13\,\%$ for the max phase function. In the evening, relative differences of $13\,\%$ can be found for the mean curve, while it is up to $-20\,\%$ for the max phase function. The bottom extinction curves are more variable which is in agreement with changes of the asymmetry factor. The relative differences are usually $< \pm 25\,\%$ but can exceed this value for specific measurements. Again a large negative relative difference can be found around noon for all curves ($< \sim 25\,\%$) indicating the underestimation of the profiles when using Henyey-Greenstein. The early morning and late afternoon results seem to be highly variable with a change in sign within few hours for all days ($\pm\ \sim 30\,\%$). This is interesting as it indicates that also one measured but fixed phase function might be insufficient for an accurate treatment of the aerosol scattering distribution throughout a longer time period.

The orange lines show the before discussed mean, min and max phase function in addition to the mean, min and max SSA values. Again, the SSA is less important compared to the phase function but strong deviations for specific measurements are possible which might

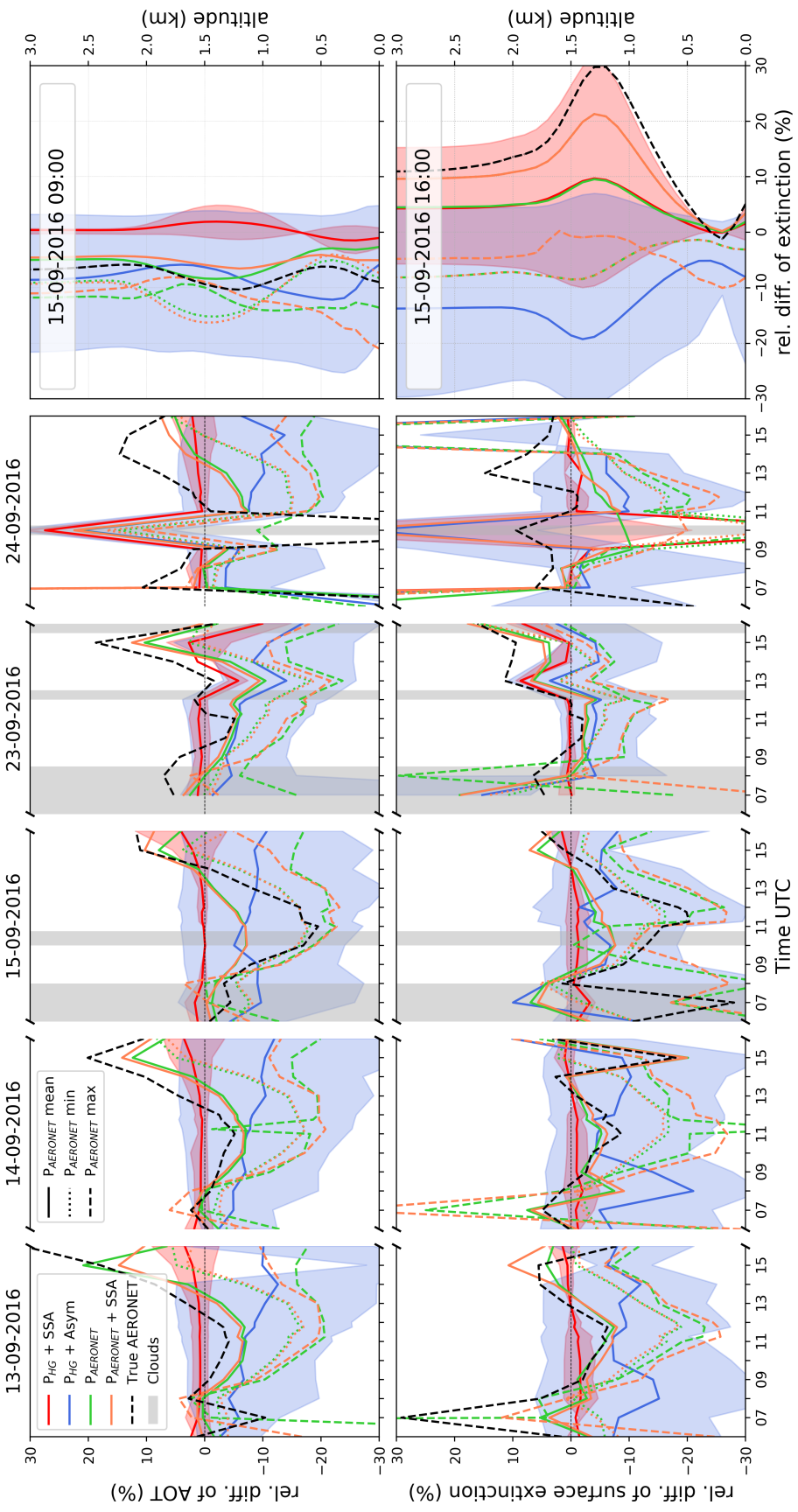


Fig. 4.41: *Relative differences between the results from* BÖSCH et al. [2018] *and the results when certain parameters are changed for five different days during the CINDI-2 campaign.* ***Left:*** *Curves for relative differences of the AOT (top) and of the surface extinction values (bottom).* ***Right:*** *Relative differences of two profiles of the 15th of September at 09:00 (top) and 16:00 UTC (bottom).*

be due to problematic measurements or because of the insufficient treatment when applying fixed aerosol quantities. The dashed, black lines in Figure 4.41 show the true AERONET aerosol informations (phase function and SSA) closest in time to the specific measurements. Due to the availability of AERONET data, 7 different phase functions and SSA values can be used for the days 13.09. and 14.09. (each day from 7:00 to 17:00) while only 3 data points were provided for the 15th of September (from 9:38 to 13:38). On the 23rd of September, four data points are available (08:35 - 12:35) but no data was provided on the next day by the AERONET webpage[9]. The curves follow the discussion of the phase function above with some smaller deviations. For nearly all days, the AOT around noon was highly underestimated by using Henyey-Greenstein (up to $-20\,\%$) while the afternoon measurements led to an overestimation (up to $+30\,\%$). The behaviour in the morning differs throughout the investigated days. Relative differences to the standard settings are around $\pm 10\,\%$, under cloud free conditions. For the surface extinction curves, the behaviour follows mostly the phase function discussion above. In general, the results are close to the mean phase function curves (solid green or orange curves, 13. - 14.09.) but larger deviations to all curves can be found at the 15th of September and in the morning of the 13th. From the 13th to the 15th of September, the surface extinction was mostly underestimated with some exception (morning of 13.09.). For the two cloudy days (23. - 24.09.), positive relative differences are dominant. In general, under cloudy conditions, even larger differences for all aerosol parameters are possible (24.09.) but cannot always be found (23.09.). Even though the temporal coverage at the 15th of September is lower than at the first two days, the relative differences are comparable around morning and evening but slightly larger around noon (when the AERONET data was provided).
The comparison of the specific profile differences (two right-hand side sub-plots in Fig. 4.41) show that the altitude dependence can vary. At 09:00 there are only smaller differences for varying altitudes while the 16:00 profile shows a large sensitivity on the specific aerosol parameters in an altitude between $1\,\mathrm{km}$ to $1.5\,\mathrm{km}$. Here, also the SSA variation has a larger impact on the profile shape with relative differences up to $\sim 30\,\%$. A change of asymmetry factors leads to larger negative differences.

As an intermediate conclusion it can be stated that accurate aerosol information should be used within aerosol MAX-DOAS retrievals and only in the absence of AERONET data, use of the Henyey-Greenstein approximation is suggest. However, in this case, a proper asymmetry factor is much more important than an accurate SSA value. In general, mean values can be used but should be chosen carefully for a specific location and prevalent aerosol type as results might vary in time when the aerosol properties and geometry changes. Measured phase functions should always be selected if available but the larger relative dif-

[9] https://aeronet.gsfc.nasa.gov/cgi-bin/data_display_inv_v3?site=Cabauw&nachal=2&level=2&place_code=10&DATA_TYPE=76

ferences indicate that the usage of an early morning phase function around noon or in the afternoon might be even worse than Henyey-Greenstein.

In order to prove the statements above, the Figures AF.54 - AF.57 in the appendix show time series as well as correlation plots for the AOT, the surface extinction, the VCD and the surface concentration for the previous results from BÖSCH et al. [2018] in red and for the same settings but with the actual AERONET phase functions and SSA values in light green. The regression line parameters and Pearson correlation coefficients can be found in Tables AT.6 and AT.7. The correlation coefficients are also given in the corresponding plots . Here the numbers in brackets are for the convergence-filtered data points while the numbers outside of the brackets were calculated for the full datasets. Dark colours correspond to the old settings while bright colours indicate the retrieval with proper AERONET data. Note that the correlation numbers of the old BOREAS results in red do not match those from similar figures published within the AMTD paper. Here, an error was found because the published numbers include also data from the 12th of September.
It can be seen that the use of accurate AERONET data always deteriorates the correlation coefficients when applying the same settings. For the integrated values (AOT, VCD) the decrease in correlation is small while the bottom values show a larger deviation. The temporal comparison of both curves for all plots indicate that the deviations are large only for the (partly) cloudy days. The results for the 13th and 14th are similar for the bottom values and more or less identical for the integrated values. The remaining three days show the largest differences for measurements which were marked as cloudy. However, some features as e.g. the increase of AOT at the 15th before noon can be seen with the new data. The comparison of the filtered and unfiltered correlation cover these findings as the values are strongly improved when the cloudy (and often not converging) scenes are filtered. Nevertheless, the correlation coefficients from the results with included AERONET data are at best similar high but sometimes significantly lower (cf. AOT and NAQMN surface extinction correlation). At this point, no clear reason for this results can be found. However, the applied settings were optimized with respect to the Henyey-Greenstein approximation. One possible explanation would be that the old settings with an appropriate aerosol description demand the need for larger constraints within the retrieval. **Test 3** will introduce new settings for the *prev*-mode (cf. **Test 4** in Sec. 4.3.1.1) which were optimized with respect to AERONET aerosol information.

Test 2: Impact of pressure and temperature on the retrieval

It was already stated by WANG et al. [2017] that accurate pressure and temperature profiles might have a large impact on the retrieval results. In order to confirm this statement, different climatology treatments are tested below.

Within the publication introduced in the previous section, vertical profiles of pressure and temperature were used, created by taking the mean of 16 different sonde measurements taken during the years 2013-2015 in De Bilt[10] (the Netherlands). In addition, the US standard atmosphere was used (NASA, 1976). Furthermore, the before mentioned NO_2 sondes (see Sec. 4.4.1) also measured pressure and temperature profiles which were used below[11]. These profiles were measured only up to the tropopause so that KNMI radio sonde[12] data was used in order to extend the vertical profiles of the NO_2 sondes to higher altitudes for the RTM calculations. Since the distance to Cabauw was rather small, no larger differences in pressure and temperature can be expected for the stratosphere. Because these profiles are not accurate in the lower troposphere when daily variations are considered, surface pressure and temperature measurements from the CESAR observatory (CESAR, 2018) were used in a similar way as suggested by WAGNER et al. [2018] and are explained as follows:

1. A constant lapse rate γ is calculated with the assumption of a certain tropopause temperature (T_{tp}) and height ($z_{tp} = 11\,\mathrm{km}$) and the surface temperature T_0: $\gamma = (T_{tp} - T_0)/(z_{tp} - z_0)$.

2. The temperature profile at z_{tp} and above follows an underlying atmospheric profile as e.g the US standard atmosphere or data from the radio sondes.

3. Then, this lapse rate is used to calculate the pressure profile up to z_{tp} by:

$$p(z) = p_0 \left(\frac{T_0 - \gamma(z - z_0)}{T_0} \right)^{g/(R_d \gamma)} . \tag{4.19}$$

   Here, g is the gravitational acceleration ($g = 9.81\,\mathrm{m/s^2}$), p_0 is the surface pressure at the surface altitude $z = 0$ and R_d is the gas constant for dry air ($287\,\mathrm{J/K/kg}$).

In Figure 4.42, different pressure, temperature, and O_4 profiles are shown for the before mentioned datasets. The US standard atmosphere (black) shows the lowest tropospheric temperature profile, and therefore, the largest O_4 profile. The mean KNMI sonde profile (red) shows slightly larger temperatures with more altitude depending variations. This directly transforms into a slight decrease in O_4 concentration values. The comparison with sonde profiles, measured at this specific date, show that the temperature profile varies strongly with different altitudes. Even though the KNMI radio sonde measured in a larger distance to Cabauw, similar vertical features compared to the NO_2 sondes can be observed. The strong increase in temperature at around $2.5\,\mathrm{km}$ is interesting as it seems to be stable in time and it might affect the profiling results because the mean KNMI profile does not

[10] The data was kindly provided by François Hendrick and Marc Allaart from KNMI.
[11] The data was kindly provided by Mirjam den Hoed, KNMI.
[12] The sondes were launched at KNMI (de Bilt, the Netherlands) with an approximate distance of $24\,\mathrm{km}$ to Cabauw. The data was kindly provided by KNMI.

follow this pattern. The larger temperatures in the troposphere found by the sondes are directly linked to smaller O_4 values which will lead to a decrease in extinction for the aerosol profiles within the retrieval (compared to the mean KNMI profile). The impact of the diurnal temperature variation can be seen when comparing the two NO_2 sonde curves near the surface. Here, the US standard atmosphere as well as the mean KNMI profile are closer to the morning values, deviations around noon and in the afternoon can be expected.

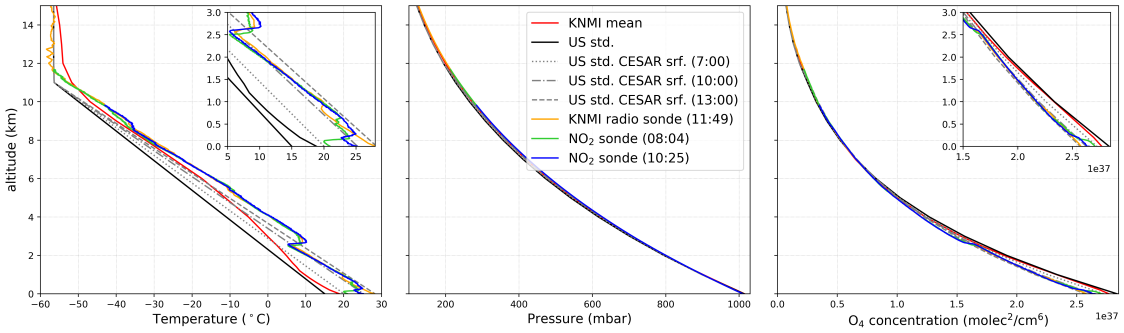


Fig. 4.42: *Temperature (**left**), pressure (**mid**) and O_4-concentration (**right**) vertical profiles for different data sets. The smaller inlay sub-plots show zoom-ins into the lowermost $3\,\mathrm{km}$.*

The grey dashed and pointed lines represent the before mentioned surface scaling with the CESAR surface temperature and pressure values for 07:00, 10:00 and 13:00 UTC. The US standard atmosphere was used for altitudes larger than $11\,\mathrm{km}$. It can be seen that even though the surface temperatures are similar to the sonde profiles, the O_4 concentration profiles differ with altitude as the real temperature profile is not as linear as the surface scaling approach assures.

Figure 4.43 shows the relative differences of AOT and surface extinction values for the 15th of September for the standard approach of the paper and the before mentioned different pressure and temperature profile dataset results. On the right-hand side, relative differences of profiles at 10:00 and 16:00 UTC are depicted. The use of the US standard atmosphere is the only climatology approach which leads to negative relative differences indicating an increase in both, the AOT and the surface extinction. The two NO_2 sonde lines (solid green and blue) show a slight increase on AOT in the morning, a more or less constant slope from 08:00 to 15:00 but a strong increase for 16:00 UTC. The surface extinction shows larger deviations between these two lines but the general behaviour is similar. Around noon, the largest relative differences can be seen with values up to $12\,\%$. The differences in the morning are the smallest while the afternoon differences are in between the noon and the morning differences here. The dashed curves show the surface scaling approach. The only difference is the underlying profile above $11\,\mathrm{km}$ where the surface scaling

ends. Here, the KNMI mean, the US standard atmosphere as well as the radio and NO_2 sonde differ only slightly (cf. Fig. 4.42). However, no larger differences between the three curves can be identified indicating that both, the choice of the tropopause height as well as the climatology profiles above do not have a large impact on the results. However, a strong increasing trend for all curves in time can be observed indicating that the warming up of the lower troposphere throughout the day directly propagates through the retrieval by decreasing the overall aerosol profiles. As the diurnal variation of AOT and BOT curves in the previous section was found to be good, this result is highly questionable. Even though the surface temperature and pressure values can be assumed to be accurate, the wrong vertical profile seems to lead to a diurnal bias which should be avoided. This is an unwanted result as it still does not answer the question of which climatology profiles should be used when no accurate temperature and pressure profiles are available. Since the number of sonde launches during CINDI-2 was limited, the mean KNMI profile is still assumed to be the best option so that no changes in the climatology treatment can be suggested here.

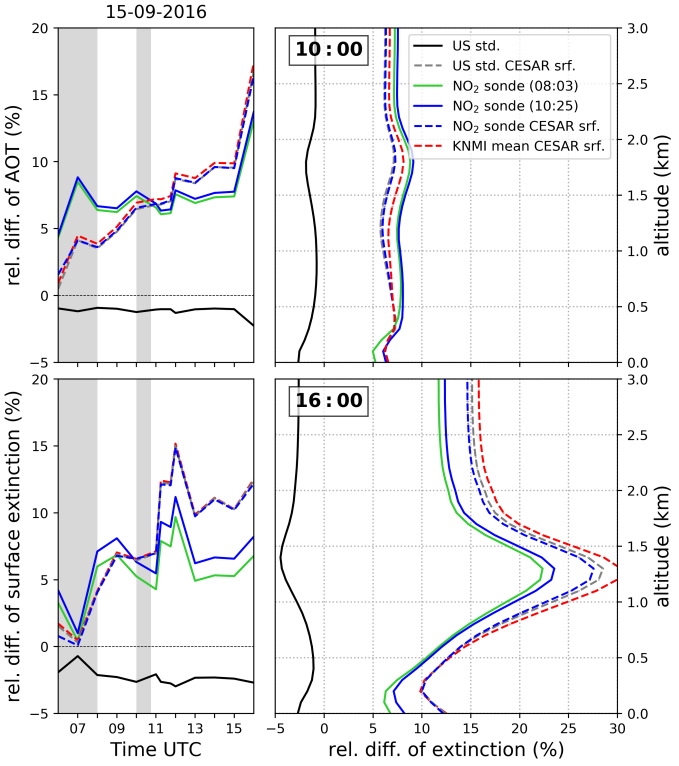


Fig. 4.43: *Comparison of profiling results for different approaches used as climatology within the profile retrieval.* ***Left****: Relative difference of the standard settings and the specific climatology result for the AOT (top) and the surface extinction (bottom).* ***Right****: Relative difference of profiles at 10:00 and 16:00 UTC.*

Test 3: Improvements of settings and application of the LM option

Here, the findings from the synthetic study (see Section 4.3) are used in order to improve the retrieval settings, and, therefore, the correlations with ancillary measurements from the previous sub-sections.

Furthermore, the appropriate aerosol information from the AERONET web page (see **Test 2** above) are applied, meaning that the phase function and the SSA closest in time to the specific measurement were used. For the aerosol retrieval, the settings were changed as

follows:

- The retrieval mode was changed from *apri* to *prev* (see **Test 4** in Sec. 4.3.1.1).
- The a priori variance was slightly decreased in the lower $2\,\mathrm{km}$ from $150\,\%$ to $100\,\%$ because strong deviations from the pre-scaled a priori can not be expected.
- The Tikhonov parameter was increased from 2 to 15 which accounts for the *prev* mode.
- The SNR was slightly increased from 2500 to 3000.
- Since the AOT was underestimated compared to AERONET, the exponential a priori profile was defined with a scale height of $1.25\,\mathrm{km}$ instead of $1\,\mathrm{km}$.

The trace gas retrieval was performed by applying the LM option (see Sec. 4.3.2.1). Here, the a priori profile scale height was changed as well. Since a slight overestimation compared to the Pandora instrument was found (cf. Sec. 4.4.1.2) the scale height was reduced from $1\,\mathrm{km}$ to $0.75\,\mathrm{km}$. It might appear inconsistent to use different scale heights for the aerosol and the trace gas retrieval. However, there is no reason to believe that both atmospheric constituents have the same origin and do follow similar spatial distributions. Furthermore, the most important sink for aerosols is wet deposition which means that strong vertical mixing is possible in the absence of precipitation. On the other hand, NO_2 is located mostly in the PBL as the lifetime (in the absence of reservoir species) is usually not long enough for mixing on large spatial scales.
In addition to the change in a priori scaling and the new LM retrieval option, only the variance was slightly changed within the trace gas retrieval. Here, the stronger altitude dependence was removed as the improvements were rather small. The variance was set constant to 0.5 for altitudes smaller than $3\,\mathrm{km}$.

The Figures 4.44 - 4.47 show again the time series for relevant parameters for the results published in BÖSCH et al. [2018] (red) and the new settings including accurate AERONET phase functions and SSA values in light green.

The first plot depicts a slight improvement of the BOREAS AOT underestimation which was found before (cf. Sec. 4.4.1.2). On the 14th and 15th of September, the green curves follow the AERONET data better than the original results. However, no strong improvement was found for the other days. The larger differences in the afternoon are still present showing that this is not due to an inaccurate phase function. The correlation coefficients for the unfiltered dataset (outside the brackets) is slightly smaller which can be explained by the larger deviations on the 15th, especially in the morning, when clouds are present. The correlation for the filtered data (inside the brackets) has also not changed significantly. However, more profiles converge with the new settings which might affect the correlation as hard-to-retrieve-scenarios might not be filtered any more, which might

decrease the correlation even though the curves seem to be improved by eyes. Nevertheless, an underestimation of BOREAS' AOT compared to AERONET can be observed but not explained. Here, the reader is referred to the publication of WAGNER et al. [2018] who discussed the need for an O_4 scaling factor in order to cope for differences in measurement and RTM calculations in specific cases. Here, no further analyses are done on this matter.

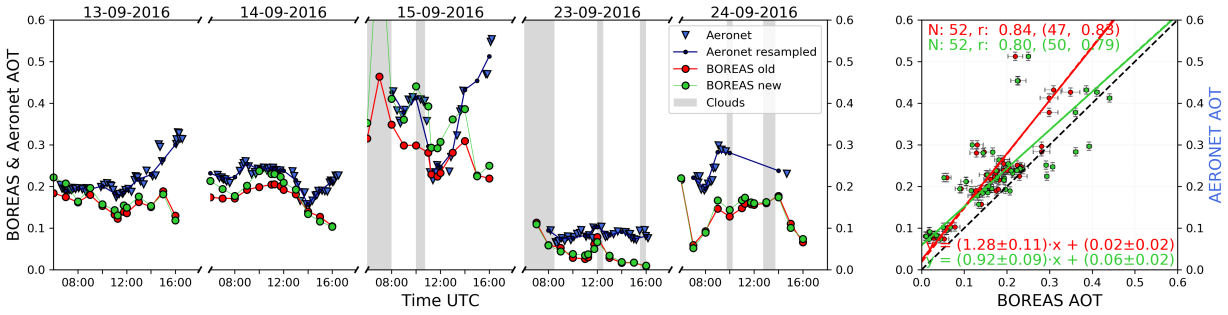


***Fig. 4.44: Left**: time series of AERONET (blue) and BOREAS AOT (red: paper data, green: new settings with AERONET phase fct and SSA). Small triangles show the original AERONET measurement, small dots with connecting lines depict the resampled data. Grey areas indicate clouds. **Right**: scatter plot of both datasets including parameters of the orthogonal regression and Pearsons correlation coefficient. The numbers outside (inside) of the brackets show correlations for unfiltered (filtered) data. The solid (dashed) lines show the regression lines of the unfiltered (filtered) datasets.*

Figure 4.45 shows the time series of surface extinction values. Here, the results on the first two days are again similar but some smaller differences can be observed in the mornings. The results on the remaining days are more variable. On the 15th, a strong overestimation was found in the morning, when clouds were present. Later that day, the bottom extinctions show a more variable behaviour compared to the previous results. Here, it is hard to say which profiles are better as both curves do not follow the Ceilometer or the in situ station. On the 23rd, a more or less constant extinction level is found, except for the first two data points. Here, the reader should note that an issue for small AOT was fixed within BOREAS aerosol retrieval. For AOTs smaller than 0.1, the retrieval sometimes did not converge and led to strongly oscillating profiles due to the limited information content for such small aerosol loads. In theses cases, BOREAS increases smoothing (by increasing the Tikhonov parameter) when three oscillations in the RMS of dSOT differences are found together with AOTs smaller than 0.1. The comparison with ancillary Ceilometer measurements show that the newly retrieved results can be considered as more accurate. For the 24th, the old and new results are similar except for the morning values which follow the in situ data better than before. Additionally, this means that the agreement with Ceilometer near-surface values is not as good any more as both ancillary datasets differ strongly on this day.

The correlation coefficients for the new profiles results with NAQMN and Ceilometer are generally better in the unfiltered cases but a slight deterioration can be observed for NAQMN with the filtered dataset compared to the old BOREAS results. Here, again the number of converging profiles is larger which might affect the coefficients.

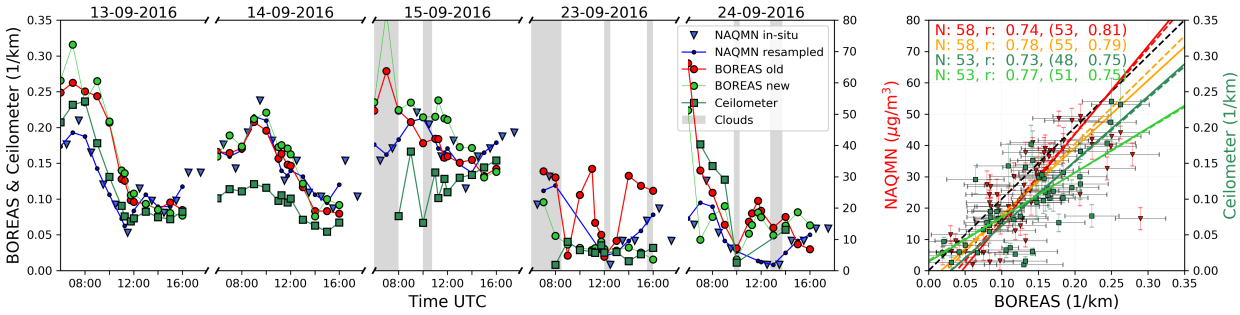


***Fig. 4.45: Left**: time series of BOREAS (red: paper data, light green dots: new data with AERONET phase fct and SSA), ceilometer (dark green squares) and NAQMN in situ (blue) and near surface aerosol parameters. Small triangles show the original hourly NAQMN measurements, small dots with connecting lines depict the resampled data. Green squares show ceilometer near-surface extinction values evaluated by averaging the 10 - 50 m coefficients.* ***Right**: as in Figure above. Regression parameters can be found in Tab. AT.6.*

Figure 4.46 depicts the results of BOREAS NO_2 vertical column densities. For all days, the differences between old and new results are small even though the applied approach has changed. As a reminder, the published dataset used the optimal estimation technique in the linear space while the new results were created by applying the Levenberg-Marquardt algorithm in the logarithmic space (see **Test 9** in Sec. 4.3.2.1).

The correlation with Pandora VCDs has slightly decreased in the unfiltered case but it has also led to a small improvement for the filtered results. The correlation with Lidar data stayed the same without filtering but it has also deteriorated with filtering which is supposed to be only due to the smaller number of data points.

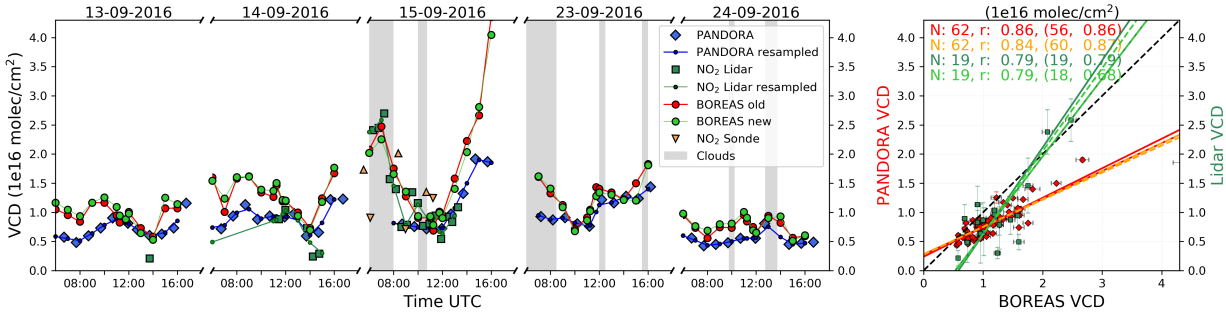


***Fig. 4.46: Left**: time series of Pandora (blue diamonds), LIDAR (green squares) and BOREAS NO_2-VCD (red: paper data, light green dots: new data with AERONET phase fct and SSA). Small dots with connecting lines depict the resampled datasets of Pandora and LIDAR. Orange triangles are the representation of the integrated NO_2 sondes measurements with the ascend and descend separated in triangles with the edge to the top or to the bottom, respectively.* ***Right**: as in Figure above. Regression parameters can be found in Tab. AT.7.*

Figure 4.47 shows the results of the NO_2 near-surface concentrations but the CAPS/ICAD in-situ dataset was removed compared to the published figure (Fig. 4.38) in order to improve the clarity of this plot. In this graph, the general results of both curves do follow similar patterns but smaller differences can be identified. The morning values of the new results are usually smaller than those from the published dataset except for the 15th of September. Here, the new bottom concentrations follow the lowest LP DOAS curves better than before. However, the general underestimation of BOREAS concentrations for the early morning is still present (cf. Sec. 4.4.1.2).
The correlation with the NAQMN in situ station and the LP DOAS instrument has slightly decreased while it has improved for the Lidar data.

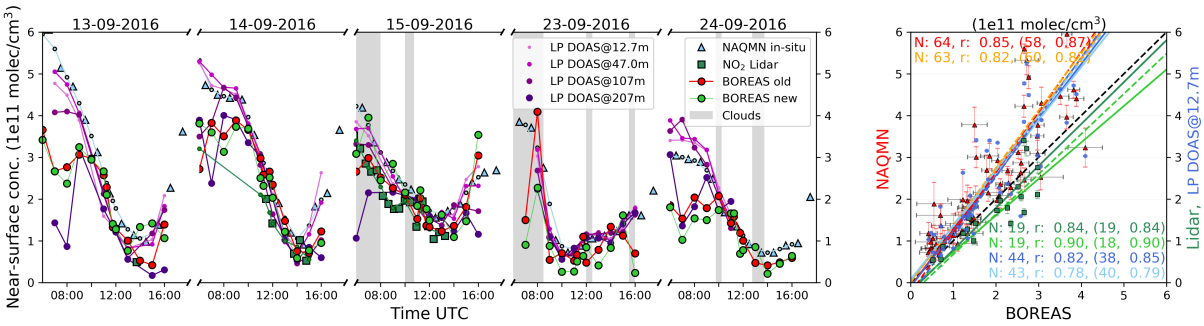


***Fig. 4.47: Left**: time series of in situ and remote sensing NO_2 near surface concentrations. Triangles show in situ instruments for NAQMN as light blue triangles. The NO_2 LIDAR is plotted as green squares and the LP DOAS as different sized circles in shades of magenta (lowest altitude as smallest circle). BOREAS is shown as red (paper) and light green dots (new data with AERONET phase fct and SSA). **Right**: as in Figure above. Regression parameters can be found in Tab. AT.7.*

As a conclusion it can be noted that the new settings led in total to similar correlations compared to the old results even though a more accurate aerosol description was applied. However, the correlations compared to the change in aerosol phase function and SSA only has strongly increased indicating that also an improvement of the optimal estimation results should be possible (e.g. by increasing the smoothing constraints). The general differences (such as the general underestimation of AOT or bottom concentration in the morning) to ancillary measurements could not be fixed with this new dataset showing that this is rather an issue due to different air masses probed, geometry or the limited vertical resolution of MAX-DOAS profiling.

Test 4: NO_2 in the UV

As a lookahead on the discussion of NO_2 profiles retrieved in the visible and UV spectral range for Bremen data and a similar discussion presented in the upcoming paper by Tirpitz et al. (2018), a brief comparison of results during the CINDI-2 campaign is presented. While the NO_2 profiles should be similar for all wavelengths under the assumption of a

homogeneously distributed trace gas layering, the aerosol extinction coefficient profile depends strongly on the wavelength (cf. Eq. 2.95). Here, a qualitative profile comparison for the 14th of September only is shown.

Figure 4.48 shows the profiles, AOT and bottom extinction values for the results retrieved in the visible and ultra-violet spectral range. Note that only for the profile at 07:00 UTC in the UV, no convergence was reached.
The comparison of profile shapes indicate that similar temporal features were retrieved for both spectral ranges. In the morning, the aerosol load is located close to the surface with large extinctions. Throughout the day, the lowest values decrease while an elevated layer seems to form. This is more pronounced for the UV, where the AOT is generally larger. Note that for the VIS profiles between 14:00 and 16:00 UTC the AOT is smallest indicating that these profiles were retrieved with a larger Tikhonov parameter (see explanation in

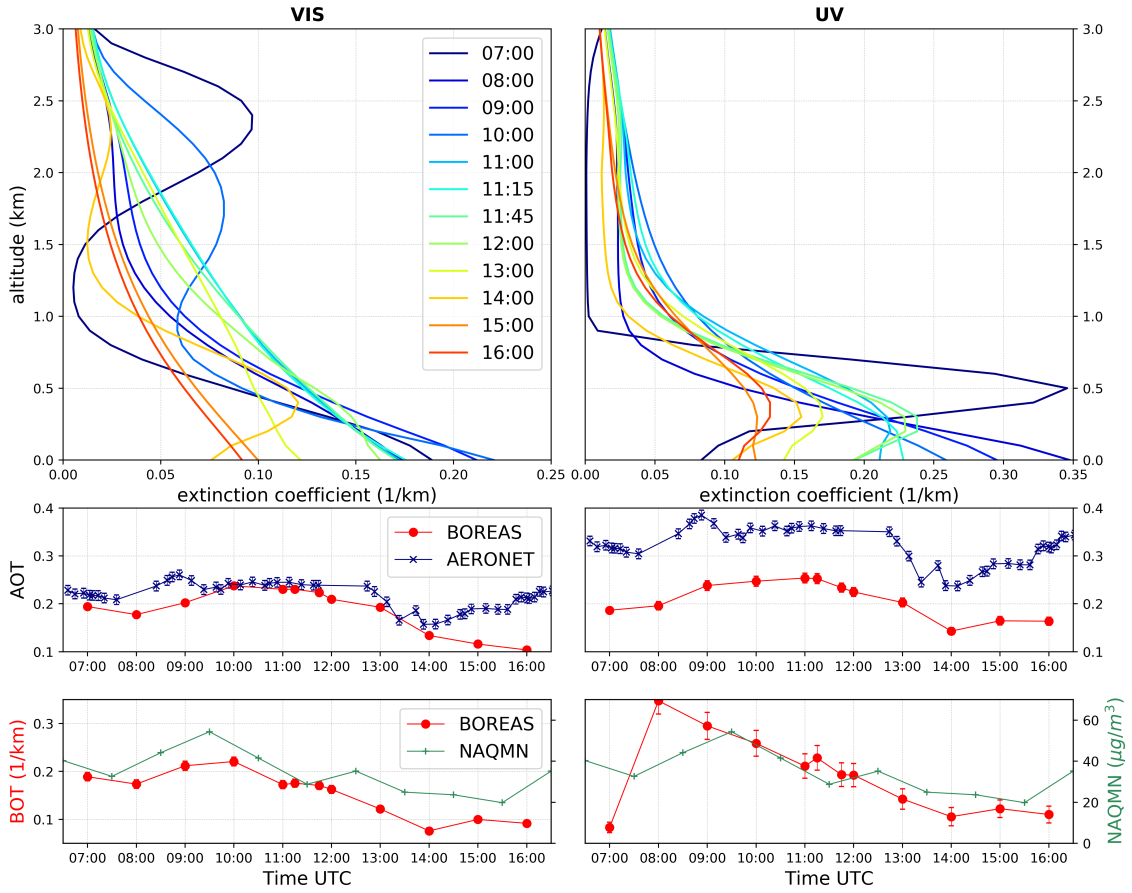


Fig. 4.48: *Comparison of aerosol profiles (**top**), AOT (**mid**) and bottom extinction (BOT, **bottom**) in the visible (**left**) and ultra-violet (**right**) spectral range. AERONET AOTs were calculated on* 477 nm *and* 360 nm *based on the Angström exponent between* 440 nm *and* 675 nm. *PM10 in situ concentrations were plotted from the NAQMN site in Cabauw.*

Test 3). The elevated layers do not necessarily have to be detached from the surface. The retrieval response for a box-like profile would also vary between an exponential-like shape and oscillating features where the box can be assumed, depending on the regularization parameters. Here, it is quite likely that a rising PBL is accompanied by box-like aerosol vertical distributions which propagate through the retrieval to appear as elevated layers. An example for this behaviour can be seen for synthetic data in Figure 4.18 for aerosols and in Figure 4.28 even better for NO_2.
The comparison of BOREAS with AERONET AOTs reveal that the agreement is much better for VIS than for UV. The VIS results show more or less similar AOTs with larger deviations in the afternoon. However, the UV values are strongly underestimated by BOREAS. Two possible reasons can be found. Either horizontal inhomogeneities were present or the applied AERONET Angström exponent for the linear extrapolation to $360\,\mathrm{nm}$ was not an accurate representation of the wavelength dependency in the UV. Nevertheless, the temporal curve behaviour is generally well described for both spectral ranges but the strong increase in the afternoon is not found as strongly for BOREAS as the AERONET data indicates. The reason for that is still unclear.

Figure 4.49 shows a similar plot as above but for NO_2. Here, the general profile shapes and temporal variations are similar to the aerosol profiles but the main concentration is closer to the surface for the whole day. Again, an increase in PBL might be accompanied by profiles which seem to find layers detached from the surface. In contrast to the aerosol profiles in the visible range, the NO_2 profile do also show this elevated layer for the afternoon profiles supporting the argument of an increased Tikhonov parameter due to low AOTs. The comparison between both spectral range results show that the concentrations in the morning are not the same. The 10:00 UTC profile is the first result which appears to be similar for VIS and UV. Here, the before mentioned limited vertical resolution might have an impact on the results when the aerosol and trace gas is mainly concentrated close to the surface. Furthermore, since the effective light paths in the UV are much shorter, deviations because of horizontal inhomogeneities are possible.
The VCD comparison shows similar columns for both spectral ranges but an overestimation compared to the Pandora tropospheric column. It is interesting to see that even though the bottom concentration differs in the morning, the vertical columns agree well. The bottom concentration comparison with the NAQMN in situ data reveals a general good agreement between both instruments and for the two spectral ranges with the exception of the first three morning values. Especially, the 07:00 UTC profile in the UV deviates stronger from NAQMN and VIS results which is due to the before mentioned missing convergence here.

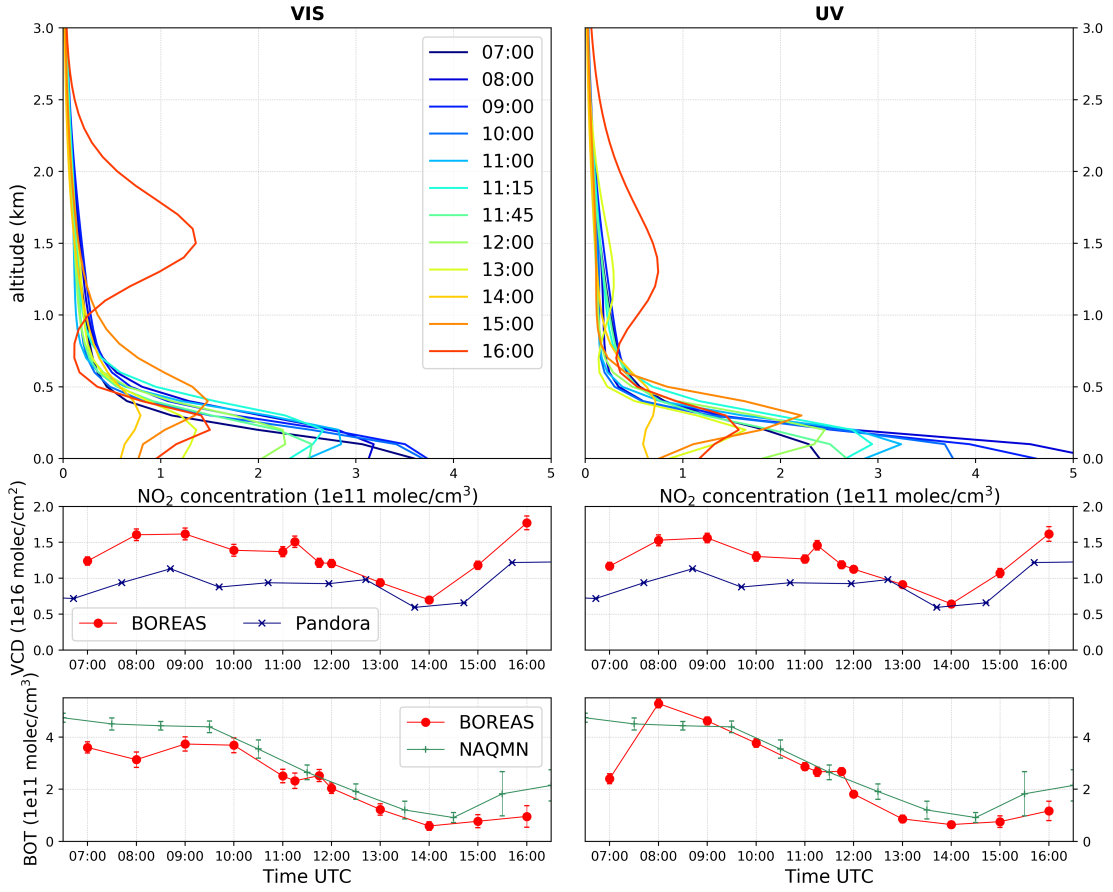


Fig. 4.49: *Comparison of NO_2 profiles (**top**), VCD (**mid**) and bottom concentration (BOT, **bottom**) in the visible (**left**) and ultra-violet (**right**) spectral range.*

In general, the agreement between the results of both spectral ranges is good showing that similar results can be expected when inhomogeneities can be excluded. Furthermore, the NO_2 profiles are similar even though the AOT comparison indicated that the BOREAS UV AOT might be underestimated stronger than the VIS AOT. This leads to two different conclusions. Either the underlying aerosol profile is not as important as it was expected or the calculation of the $360\,\mathrm{nm}$ AERONET AOT was wrong due to an inaccurate UV Angström Exponent. Here, no finale statement about this discrepancy can be made, further research is necessary.

4.5 Retrieval of profiles of MAX-DOAS measurements in Bremen

After the successful validation of the BOREAS algorithm, a longer time-series of aerosol, NO_2 and $HCHO$ profiles is presented here. The applied settings were introduced in the previous Section 4.4.2 in **Test 3**. However, even though it was shown that the aerosol phase function has a an impact on the retrieval, no AERONET data is available for Bremen. The closest AERONET measurement site is Hamburg (distance $\sim 93\,\mathrm{km}$) which is assumed to have similar atmospheric and weather conditions. Due to the generally large percentages of cloudy scenes, the availability of aerosol information throughout the year is strongly limited. Because of this lack in an accurate temporal coverage of AERONET phase function and SSA, the Henyey-Greenstein approximation is applied with mean SSA and asymmetry factor values calculated from the AERONET data of the Hamburg site. The values were found to be 0.955 and 0.710 (VIS), and 0.956 and 0.734 (UV), for SSA and g, respectively.
It was already shown that direction b and c led to similar results so that the retrieval is only applied to direction c to decrease the computational time. Direction s was also discarded as it suffers strongly from pointing towards the sun which implies the necessity of knowledge of accurate aerosol phase functions in order to retrieve realistic profiles.
The retrieval was started for the fit windows VIS and UV1 (see Tab. 3.3) due to computation time reasons and because UV2 led to unstable and sometimes questionable results in the winter months (see Sec. 3.8.3).
The resulting aerosol profiles retrieved in the visible and ultra-violet fitting windows are compared with AERONET results and with data from the BLUES network, Bremen (see Bremen map, Fig. 3.2). The trace gas profiles are additionally compared to the BLUES in situ dataset but also to the vertical column and onion peeling results from the previous chapters (see Sec. 3.8 and 3.7).
Note that several filters were applied to the dSCD data (Sec. 3.5.2). Furthermore, since the applied filters still do not account for all problematic scenarios, measurements, and atmospheric conditions, various filters are additionally applied to the profiling results and introduced in the following sub-sections.

4.5.1 Retrieval of aerosol profiles

For the year 2015 - 2017, aerosol profiles were retrieved at $360\,\mathrm{nm}$ and $477\,\mathrm{nm}$ for the fitting windows VIS and UV2 introduced in Table 3.3. BOREAS' aerosol retrieval was applied with the *prev*-option with including a priori pre-scaling. The unscaled a priori profile was exponential with a scale height of $1.25\,\mathrm{km}$ and a surface extinction of $0.183\,\mathrm{km}^{-1}$. The a priori variance was set to $100\,\%$ and the Tikhonov parameter to 15 and 5, for VIS and

UV, respectively. The signal to noise ratio is input for the measurement covariance matrix was fixed to 3000.
The resulting profiles were filtered based on the following criteria:

1. Profiles were discarded if the maximum iteration number was reached.
2. If the degrees of freedom were smaller than 1.0, the profiles were removed as this is a sign for bad weather conditions.
3. Profiles retrieved for measurements with an SZA $> 85°$ were neglected.
4. Additionally, profiles later than 18:00 LT were filtered as for these measurements the instrument was pointing towards the sun.
5. The RMS of O_4 dSCD differences was not allowed to be larger than 1000.
6. Profiles with AOTs larger than 1.0 were discarded as these led to poor convergence, e.g. due to difficult weather conditions (e.g. fog).

Several of these filters account for the same problematic scenarios but some of them indicate surprising difficulties. For example, the AOT filter was necessary as some results led to unrealistically large extinction profiles even though no other filter criterion was able to remove these profiles. However, since neither AERONET AOT nor the in situ PM10 measurements indicated large aerosol loads for any of these days, the specific measurements are assumed to suffer from unknown issues (e.g. fog, clouds). Note that data from ancillary measurements was not allowed to be farer away in time than 1 hour compared to the MAX-DOAS scan.

The Figures 4.50 and AF.58 (appendix) show time series of monthly mean AOTs of BOREAS and AERONET products as well as the correlation between the datasets for VIS and UV, respectively. Note that two different datasets are depicted for AERONET. The first dataset consists of the Almucantar inversion product (Level 2.0, Version 2) which is referred to as AERONET inversion here. The second dataset is the AERONET Version 3 AOD (Level 2.0) direct sun product. For the first one, less values are usually available as the inversion product needs further filtering in order to apply the retrieval from specific Almucantar scans. Thus, a clear sky for all sky radiance measurements can be assumed for the inversion products while the direct sun values needs only proper conditions for one measurement in the direction of the sun. Both datasets were interpolated to $477\,\mathrm{nm}$ by applying the Angström Exponent calculated between $440\,\mathrm{nm}$ and $500\,\mathrm{nm}$ for VIS and to $360\,\mathrm{nm}$ by applying the $380\,\mathrm{nm}$ and $440\,\mathrm{nm}$ Angström exponents.

The time series for all datasets (left) show a large variability throughout the months. No clear seasonality can be observed indicating that the aerosol load is rather not on a constant background level but it is dominated by larger aerosol events as e.g. at the end

of 2017. The comparison between AERONET and BOREAS AOT indicates that the peak values were retrieved fair while the smaller AERONET background can not be confirmed. However, the standard deviation of BOREAS AOT (red shaded area) reveals large variability with specific scenarios in the range of the AERONET background level. The only fair agreement between the time series is not surprising considering the generally large distance between the MAX-DOAS measurements location and the AERONET site (roughly $93\,\mathrm{km}$). Furthermore, the distance of the AERONET station to the airport and the industrial harbour in Hamburg is short indicating that aerosol peak loads can not be considered to match those in Bremen.

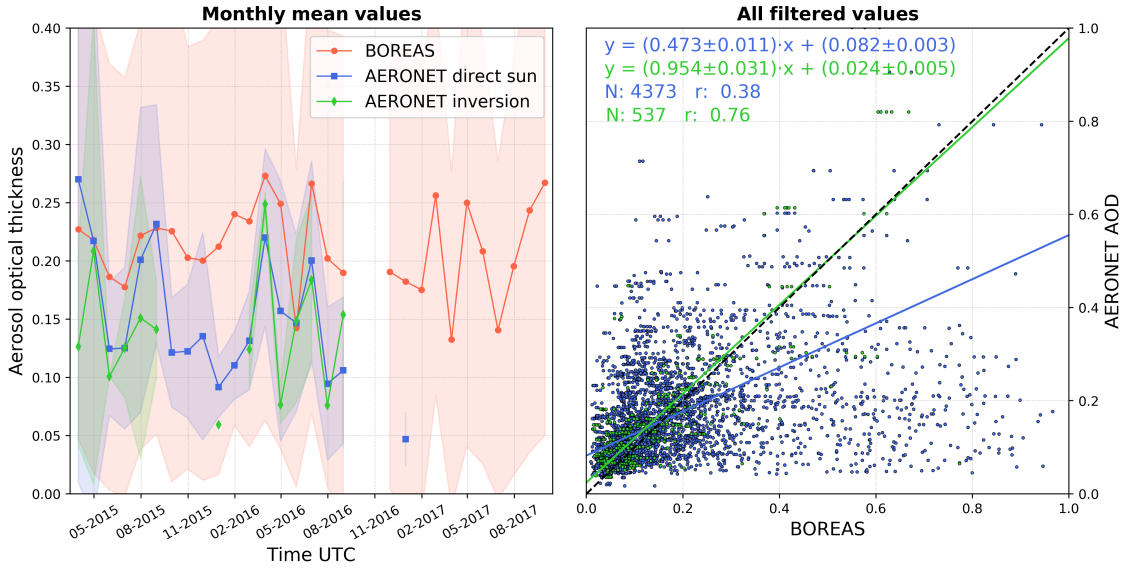


***Fig. 4.50: Left**: Time series of monthly mean BOREAS AOT and AERONET direct sun and inversion AODs. **Right**: Scatter plot including correlation coefficients and regression lines for all filtered values in the visible spectral range.*

Nevertheless, the correlation coefficients in the visible are fair and good for the direct sun and inversion products, respectively. The coefficients for the UV retrieval are slightly smaller but the regression line for the Aeronet direct sun is closer to 1. This is interesting as it is contrary to the findings for the CINDI-2 dataset. The slope of the regression line for the inversion product is slightly better in the visible but still small enough indicating a good agreement in both spectral ranges. The general underestimation of MAX-DOAS profiling AOTs is in agreement with other studies (e.g. CLÉMER et al., 2010; FRIESS et al., 2016). However, both correlation plots show large BOREAS AOTs which are mostly not accompanied with similar values from the AERONET products. Due to this, the correlation is not as good as the bulk of data points indicate. If these AOT were retrieved during a prevalent large aerosol load or a somehow problematic scenario can not be deduced at this point.

The Figures 4.51 and AF.59 show BOREAS surface extinction values and BLUES PM10 concentrations as monthly mean time series for the VIS and UV, respectively.
Both plots show a generally good temporal agreement between BOREAS and in situ PM10 concentrations. The differences are largest in winter 2016/17 with higher values for the in situ site than for BOREAS. On the other hand, in winter 2015/16, the BOREAS results exceed the PM10 concentrations. The comparison with the UV plot shows that the peak differences between BOREAS and BLUES are larger for the winter 2015/16 while it stays similar for 2016/17. For both spectral ranges, the winter surface values are usually higher than the summer values indicating a clear seasonality here, which was not found for the AOT. The correlations with the in situ station are good for both spectral ranges indicating a better agreement than compared to AERONET. Note that the number of data points is much larger for the surface aerosol plots due to the high temporal coverage of the in situ station. Exceptionally high BOREAS values which are not accompanied by ancillary data (as for the AERONET AOT) can not be found. However, deviations from the regression lines outliers can be found for both instruments and fitting windows.

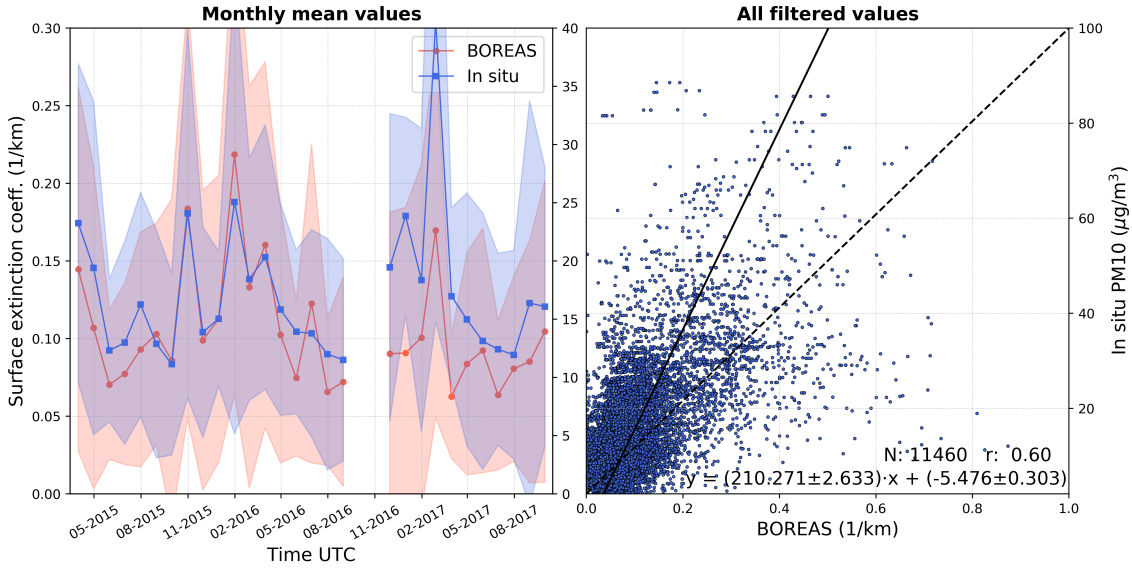


***Fig. 4.51: Left**: Time series of monthly mean BOREAS surface extinction values and in situ PM10 concentrations for the in situ station I_2. **Right**: Scatter plot including correlation coefficients and regression lines for all filtered values in the visible spectral range.*

In order to evaluate the seasonal variability, Figures 4.52 and AF.60 (appendix) depict the diurnal surface aerosol variation for each month and for VIS and UV, respectively. The shaded areas represent the standard deviations and the black vertical lines separate the individual months. Both instruments depict similar diurnal variations with a larger variability for the winter months, when the number of data points is smaller. Nevertheless, the morning hours are nearly always accompanied with the largest surface aerosol load which was also found for specific days within other studies (LI et al., 2010; FRIESS et

al., 2016). This was already expected as the effective light path lengths were found to be small in the morning (Fig. AF.34) and the PBL has not yet evolved (Fig. AF.8). After that initial high morning aerosol load, a decrease can be seen throughout the average days and for all months with a slight increase in the evening for some months. This evening enhancement could be explained by commuter traffic emissions. However, this explanation is not satisfying as these higher values can not be observed for all months. The underlying diurnal cycle can not be seen as well for the winter months as for the summer but a general seasonal variation of the aerosol background level (similar to WANG et al., 2017; MAMALI et al., 2018) can be identified. This can be observed even better for the UV, when the typical aerosol extinctions are higher. Here, January and February show the largest morning values while July and August represent the lowest morning extinctions. The December values are also low which can be due to smaller number of data points but usually also higher temperatures, and therefore, no necessity for large scale heating of resident's houses, meaning less particle emission.

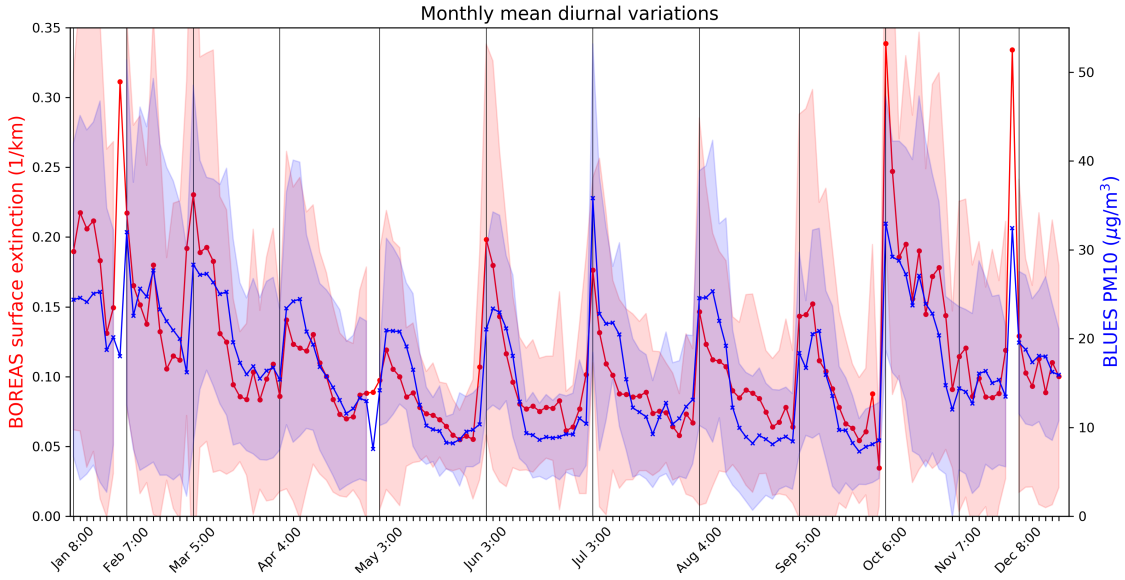


Fig. 4.52: *Comparison of monthly mean diurnal surface extinction and PM10 concentrations for VIS. The shaded areas indicate the standard deviations.*

Figures 4.53 and AF.61 show hourly mean profiles averaged for the full time period in the visible and ultra-violet spectral range, respectively. In addition, BLUES PM10 in situ data is shown as the lowermost row. The red lines depict similarly averaged AOT for the specific profiles with the man daily values as dashed line. It can be seen that the largest aerosol load can be found close to the surface in the mornings for each day. These findings can be confirmed with the in situ dataset which shows a similar surface decrease throughout the day which is accompanied by a lifting of the aerosol most likely due to the rising PBL. On Wednesdays, the largest extinctions can be found near the surface while

on Fridays, higher values are in elevated altitudes. It is interesting to note that the beforementioned detachment from the surface can again be identified. This was also found during the CINDI-2 discussion in Section 4.4.2 (cf. **Test 4**) where it was left as unclear if this is the retrieval response to a box like profile or a real detached layer. However, in the above discussion about near-surface aerosol and the coincident in situ data in the plot below, a good agreement can be found, indicating that the lower aerosol extinction between noon and afternoon at the surface is accurately retrieved. Unfortunately, this does not confirm if the higher altitude extinctions were retrieved well. However, the UV profiles support an elevated layer even more clearly as the visible profiles. Here, the Wednesday morning values are also large but the highest extinction values can be found on Saturday morning. On the other hand, the visual agreement with the in situ dataset is less good, most likely due to the smaller number of profiles in the UV. On the other hand, this could also indicate possible horizontal inhomogeneities which were also found with the onion peeling approach (see Sec. 3.8). The AOT for the UV seems to obey an increasing trend throughout nearly all days while it is more variable in the visible spectral range. Here, large values in the morning and around noon are dominant.

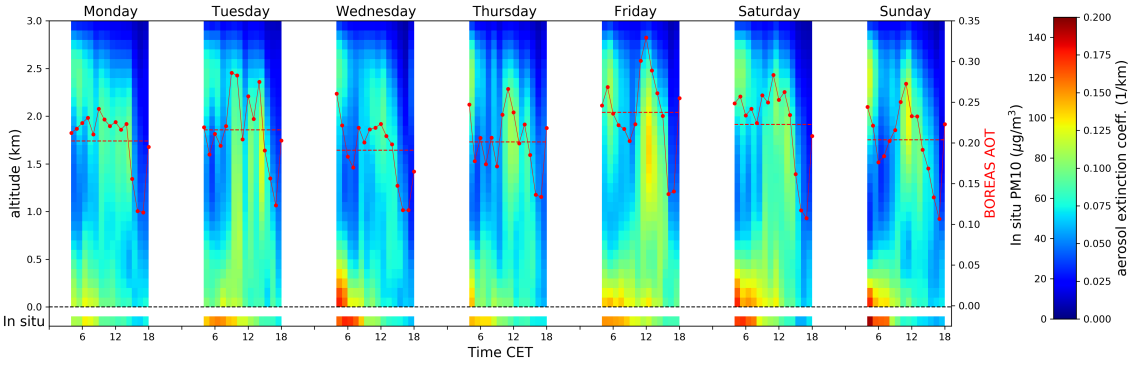


Fig. 4.53: *Average extinction coefficient profiles (colour-coded) for the days of the week retrieved in the visible spectral range. The lowest row indicates the in situ PM10 conc. for BLUES site I_2. The red lines show BOREAS AOT's with the mean level depicted as dashed line.*

The mean diurnal profiles for the four seasons of the year are depicted in the Figures 4.54 and AF.62 for VIS and UV, respectively. Here, the temporal sampling is 15 min instead of 1 hours as above. The numbers in the corners give the numbers of data points used for the averaging.

Again, larger aerosol extinctions can be seen for all seasons in the morning with an uplift of aerosol in the afternoon. However, this elevated layer can not be observed as well for winter and spring as for summer. Especially in winter, large extinction values close to the surface were found on average throughout the whole day (VIS). For the UV, these large winter surface values can also be seen but smaller high altitudes enhancements can be identified, additionally. This difference between VIS and UV supports the statement about horizontal inhomogeneities. The comparison for spring shows a rather box-like afternoon

profile in the UV while it appears slightly elevated for VIS. Interestingly, the summer profiles show a detached layer for both spectral ranges. When looking at the summer mean diurnal profiles in the UV, the impression of a plume-like structure appears. Here, the reader might be reminded that the average onion peeling map also shows dominant point-sources (cf. Fig. 3.30). However, the true nature of this issue can unfortunately not be solved without further analyses and measurements of ancillary instruments.

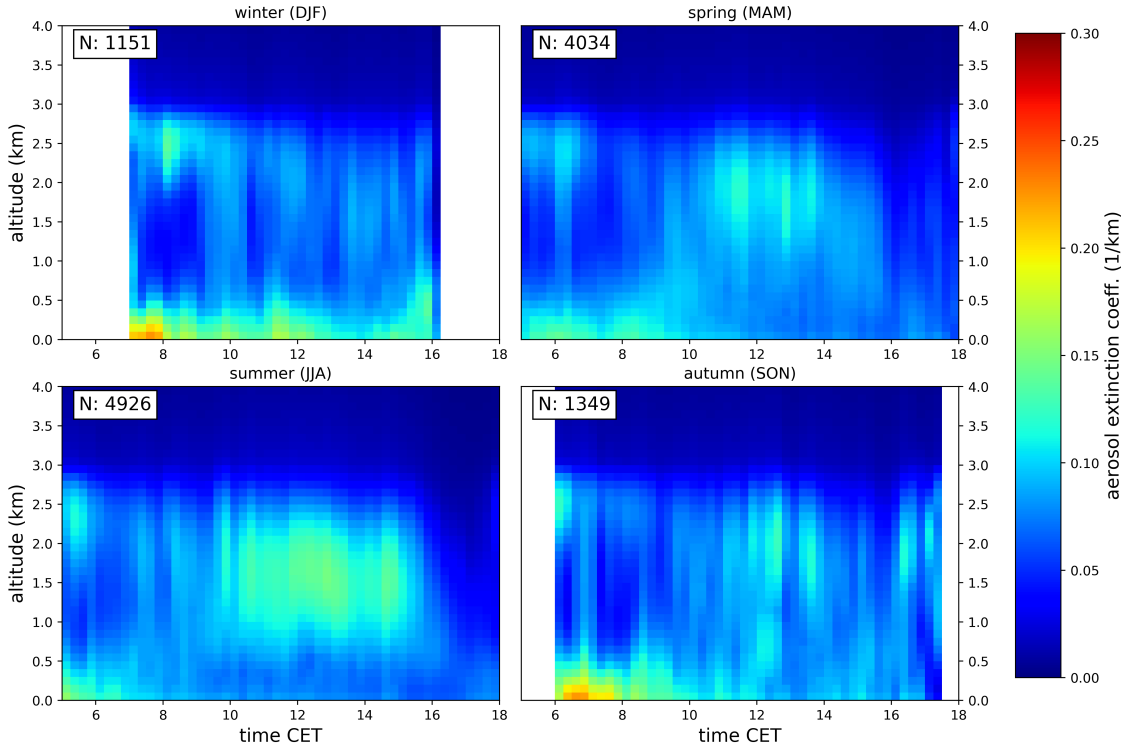


Fig. 4.54: *Average extinction coefficient profiles (colour-coded) for the seasons of the year retrieved in the visible spectral range. The numbers in the corners state the number of data points used for the averaging.*

4.5.2 Retrieval of NO_2 profiles

NO_2 concentration profiles were retrieved from MAX-DOAS data with fit settings introduced in Table 3.3 in the visible and ultra-violet spectral range (VIS, UV1). The Levenberg-Marquardt option was used including a priori pre scaling. The unscaled a priori profile was exponential with a scale height of $0.75\,\mathrm{km}$ and a surface concentration of 9.13×10^{10} $\mathrm{molec/cm^3}$. The a priori variance was set to $50\,\%$ for both fitting windows.
The resulting profiles were filtered based on the following criteria:

1. Profiles were discarded if the maximum iteration number for the underlying aerosol profile was reached.

2. Additionally, not more than 5000 iterations were allowed for the Levenberg-Marquardt algorithm.
3. The RMS of dSOT differences of the underlying aerosol retrieval was not allowed to be larger than 1000.
4. Underlying aerosol profiles with AOTs larger than 1.0 were discarded as these led to poor convergence, e.g. due to difficult weather conditions (like fog).

Unless stated otherwise in this section, all UV quantities were retrieved for UV1 data only.

Figures 4.55 and AF.63 (appendix) show NO_2 vertical column densities retrieved for the c direction (east-north-east) for VIS and UV, respectively. The corrected geometric vertical column densities VC_{geom} introduced in Section 3.5.1 and discussed in Section 3.7.3 are shown (black) in addition to BOREAS VC (red), CAMS tropospheric columns (green) and BLUES in situ concentrations (blue), as monthly mean values (left). The temporal behaviour of all curves is similar with large NO_2 columns in winter and smaller values in summer, which was also found by other studies (KRAMER et al., 2008; MA et al., 2013; MENDOLIA et al., 2013). The agreement between MAX-DOAS VC and profiling integrated

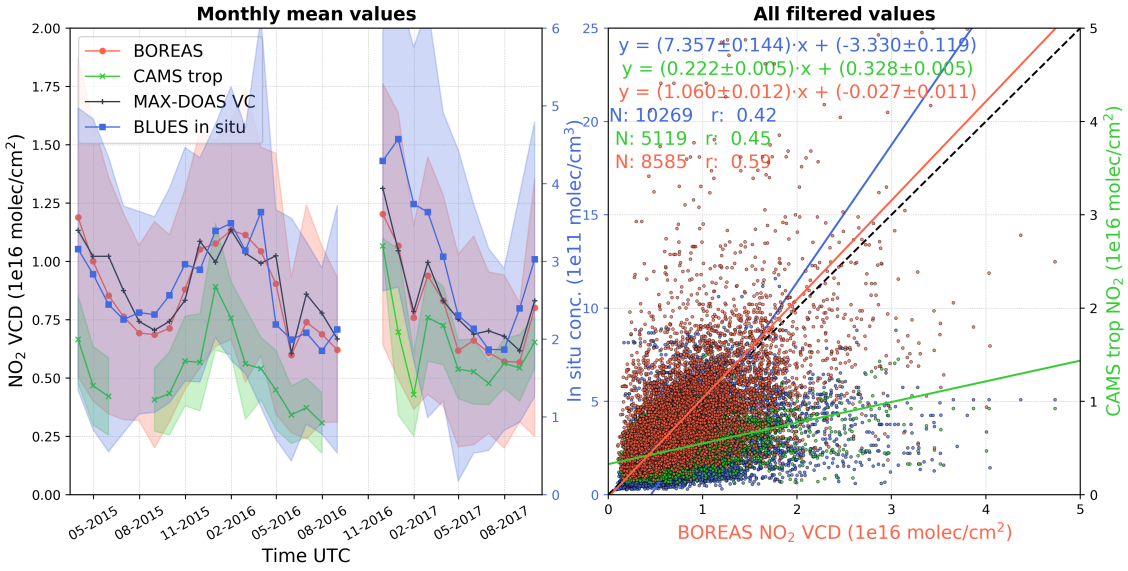


***Fig. 4.55: Left**: Time series of monthly mean BOREAS, MAX-DOAS (geom. corrected VC), and CAMS VC as well as BLUES in situ NO_2 concentrations. **Right**: Scatter plot including correlation coefficients and regression lines for all filtered values in the visible spectral range.*

values is good but smaller deviations can be identified showing a large variability in both datasets. It can be assumed that problematic scenarios which were not filtered for the VC_{geom} calculation were also left unfiltered by the BOREAS algorithm. In these cases, the retrieval response is unpredictable which might be the reason for differences between both datasets. Even though the BLUES surface concentrations agree well when considering the

overall temporal curve shape, an interesting feature can be observed for February 2017. Here, all datasets except BLUES show a strong decrease which is not covered by the in situ site. This is interesting as it might indicate that long-distance NO_2 transport might have happened which is unmeasurable by near-surface stations (as BLUES) but can be seen with BOREAS. Additionally, CAMS tropospheric columns find this as well but show a general strong underestimation of the total column which was also found by BLECHSCHMIDT et al., 2017. The correlations are fair with the highest agreement for MAX-DOAS and BOREAS results indicating the consistency of the different analysis techniques. For these scatter points, the regression line is close to one while it is smaller than one for CAMS indicating the before-mentioned underestimation by the model. The correlation with the in situ site is the smallest which can be attributed to the difference in the compared quantities and outliers for both datasets.

In Figures 4.56 and AF.64, surface concentrations of BOREAS (red), the BLUES in situ site (blue) and the different onion peeling results (black to green) are depicted. For all dataset, high values can be seen in winter and low values in summer confirming the results of the vertical column comparison above. This seasonal behaviour for near-surface NO_2 concentrations was similarly found by other studies (e.g. MENDOLIA et al., 2013; VLEMMIX et al., 2015). Furthermore, all curves except for the BLUES in situ dataset show the before mentioned decrease in February 2017. This somehow disproves the long range transport as the retrieved quantities are considered as near-surface even though the 1° LOS, used for the onion peeling approach, also includes additional altitude information. Since the

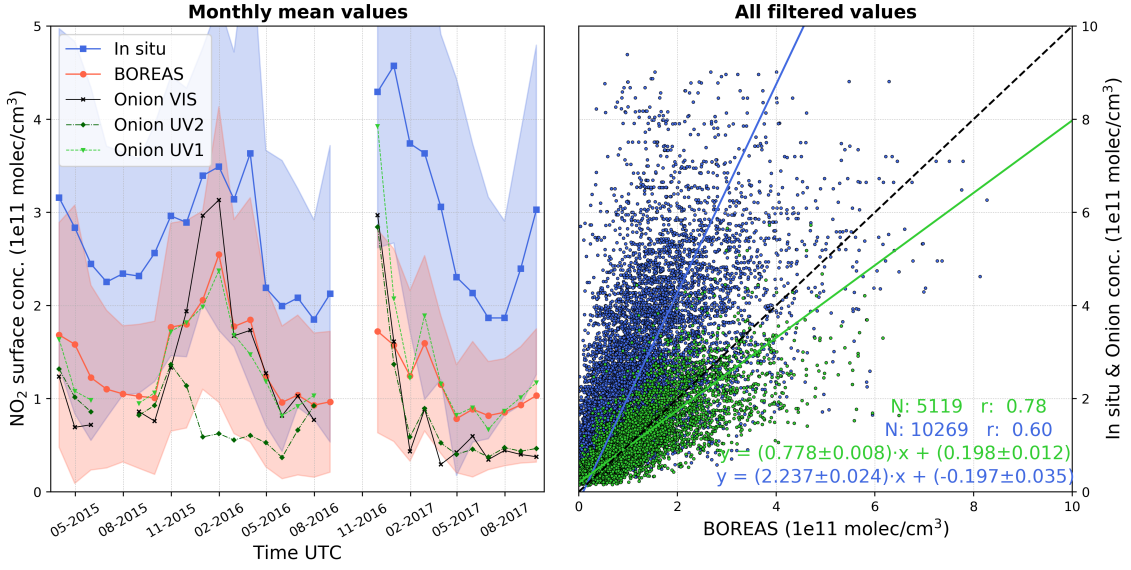


***Fig. 4.56: Left**: Time series of monthly mean BOREAS NO_2 and BLUES in situ surface concentrations (in situ station I_2) in addition to the NO_2 concentrations retrieved with the onion peeling approach. **Right**: Scatter plot including correlation coefficients and regression lines for all filtered values in the visible spectral range.*

in situ station does not show this decrease, a horizontal inhomogeneity might explain the results meaning that either the prevalent wind direction blew NO_2 away from the in situ site into the field of view of the telescope or a close temporary source was responsible (when neglecting instrumental issues). However, it is questionable that a strong source emitted only in this specific time period. In general, the in situ data is underestimated by BOREAS and the onion peeling results for the full period except for one peak value in the UV. It is interesting to see that the UV1 onion peeling approach agrees best with BOREAS VIS results even though they were retrieved in different spectral ranges. This gives an indication of the effective horizontal distance covered by MAX-DOAS profiles which include the long-range horizontal light-paths of the lower elevation angles but also the horizontal light paths for the higher LOS, which are expected to have shorter horizontal overground lengths. Therefore, the profile representing effective light path can be assumed as being a weighted mean light path which is obviously closer to the UV1 path lengths (usually smaller than $10\,\mathrm{km}$, cf. Fig. AF.34).
The correlation with the in situ site is good and it is even better with the UV1 onion peeling concentration showing again the consistency of both approaches even though they work in completely different ways. The correlation with onion peeling is slightly better for the UV but a smaller slope can be observed (VIS: 0.778, UV: 0.548). This is interesting as the matching monthly mean curves agree well except for the winter 2015 for UV. Here, BOREAS retrieves higher concentrations which come close to the in situ dataset. However, the general number of data points can be assumed to be much lower for winter indicating that outliers might have affected on the results.

Figures 4.57 and AF.65 show the diurnal surface concentration variation for each month and for VIS and UV, respectively. This plot is in agreement with the similar aerosol plots (Fig. 4.52 and AF.60) and depicts higher values in the morning and a decreasing trend throughout the day. In the evening of some months, a slight increase can be observed. The diurnal curve shape can be again explained by low PBL's in the morning accompanied with large emissions by commuter traffic. The evolving PBL and less emissions around noon lead to the decreasing trend. The slight increase in the afternoon can be again attributed to traffic emissions. The background level is highest for the winter months and small in summer. The comparison of all datasets shows a good agreement between BOREAS and the onion peeling results but higher values for the in situ site.

The Figures 4.58 and AF.66 depict the average NO_2 concentration profiles for the individual days of the week and for VIS and UV, respectively. In addition, the matching averaged VC curves are shown in red, with the dashed red line representing the diurnal mean level. The comparison with the similar aerosol plots (Fig. 4.53 and AF.61) show that NO_2 is mostly concentrated in lower altitudes ($< 1\,\mathrm{km}$) whereas aerosols were also

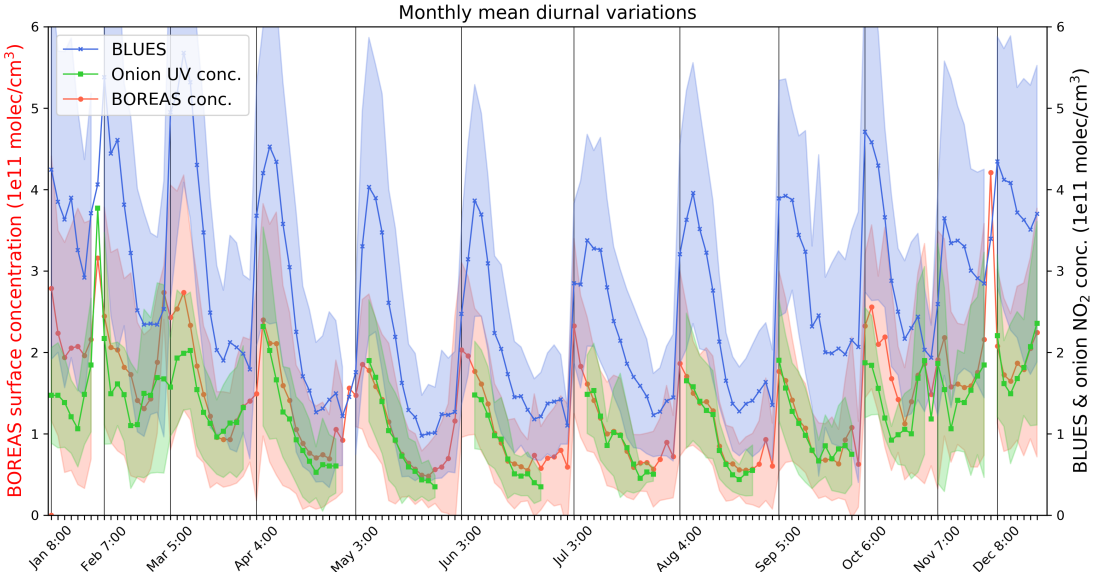


Fig. 4.57: *Comparison of monthly mean diurnal NO_2 surface concentrations retrieved with BOREAS and the onion peeling approach (for UV1) as well as data from the BLUES in situ site I_2 (VIS). The shaded areas indicate the standard deviations.*

found at higher altitudes (up to $3\,\mathrm{km}$). This shows that aerosol undergoes a more efficient mixing due to the longer residence time while NO_2 is mainly located close to its source (horizontally and vertically) and reservoir species can be considered as having a small impact (when present). Consequently, NO_2 is mainly found in the mornings of each day with the smallest concentrations on Sunday, when commuter traffic emission can be neglected. Especially for the working days, an evolving PBL is represented by increasing concentrations in altitudes up to $1.5\,\mathrm{km}$. This is more pronounced for the VIS results than for UV indicating that the corresponding emitters are not close to the measurement site and that the effective light paths do not fully cover the distance to the power plants and residence areas on the other side of Bremen's Bürgerpark (cf. map of Bremen, Fig. 3.2).

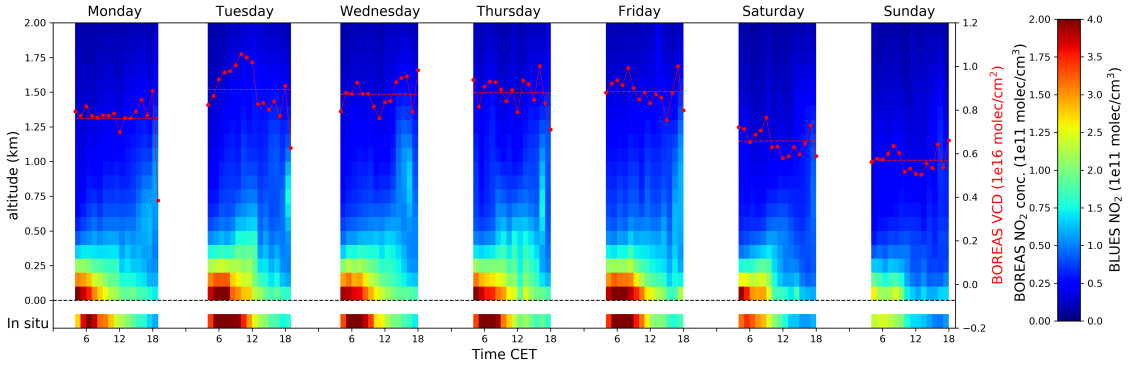


Fig. 4.58: *Average NO_2 concentration profiles (colour-coded) for the days of the week retrieved in the visible spectral range. The lowest rows indicate the in situ concentrations for BLUES site I_2 with a second colour-scale on the right-hand side. The red lines show BOREAS VCDs with the mean level depicted as dashed line.*

The seasonally averaged profiles in the Figures 4.59 and AF.67 confirm the results above. NO_2 is mainly concentrated close to the surface (similarly found by WANG et al., 2017; CHAN et al., 2018) but higher altitude concentrations are possible, when the PBL evolves (winter). The other seasons, however, do not show larger values in higher altitudes which might be due to generally smaller emissions coupled with a larger vertical space by a higher upper PBL boundary and the enhanced loss via reaction with OH (cf. Eq. 2.70), in seasons when more daylight is present. The VCD is usually larger in the mornings and in the late afternoon than the mean daily NO_2 VCD level. Interestingly, this mean level is the highest for Tuesdays (VIS) at it was already found in Section 3.7.3. This levels stays similarly high until Friday and decreases over the weekend. For the UV, the highest value can be observed on Thursday and again is the lowest over the weekend.

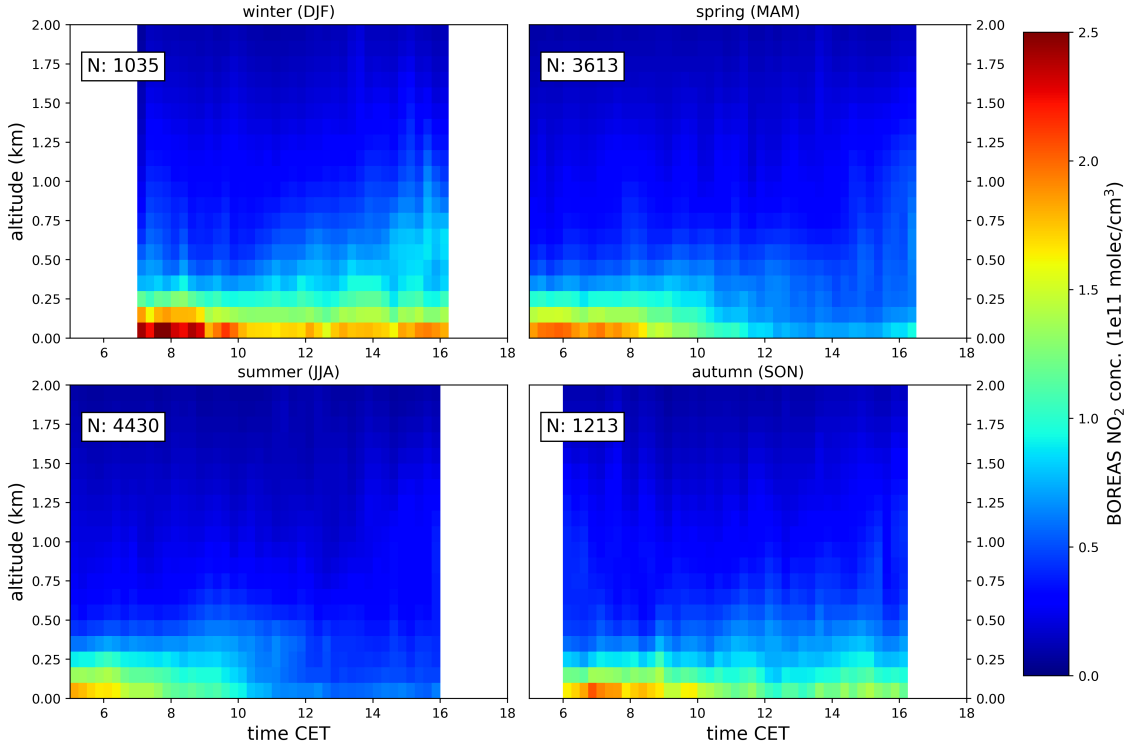


Fig. 4.59: *Average NO_2 concentration profiles (colour-coded) for the seasons of the year retrieved in the visible spectral range. The numbers in the corners state the number of data points used for the averaging.*

4.5.3 Retrieval of HCHO profiles

The HCHO Levenberg-Marquardt retrieval was started with underlying aerosol profiles already retrieved with the UV1 fitting window and which were extrapolated from $360\,\mathrm{nm}$ to $343\,\mathrm{nm}$ by using the mean Angström-Exponent of 1.365. All settings and filter options are similar to the NO_2 retrieval (Sec. 4.5.2) but the a priori scale height was chosen to be $1\,\mathrm{km}$ with a surface concentration of $8 \times 10^{10}\,\mathrm{molec/cm^3}$.

Figure 4.60 shows HCHO monthly mean values of surface concentrations and vertical column densities, left and right, respectively. In addition, temperature and BLUES NO_2 in situ concentrations are plotted in both sub-figures. Furthermore, the MAX-DOAS geometric VC is shown on the right-hand side of the figure. Note that the NO_2 curves were scaled by 4×10^{11} instead of 1×10^{11} (for HCHO) to make the comparison easier.
The largest HCHO values can be seen in winter 2015/16. This is surprising as biogenic emissions can be excluded in winter and high concentration values are most likely due to anthropogenic emissions. The corresponding high standard deviations indicate that there is a large variability, meaning that these results in winter are questionable. The winter 2016/17 shows some of the lowest values indicating a multi-variable dependency on different factors. The maximum peak agrees well with high NO_2 concentrations while the in situ and HCHO curves for the year 2017 seem to be anti-correlated. However, in this

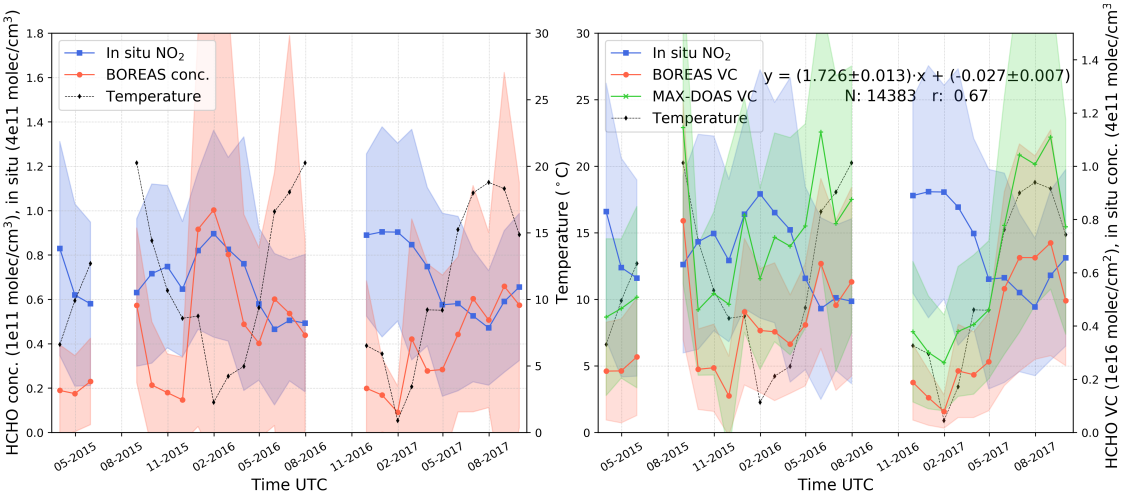


***Fig. 4.60: Left**: Time series of monthly mean BOREAS HCHO and BLUES in situ NO_2 concentrations, in addition to the mean ambient temperature curve. **Right**: Time series of monthly mean BOREAS and MAX-DOAS HCHO VC, BLUES in situ NO_2 concentrations, in addition to the mean ambient temperature curve. The regression and correlation parameters compare BOREAS and MAX-DOAS VC.*

time period, the temperature fits quite well to the HCHO data with low values in winter and high concentrations in summer. On the other hand, the temperature indicates an anti-correlation with HCHO for the years 2015/16. The VC values on the right side show a similar curve shape as the surface concentrations but the maximum peak in winter 2015/16 does not lead to high VC. However, the geometric vertical column densities of the 30° LOS agree well with BOREAS results which are slightly underestimated. The regression line between BOREAS and MAX-DOAS results has a slope of 1.73 with a correlation coefficient of 0.67 showing a good agreement, similar to the NO_2 results. Except for the high 2015/16 winter values, the findings show a seasonal variability similar to other studies (VLEMMIX et al., 2015; WANG et al., 2017; TIAN et al., 2018; KHAN et al., 2018).

The double-dependence of HCHO with NO_2 and the temperature can also be seen in Figure 4.61. Here, monthly mean diurnal HCHO and NO_2 (in situ) surface concentrations are shown together with similar averaged temperature curves. In some months, the agreement between HCHO and NO_2 seems to be excellent (March, April) while there is a total disagreement in summer.

In general, the spring months are accompanied by large HCHO values in the morning and afternoon with a minimum around noon. The summer months show more variations with a high background level during the full mean day. The autumn curves look similar to the spring months while the winter months are extremely noisy, which prevents a discussion of the diurnal shape. In order to assess the HCHO seasonal and diurnal variation, polynomials with the order of 4 were fitted to all three datasets to show the background level of the corresponding curves. As expected, the temperature curve has a maximum in summer and a minimum in winter. NO_2 shows the opposite behaviour. The summer is accompanied by a minimum while the winter shows the largest NO_2 background level. As already discussed in Section 4.5.2 and 3.7, this is due to the enhanced emissions by fossil fuel combustion processes (heating of residential houses, increasing energy consumption when less daylight is available), increasing traffic emissions and a generally larger lifetime in winter. The HCHO is expected to show a more or less approximated superposition of enhanced emissions by combustion processes and temperature. While the first dependency was explained in Section 2.6, the second one is attributed to a larger biogenic activity when the temperatures are higher leading to an increase of NMVOCs which are a source

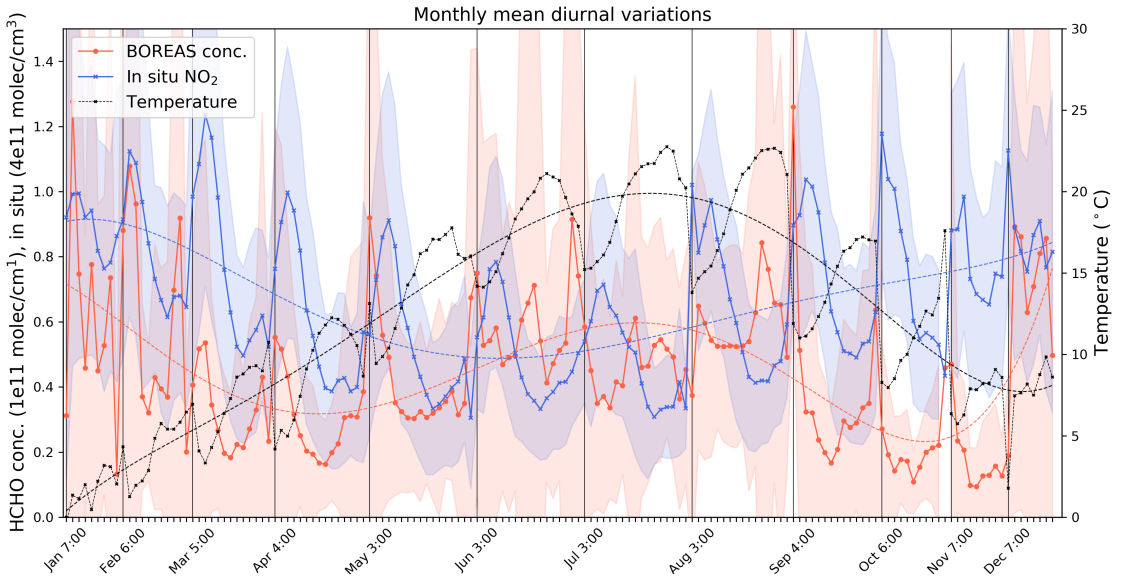


Fig. 4.61: *Comparison of monthly mean diurnal* HCHO *surface concentrations retrieved with BOREAS and* NO_2 *concentration values from the BLUES in situ site* I_2 *in addition to the ambient temperature curve. The shaded areas indicate the standard deviations while the dashed lines depict polynomials of the order 4 fitted to all three datasets.*

for HCHO. This temperature dependency was also found by PINARDI et al. [2013] and it might be responsible for the local maximum in the HCHO background curve in summer. In addition to that, an increase of HCHO background levels can partially also be attributed to the oxidation of methane when more OH radicals are present in summer.

Figures 4.62 and 4.63 show the weekday diurnal evolution of HCHO profiles as well as the seasonal mean diurnal HCHO concentration profiles. In addition, ECMWF PBL's are depicted as black dotted lines while the HCHO VC is given as red lines in the first graph. Both figures show the largest values close to the surface. Furthermore, the HCHO concentration is high in the morning and afternoon while there are temporal surface-value minima around noon. Interestingly, HCHO rises higher in altitude than the NO_2 does (cf. Fig. 4.58) which might be attributed to the longer lifetime, a generally larger background level due to the oxidation of methane and large biogenic emissions. The latter point is supported by the fact that the summer profiles go up to the PBL while the other seasons' HCHO is located close to the surface. However, it should be noted that the HCHO in higher altitudes shows only smaller concentrations of $0.2 \times 10^{11}\,\mathrm{molec/cm^3}$ to $0.4 \times 10^{11}\,\mathrm{molec/cm^3}$ and that the main HCHO load is close to the surface with average concentrations up to $1 \times 10^{11}\,\mathrm{molec/cm^3}$.
A higher rising HCHO (compared to NO_2) was also found by WANG et al. [2017]. Additionally, their findings support that a weekly cycle can not be expected.

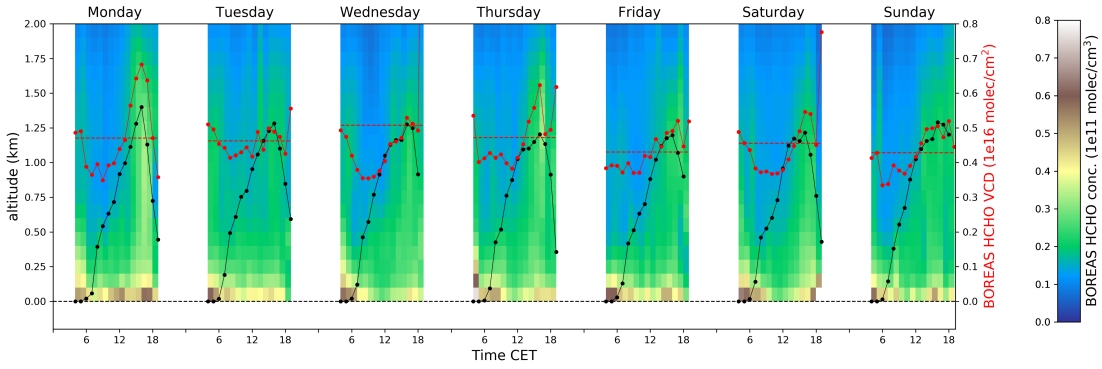


Fig. 4.62: *Average* HCHO *concentration profiles (colour-coded) for the days of the week. The red lines show BOREAS* HCHO *VCD's with the mean level depicted as dashed line while the black lines represent the PBL.*

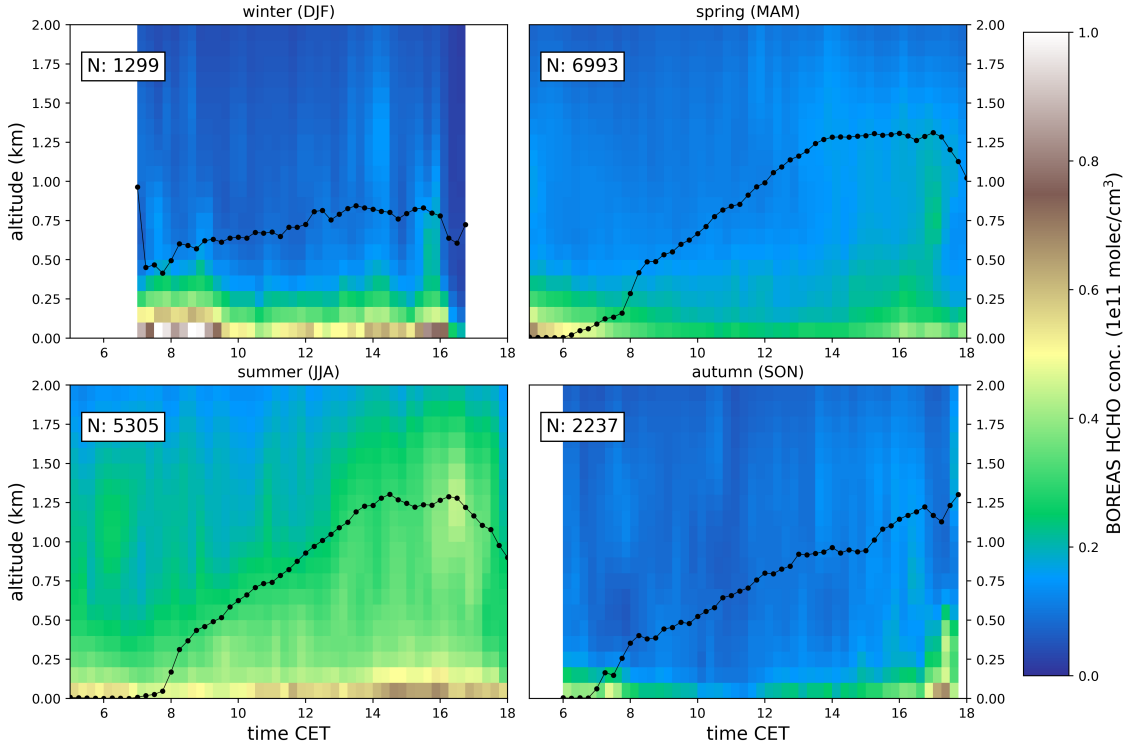


Fig. 4.63: *Average* HCHO *concentration profiles (colour-coded) for the seasons of the year. The numbers in the corners state the number of data points used for the averaging. The black lines represent the PBL.*

CHAPTER

V

CONCLUSIONS AND OUTLOOK

This thesis focusses on the analyses of spatial and temporal inhomogeneities in the distribution of tropospheric aerosols and trace gases. A new Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) profiling algorithm BOREAS (Bremen Optimal estimation REtrieval for Aerosols and trace gaseS) was developed in order to determine the vertical distribution of aerosols and trace gases. The performance and accuracy of this novel algorithm is tested with the help on synthetic data as well as with ancillary measurements on the example of data from the CINDI-2 campaign in Cabauw (the Netherlands), in 2016 (Sec. 4.4.1).
In addition, different retrieval methods, options, and settings are thoroughly analysed and discussed in order to enable the retrieval of optimal vertical profiles.
A comprehensive study of trace gases and aerosols in the area of Bremen is then presented, which utilizes different analysis techniques.
In a first preliminary discussion, vertical column densities of NO_2 and $HCHO$ are analysed in order to get a first impression on their distribution around Bremen (Sec. 3.7.3).
In a subsequent discussion, horizontal inhomogeneities of NO_2 are presented and discussed with the help of an onion peeling approach, exploiting NO_2 differential slant column densities (dSCD) calculated in three different spectral windows (Sec. 3.8).
The discussion of atmospheric constituents in Bremen utilizes the analysis of BOREAS' profiling results for NO_2, $HCHO$ and aerosols, which support the preceding findings, but gave also enhanced insights on the spatial and temporal distribution of important pollutants in the area of Bremen (Sec. 4.5).

Development and validation of BOREAS

BOREAS was developed in order to improve the results from IUP Bremen's previous MAX-DOAS profiling algorithm BREAM (WITTROCK, 2006), which used the Optimal estimation (OE) technique for the trace gas retrieval but a simplified trial-and-error approach for the calculations of aerosol extinction coefficient profiles.
For BOREAS, a novel aerosol retrieval method was developed, which uses the minimization of the difference between measured and modelled differential slant optical thickness's (dSOT) at wavelengths around an O_4 absorption band. The minimization is achieved by an iterative Newton-Gauss inversion method employing the Tikhonov regularization technique in order to constrain the solution. This inversion can be run with different retrieval

modes, each having a unique optimal regularization weighting between measurement and a priori constrains showing the complexity of MAX-DOAS profile retrievals.

The optimal aerosol retrieval settings are determined, analysed and discussed using exampled from different tests. The optimal number of retrieval altitudes and lines of sights (LOS) were investigated. Here, a strong dependence on the actual aerosol profile is found. Less LOS with smaller elevation angles or a too coarse grid deteriorate the results. Further adding of low elevation LOS is also problematic, when the telescope points close to the surface. This is due to either inaccuracies of the radiative transfer model (RTM), or, the impact of the spectral surface reflectance. Nevertheless, the omission of LOS close to the surface is not recommended as the profiling results lose information and have pour accuracy. Different a priori modes were introduced and tested. Iterations from a *zero* profile is used for the retrieval of a preliminary aerosol optical thickness (AOT) in order to scale the a priori profile for having a more accurate first guess. This a priori pre-scaling is found to perform exceptionally well, when the initial a priori profile was far away from the true conditions.
The main retrieval was tested with the *prev* and *apri*-mode (iteration from the previous and the a priori profile). While the first mode iterates from the previous retrieval result, the latter mode leaves the a priori profile unchanged during the iteration scheme. Both modes have advantages and disadvantages, but retrieve similar or sometimes identical profiles, when the underlying retrieval problem was simple (meaning that the a priori profile was already close to the true profile). The *prev*-mode is found to be better for elevated layers and profile shapes which differ strongly from the a priori shape while the *apri*-mode performs well for profiles with shapes similar to the a priori.

The trace gas retrieval was tested with the commonly used optimal estimation (OE) method, the Levenberg-Marquardt algorithm (LM), and an attempt of an RMS based OE optimization technique. For the latter option, local minima in the RMS difference of measured and simulated dSCD are analysed by changing a regularization factor g iteratively. The g-factor, which corresponds to a certain minimum in RMS, is assumed to be a better regularization than the normal one $(g = 1)$.
Comparing the three methods, the LM option in the logarithmic space leads to the best results because the retrieval of elevated layers and profile shapes, which are far away from the a priori profile, is found to be best. In order to minimize the uncertainty due to inappropriate a priori profiles, pre-scaling is additionally applied by using the $30°$ dSCD value as a scaling factor of the a priori profile. Even though the other retrieval options are found to perform well and sometimes more or less identically, the LM option in the logarithmic space is used for the further investigations due to its adaptability to various trace gas conditions.

A validation study using measurements made during the CINDI-2 campaign is presented and high correlations of BOREAS' retrieval results in the visible spectral range (VIS) with data from other instruments are found. Larger deviations for the aerosol results are possible, when clouds are present due to the strong change in light paths. BOREAS' AOT shows a general underestimation compared to Cimel sun-photometer data from an AERONET station, which is assumed to be due to: (1) different air masses probed, (2) aerosol loads in altitudes, which are not covered by the retrieval or (3) because of an unknown issue, which might be accounted for by using an O_4 scaling factor (see WAGNER et al., 2018). Nevertheless, the correlation is high (0.81).
The agreement of BOREAS' near surface extinctions with PM10 concentrations from the NAQMN in situ station and Ceilometer data is good with correlations of 0.80 and 0.75, respectively. The diurnal shape of BOREAS' surface extinctions is found to be similar to the other instruments showing the high quality of the retrieved profiles.

NO_2 VC were validated by comparison with Pandora direct sun tropospheric columns and with integrated NO_2-Lidar values. The correlation is 0.83 and 0.79, respectively. Similar diurnal variations were retrieved as by the other instruments but deviations are possible, when the underlying aerosol profile was retrieved insufficiently (e.g. when clouds were present).
BOREAS' NO_2 surface concentrations are well correlated with in situ measurements, long path DOAS results and NO_2-Lidar near surface values. The respective correlation coefficients are 0.83 and 0.86 for ICAD/CAPS and NAQMN in situ values, 0.83 for the $12.7\,\mathrm{m}$ LP DOAS concentrations and 0.84 for the NO_2-Lidar. Furthermore, the diurnal curve shape agrees well for the noon and the afternoon. However, in the mornings of two of the five investigated days, larger differences are found. These deviations between the different instruments could be observed when the NO_2 concentration was located close to the surface. Here, BOREAS limited vertical resolution is assumed to be too coarse in order to resolve these small vertical trace gas extents. This is in agreement with the LP DOAS and ICAD/CAPD results measured at different altitudes in the lowest $200\,\mathrm{m}$. Both datasets show a large variability of near-surface NO_2 in the morning hours.

An extension of the CINDI-2 validation study published in BÖSCH et al. [2018] showed that AERONET aerosol information should be used if it is provided close in time and space to the MAX-DOAS measurement. Average single scattering albedos and asymmetry factors could otherwise be used. However, the Henyey-Greenstein approximation for the aerosol phase function is found to be a poor representation of the angular aerosol scattering probability when the asymmetry factor is inaccurate.

Aerosol and trace gas distribution in the area of Bremen

An introductory analysis of NO_2 VC for three different azimuthal directions indicated that temporal and spatial differences are observed. A clear seasonal as well as weekday cycle is found with larger values in winter and on the weekdays. The summer VC are generally smaller with the lowest NO_2 VC on Saturday and Sunday.
The diurnal variation of NO_2 shows clear morning and afternoon peaks which can be explained by traffic emissions. This is in agreement with BLUES in situ data (Bremen air quality network) and with the number of vehicles leaving and entering the city of Bremen (Bundesanstalt für Straßenwesen), which show both similar maxima supporting the argument of traffic being responsible for this diurnal curve shapes.
Generally, the two westerly viewing directions show similar results, the southerly direction, however, is found to have a clear noon peak in NO_2 VC. This can be explained by the telescopes pointing towards the sun and indicating that aerosols might have a negative impact on the retrievals.
In order to evaluate horizontal inhomogeneities, the onion peeling method was applied to MAX-DOAS measurements at 1° elevation angle. Three case studies revealed that the NO_2 distribution in the area of Bremen is not as homogeneous as expected. Larger NO_2 concentrations are found in the distance of power plants and industrial complexes while in the in-between area smaller concentrations are observed. Furthermore, days with low wind speeds led to an aggregation of NO_2 close to industrial areas. An in situ station close to these industrial complexes in the west of Bremen shows a good correlation (0.68) with the UV1 NO_2 concentrations while it is slightly lower for UV2 (0.53) and much lower for VIS (0.38). However, the onion peeling case study shows that the distance of certain high NO_2 concentrations can not always be assigned to specific emitters. Furthermore, large in situ values are not necessarily observable with the onion peeling approach when the emitter is close to the surface.
Nevertheless, the concentration values retrieved with the onion peeling method were used to create a map of the NO_2 distribution in Bremen. This approach applied wind vectors calculated with the wind speed and direction (DWD site near the airport) together with NO_2 lifetimes from model-simulations (TM4-ECPL) to create areas of homogeneous NO_2 values from the onion peeling concentrations. Due to the small number of azimuthal viewing directions and the prevalent wind direction (south-westerly), the approach leads to a map of NO_2 trajectories which point to certain emitters in Bremen. What was originally developed to create a map of the NO_2 distribution in Bremen led to a back-tracing approach for dominant emitters. Additionally, because these trajectories point to the exact position of power plants and industrial complexes, the effective light paths of the onion peeling approach were successfully validated.
However, for specific cases, neither the distance nor the actual concentration values are

optimal as they suffer from wind transport of the trace gas, from several emitters close to each other mixing signals indistinguishably, and from aerosols, which have to be accounted for when calculating the applied correction factors.

Aerosol extinction and NO_2 concentration profiles were retrieved for the VIS and UV1 fitting windows from data of the westerly azimuthal direction c. Unfortunately, only PM10 in situ data is availabel as aerosol information from the BLUES network. AERONET data from the Hamburg site is available, but the large distance between the cities ($\sim 93\,\mathrm{km}$) prevents proper comparisons. Nevertheless, the correlation coefficients of AERONET and BOREAS VIS AOT are found to be 0.38 and 0.76 for the direct sun and the inversion AOT, respectively. The agreement with in situ data and BOREAS' surface extinction is better with a correlation coefficient of 0.6. Here, the comparison of monthly mean values shows for all datasets large values in winter and a minimum in summer. This is explained by an increase in traffic, residential heating and energy production by power plants in winter, together with a usually smaller PBL and a longer lifetime in winter.
The diurnal variation of extinction values agrees well with the curves of NO_2 VC. Large morning values and a decrease throughout the day can be observed. In the evening, a slight increase can be found which might be attributed to enhanced traffic.
The main aerosol load is located close to the surface in the morning but it rises sometimes similar to the developing PBL. Depending on the season, the diurnal profile comparison indicates that the aerosol layer might have been detached from the surface in summer but stays close to the surface in winter. If this detachment is a real effect or an artefact of the profile algorithm can not be answered.

NO_2 VC from BOREAS agrees well with MAX-DOAS VC calculated with the geometric approximation and the BLUES in situ data. The seasonal cycle of VC is similar to the surface concentration values but a larger variability could be observed. BOREAS VIS results show a high correlation of 0.60 with the in situ data and an even higher value of 0.78 when compared to the onion peeling result of the UV1 fitting window. This shows the consistency of the different analysis techniques which all led to more or less similar results.
The diurnal variation of NO_2 from BOREAS shows again large values in the morning and a decrease throughout the day. The profiling results show that NO_2 is mainly located close to the surface but mixing to higher altitudes in the afternoon is possible. This is found for the winter season when the photochemical loss by OH radicals and O_3 is expected to be smaller. The previously found weekday cycle in NO_2 VC is supported by the profiling results which show also a decrease on the weekend and the largest NO_2 surface concentrations during the week.

Additionally, HCHO profiles were retrieved but the validation is not as easy as for NO_2 because of a lack in validation data. The VC is found to be large in the summer months and smaller in winter. The temporal behaviour of the profiling VC is similar to the MAX-DOAS VC but an underestimation of BOREAS could be observed. The reason for that is unclear.
The comparison of HCHO surface concentrations shows an increase in summer 2015 and 2017 but an additional maximum in winter 2016. This is interesting as it indicates that two strong emission sources can be distinguished in Bremen. First, biogenic emissions are expected in summer leading to the maximum in VC and surface concentration. Second, anthropogenic emissions were found, which are attributed to be responsible for the winter peaks. Furthermore, the diurnal variation of hourly mean values for each month shows a clear double-peak in the HCHO background concentration. Corresponding temperature and NO_2 curves support the argument of biogenic and anthropogenic emissions as their individual maxima correspond well with HCHO's double peak background. The diurnal variation of vertical profiles shows increased near-surface values in the morning and rising features similar to the devolving PBL. This is found especially in summer whereas the others seasons seem to be dominated by near-surface HCHO.
Interestingly, no clear weekday cycle can be observed indicating that anthropogenic emissions are only a smaller factor compared to the biogenic impact. Furthermore, VC values for different viewing directions are found to agree well with each other which supports the argument of a homogeneously distributed HCHO in the area of Bremen.

Outlook

Several interesting features and issues are found in this thesis which require further study. The correction of vertical column densities suffers strongly from imperfect aerosol assumptions. Especially when pointing towards the sun, outliers could be observed which were due to inaccurate aerosol information within the calculation of correction factors for the geometric air mass factor (AMF). Since the alternative AMF approach from the QA4ECV project also faced this issue, more research is suggested in order to improve on the conversion to VC.
The onion peeling approach appeared more or less unreliable after first considerations but the comparison with validation measurements, as well as the back-trajectory maps, showed the potential of this method. It would be interesting to create such a map with more azimuthal viewing directions in order to evaluate, whether a complete map of average NO_2 distributions can be created. On the other hand, the applied correction factor should be analysed in more detail because it has a strong impact on the effective light paths and the concentrations.
The profiling results could be improved by adding the other azimuthal directions and,

even more importantly, the "in-between" fitting window UV2. This is interesting as the onion peeling results revealed that sometimes, large concentrations were found here but the corresponding annual variation curve showed missing features compared to VIS and UV1. Further research is necessary to evaluate the performance of this fitting window. Additionally, NO_2 dSCD retrieved with wavelengths larger than $500\,\mathrm{nm}$ could be calculated as well. Then, 3D distribution maps could be created by using profile results from different azimuthal directions and the four mentioned fitting windows.
Finally, an improvement of the assumptions on aerosols should be discussed as the insufficient coverage of AERONET data shows the need for other approximations, and/or the retrieval of necessary parameters from MAX-DOAS data. For the latter point, the reader is referred to the upcoming paper of Peters et al. (2019), who briefly discuss the possibility of retrieving aerosol information from imaging DOAS measurements.

CHAPTER

A

APPENDIX

A.1 Additional material to Chapter 3

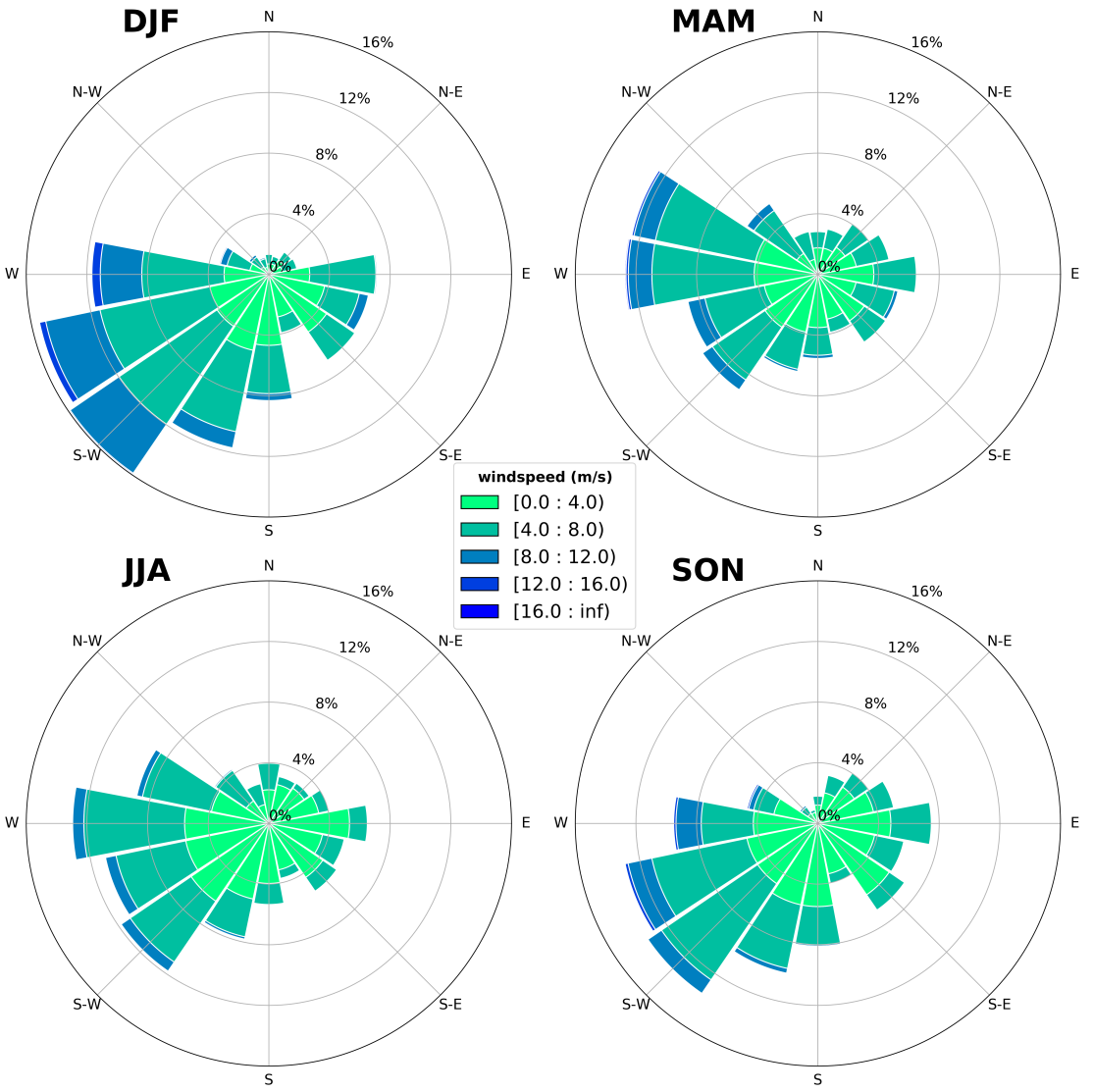


Fig. AF.1: *Windrose of hourly wind data measured at Bremen airport with wind speed colour-coded and percentages of all days as pie radius for the certain directions for the years 2015 - 2017. The four different seasons are shown (winter: DJF, spring: MAM, summer: JJA, autumn: SON), indicating a low seasonal variability.*

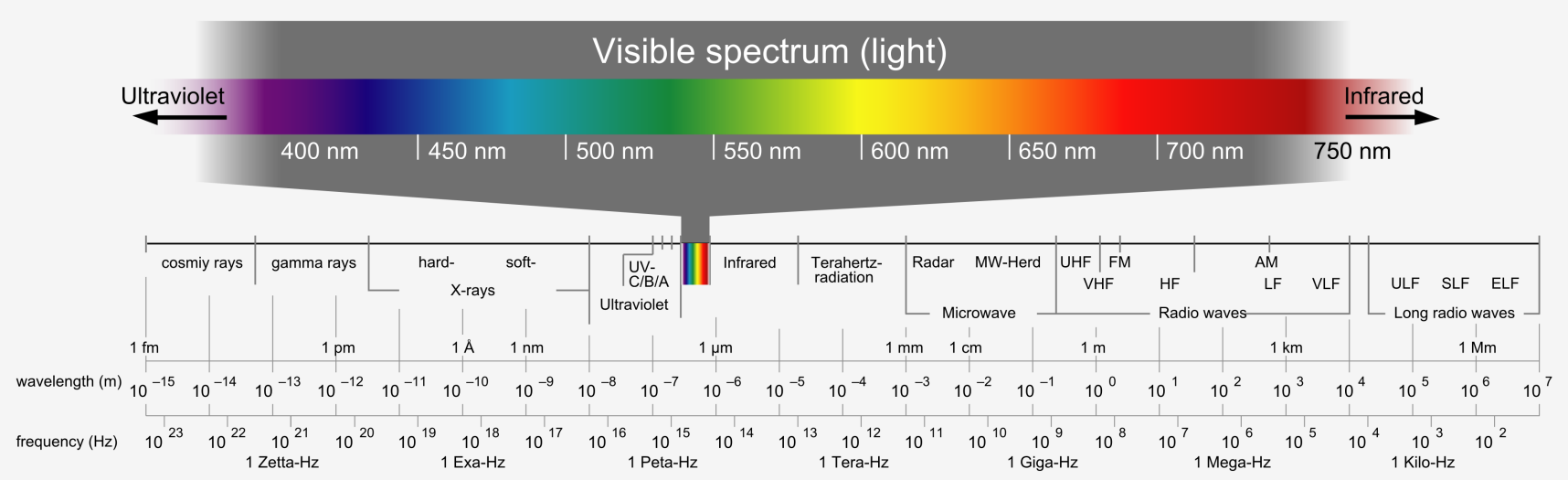


Fig. AF.2: *Representation of the electromagnetic spectrum with the most important spectral classifications including the visible spectral range.*[1]

Source: Picture of wikipedia user Horst Frank and Jailbird. URL: `https://commons.wikimedia.org/wiki/File:Electromagnetic_spectrum.svg`, Date of download: 08.05.2017

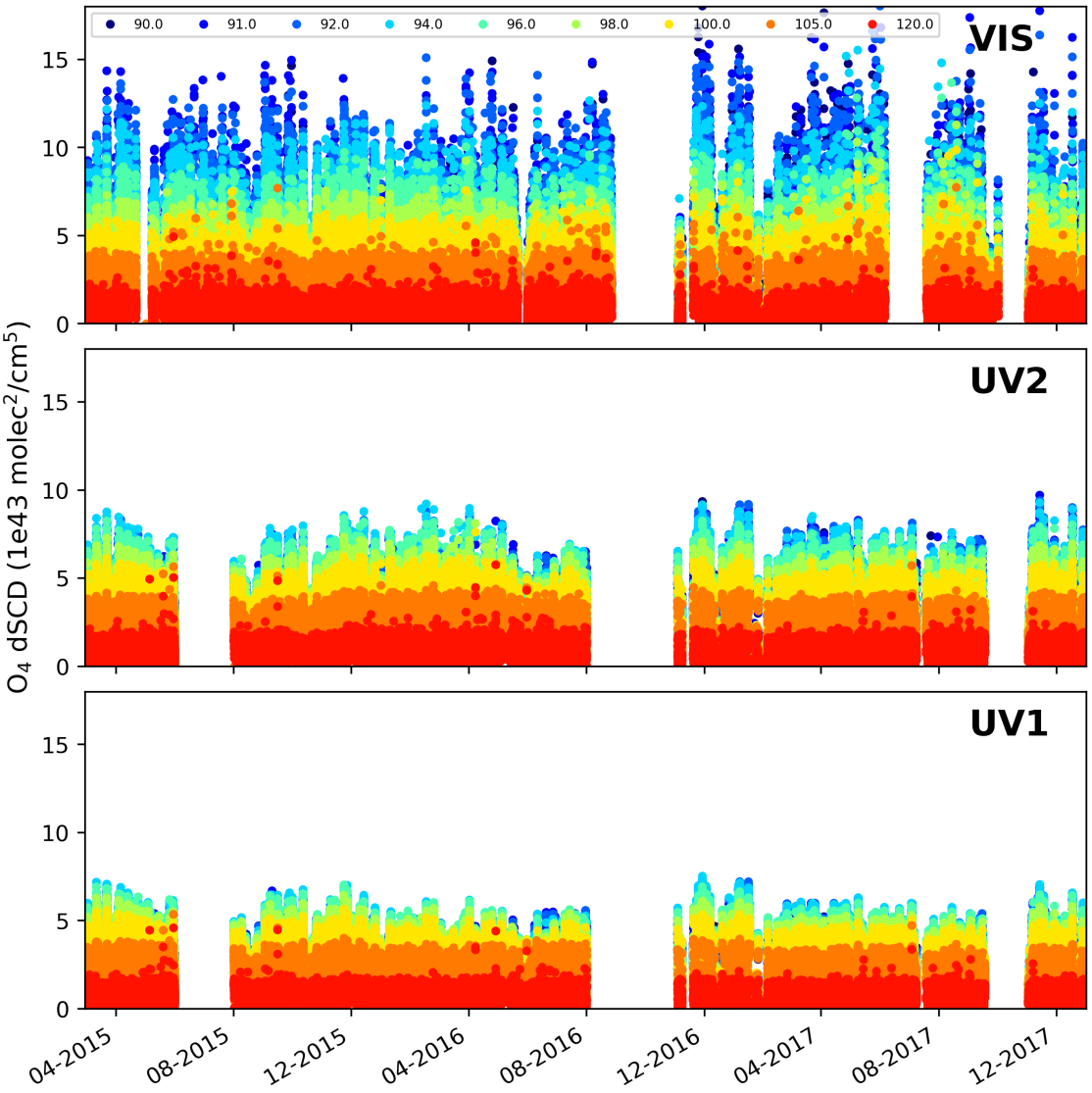


Fig. AF.3: *Time series of O_4 differential slant column densities during the full periode for different LOS, colour-coded (c direction). The three fits are shown in the individual subplots: VIS (top), UV2 (mid) and UV1 (bottom).*

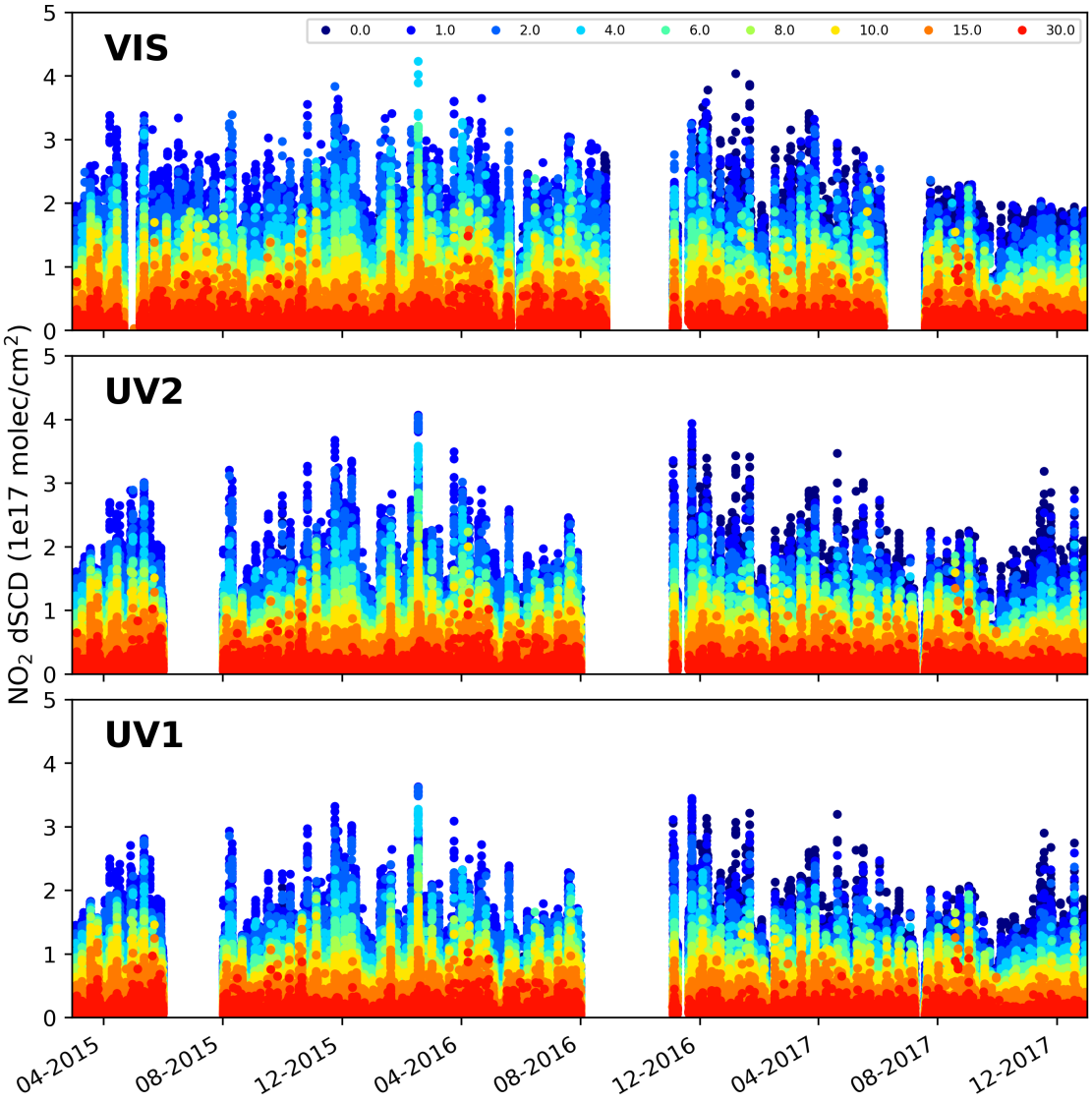


Fig. AF.4: *Time series of NO_2 differential slant column densities during the full period for different LOS, colour-coded (c direction). The three fits are shown in the individual subplots: VIS (top), UV2 (mid) and UV1 (bottom). Decrease in 2017 due to instrumental issues.*

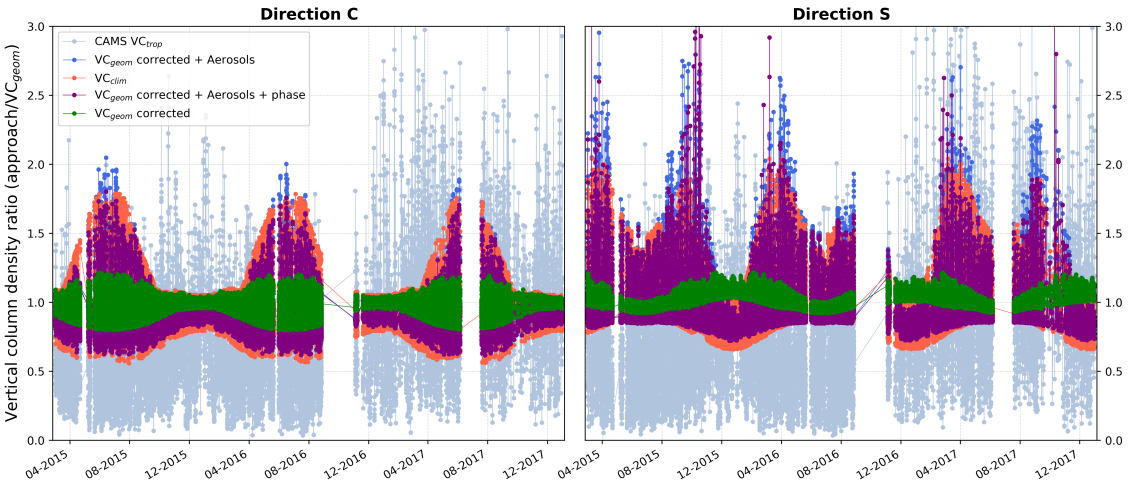


Fig. AF.5: *Comparison of the different VC calculation approaches including a correction of VC with a measured phase function (VC$_{geom}$ corrected + Aerosols + phase, in purple).*

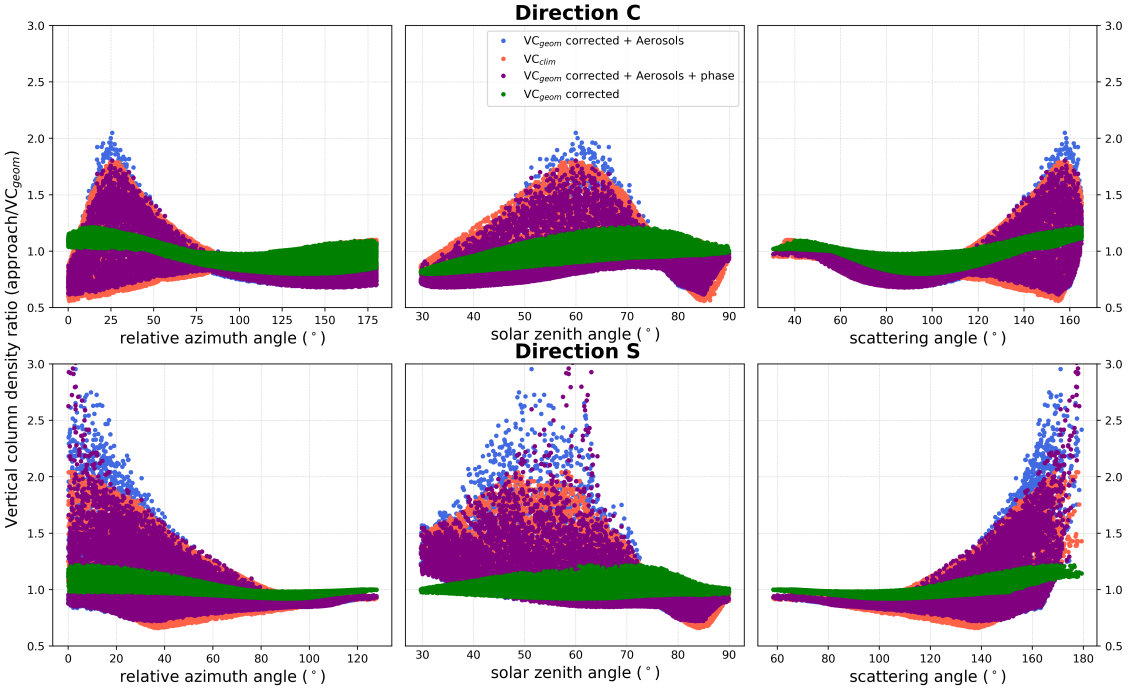


Fig. AF.6: *Comparison of the angular distribution of different VC calculation approaches including a correction of VC with a measured phase function (VC$_{geom}$ corrected + Aerosols + phase, in purple). The scattering angle Θ_s between incident and outgoing photon was calculated with $\Theta_s = \arccos\left(-\cos(\vartheta)\cos(\theta) - \sin(\vartheta)\sin(\theta)\cos(\varphi_{raa})\right)$ with the viewing azimuth angle θ, the solar zenith angle ϑ and the relative azimuth angle φ_{raa}.*

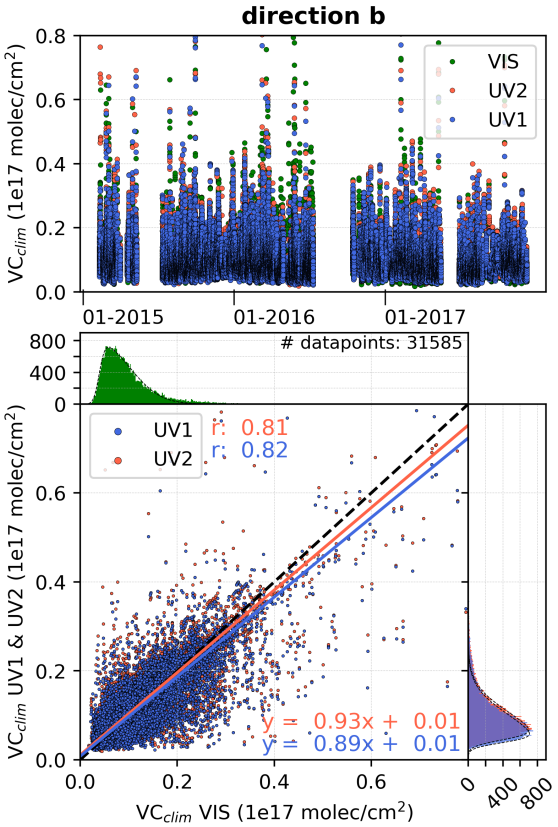


Fig. AF.7: Comparison of the different fitting windows used for the calculation of VC_{clim} for direction b. Top: Time series, Bottom: Correlation plots of visible and UV fitting windows including histograms depicting the frequency distribution of each data set.

	direction c			direction b			direction s		
	UV1	UV2	VIS	UV1	UV2	VIS	UV1	UV2	VIS
Median (%)	-3.1	4.7	-1.4	-2.8	4.5	-1.6	-2.0	4.3	-3.1
Mean (%)	-2.6	4.7	-2.1	-2.4	4.2	-1.8	-1.7	3.6	-1.6
Mean (1e15 molec/cm^2)	3.8	4.5	3.7	3.9	4.5	3.9	3.9	4.4	3.9
σ (1e15 molec/cm^2)	6.2	6.2	6.4	6.7	6.6	6.4	7.2	7.1	6.8
Skewness γ	5.7	5.2	6.9	6.2	5.3	7.1	5.7	5.4	7.4

Tab. AT.1: Mean and median relative deviations of all fitting windows (UV1, UV2, VIS) to the mean value of the specific VC for all three fits (2 top rows). Fitting parameters of skewed Gaussian distributions fitted to the histogram of data points (3 bottom rows). No RAA filtering was applied.

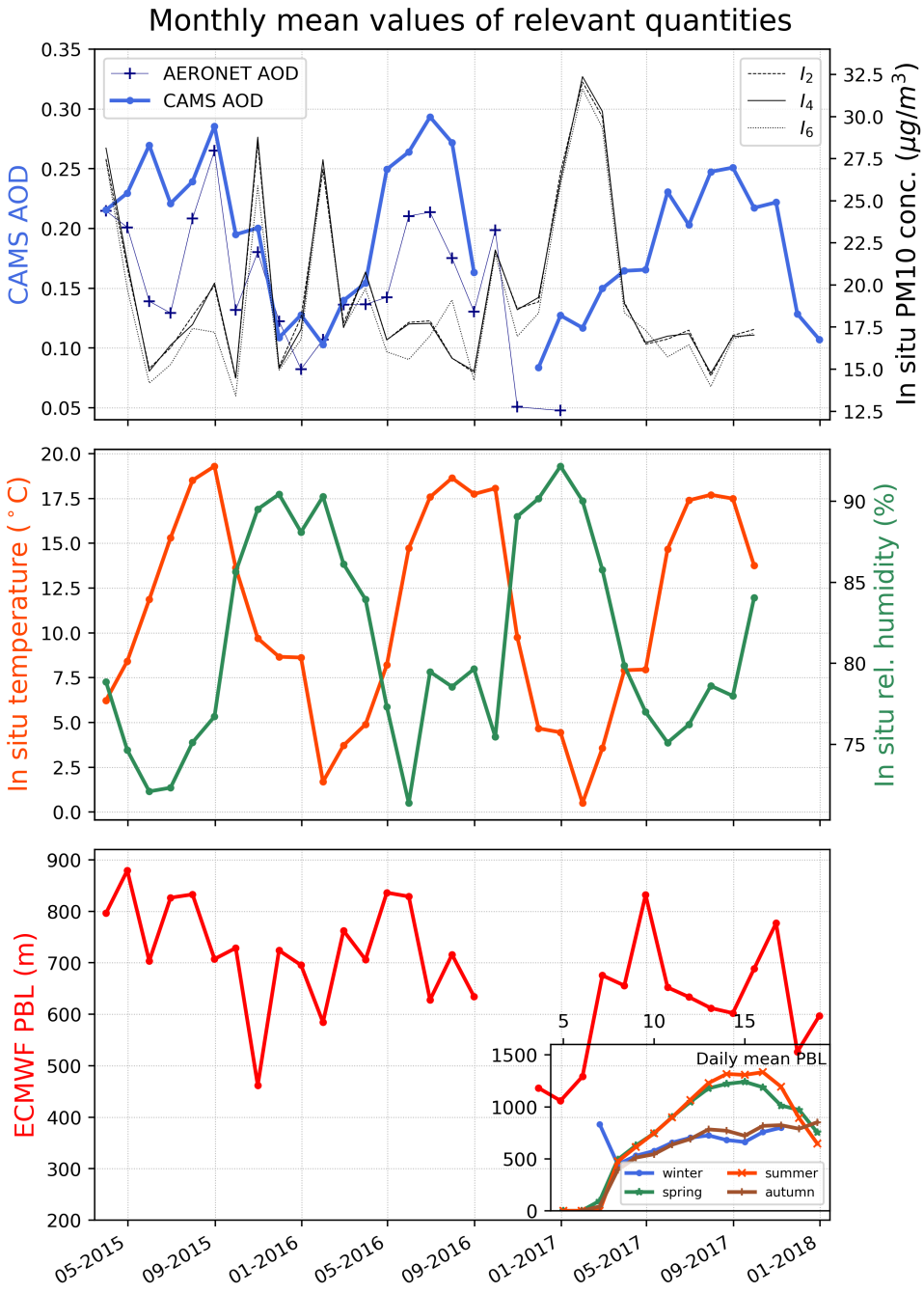


Fig. AF.8: *Time series of monthly means values. Top: CAMS AOD, in situ stations and the AERONET station in Hamburg[2]. Mid: In situ temperature and relative humidity of in situ site I_4. Bottom: ECMWF planetary boundary layer height and daily variation of seasonal PBL (small inlay plot) in UTC.*

[2] The AERONET data was kindly provided by Stefan Kinne (MPIMET, Hamburg, Germany). The reader is referred to the relevant literature for detailed information (HOLBEN et al. [1998]; DUBOVIK et al. [2000]).

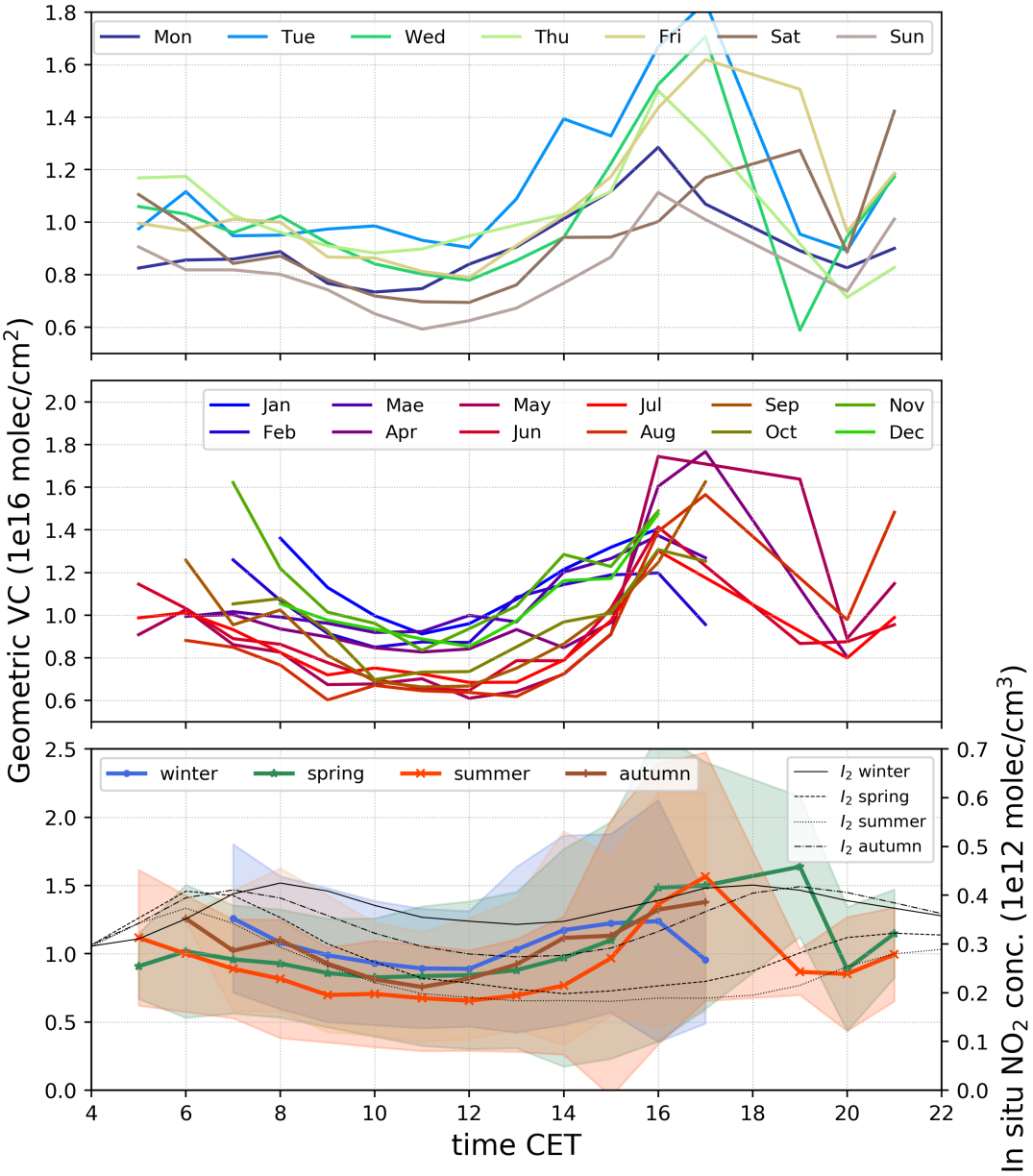


Fig. AF.9: *Vertical column densities for direction b as daily variation for, each day of the week (top), each month (mid) and each season (bottom) averaged over all three years. Additionally, the bottom plots show in situ average NO_2 concentrations for site I_2. Note, that around 18h o'clock, the data point is missing because the an applied relative azimuth angle filter for values with RAA$\leq \pm 20°$.*

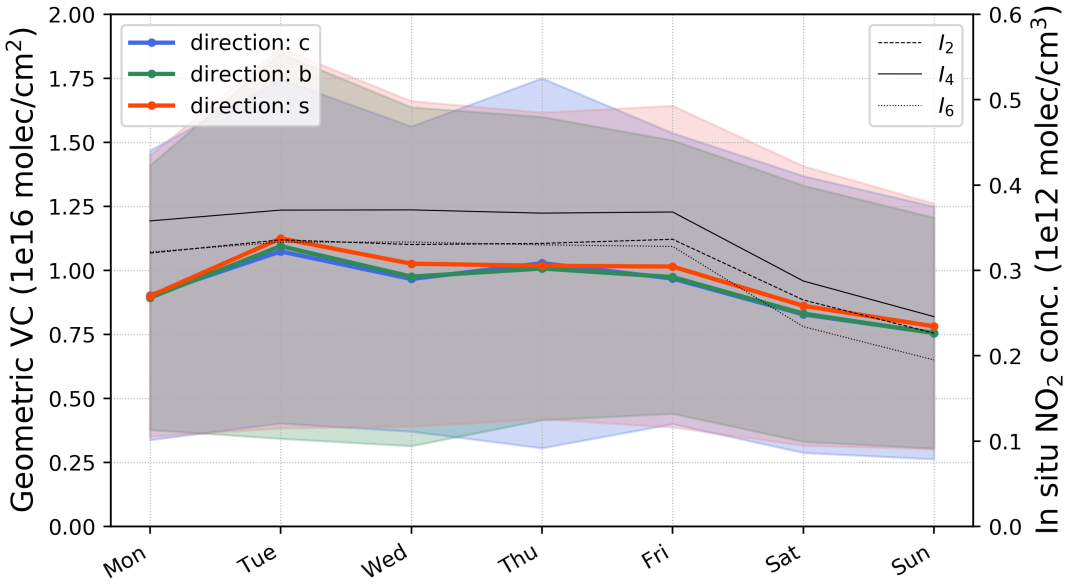


Fig. AF.10: *Vertical column density variation of NO_2 for all weekdays averaged over the full time period for all directions including the standard deviation as shaded areas.*

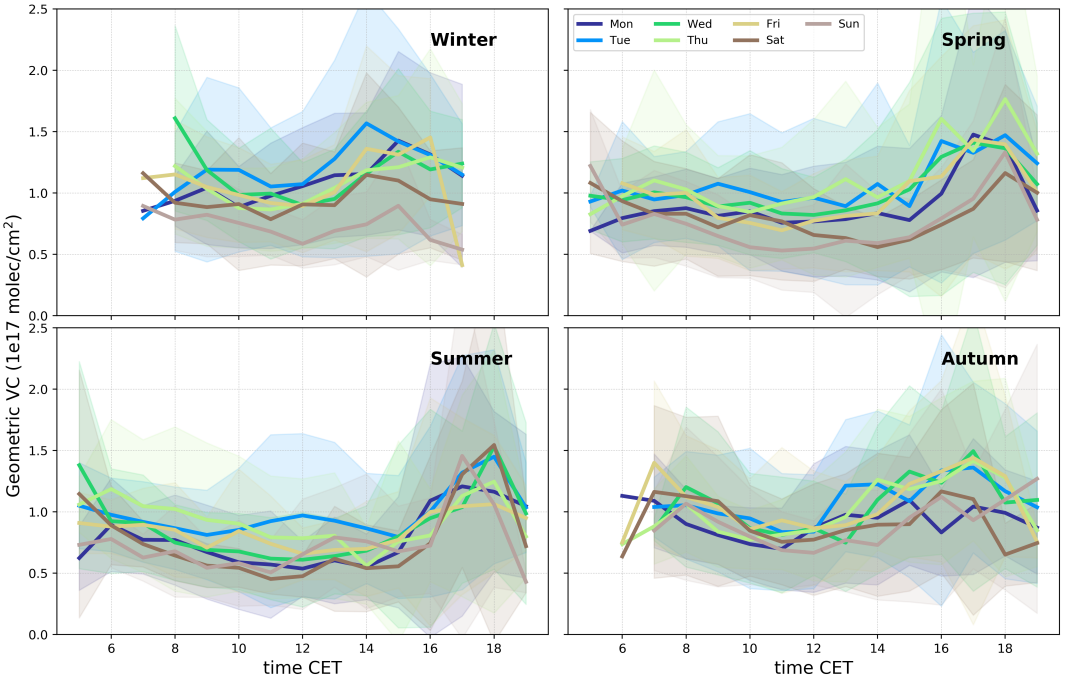


Fig. AF.11: *Vertical column density variation of NO_2 for all weekdays averaged over the full time period for all directions including the standard deviation as shaded areas.*

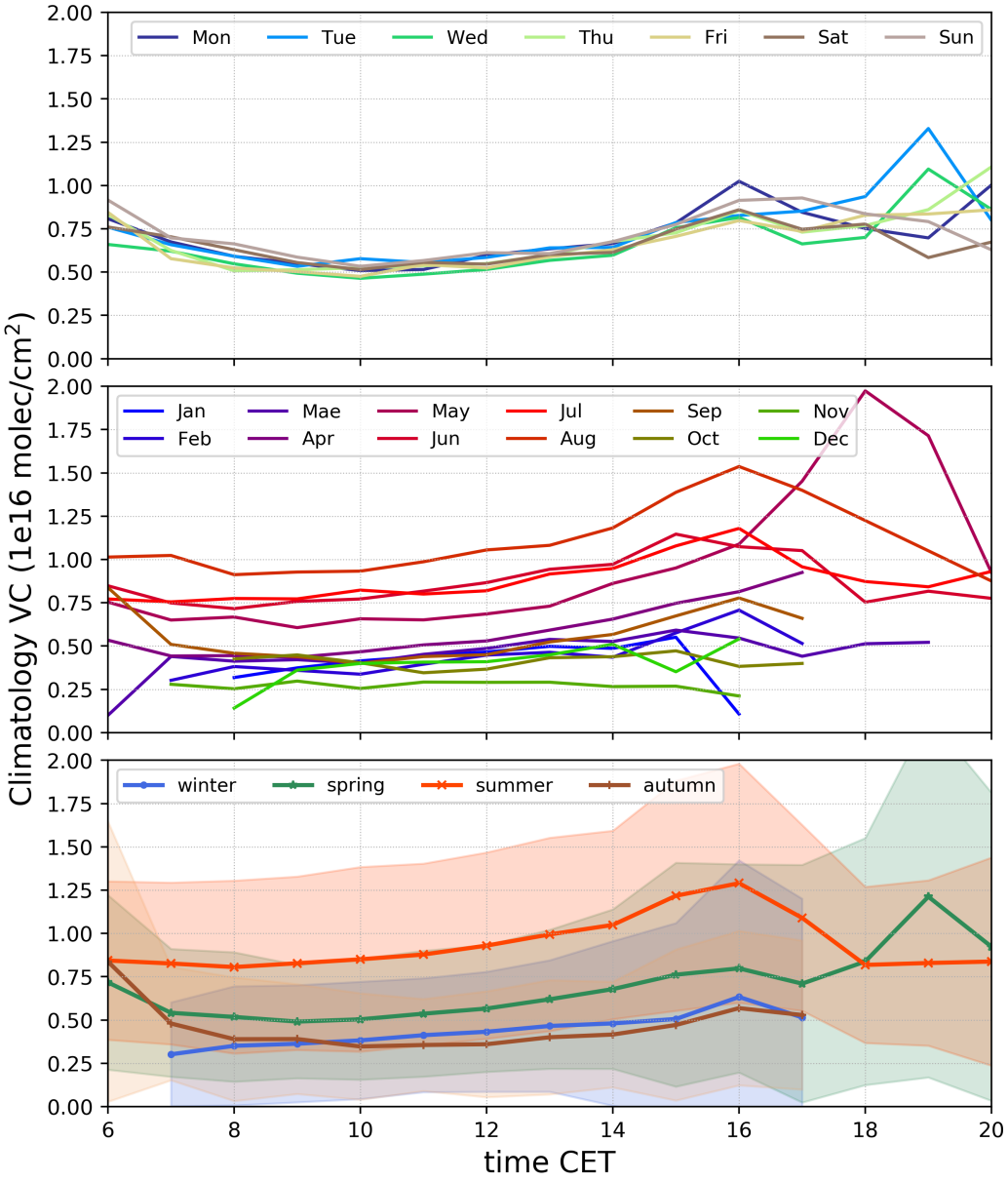


Fig. AF.12: *Vertical column densities of HCHO for direction b as daily variation for, each day of the week (top), each month (mid) and each season (bottom) averaged over all three years. Note, that around 18h o'clock, the data point is missing because the an applied relative azimuth angle filter for values with RAA* $\leq \pm 20°$.

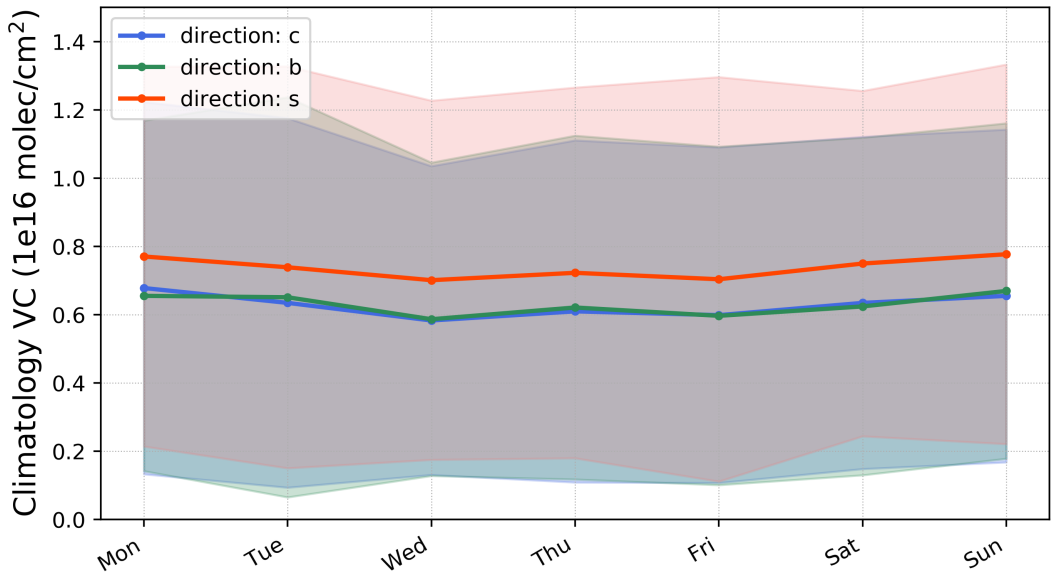


Fig. AF.13: *Vertical column density variation of HCHO for all weekdays averaged over the full time period for all directions including the standard deviation as shaded areas.*

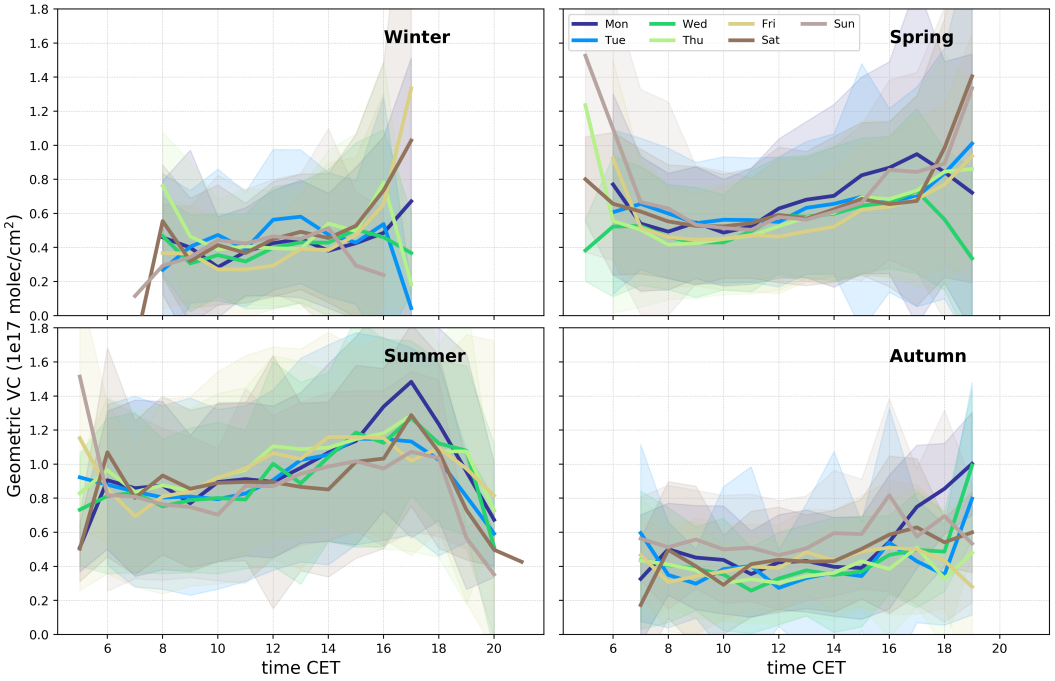


Fig. AF.14: *Vertical column density variation of HCHO for all weekdays averaged over the full time period for all directions including the standard deviation as shaded areas.*

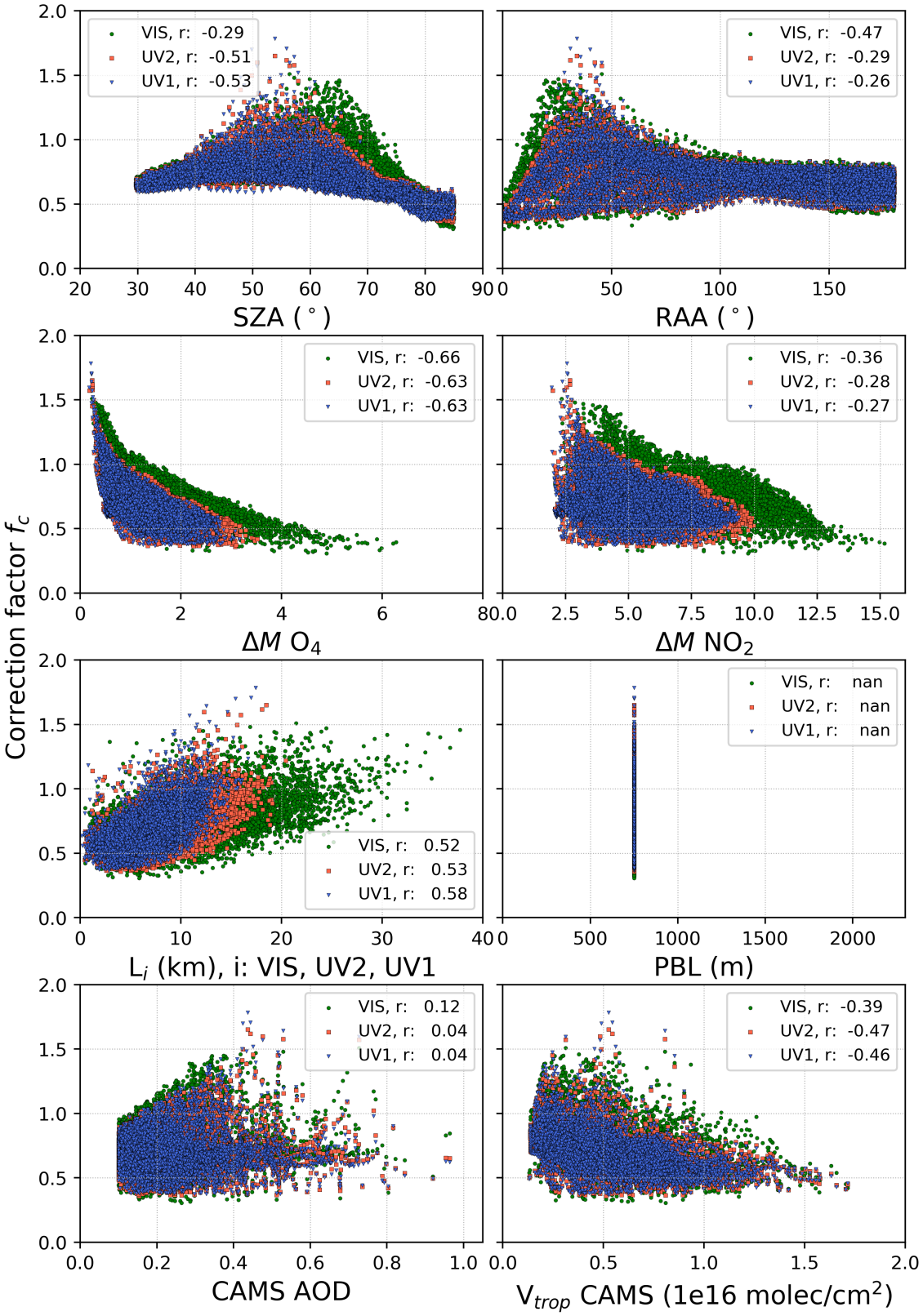


Fig. AF.15: *Scatter plot of f_c with various parameters for three different fitting windows and direction s (colour-coded). Pearson's correlation coefficient is given in the legend of each subplot. PBL and V_{O_4} were kept constant.*

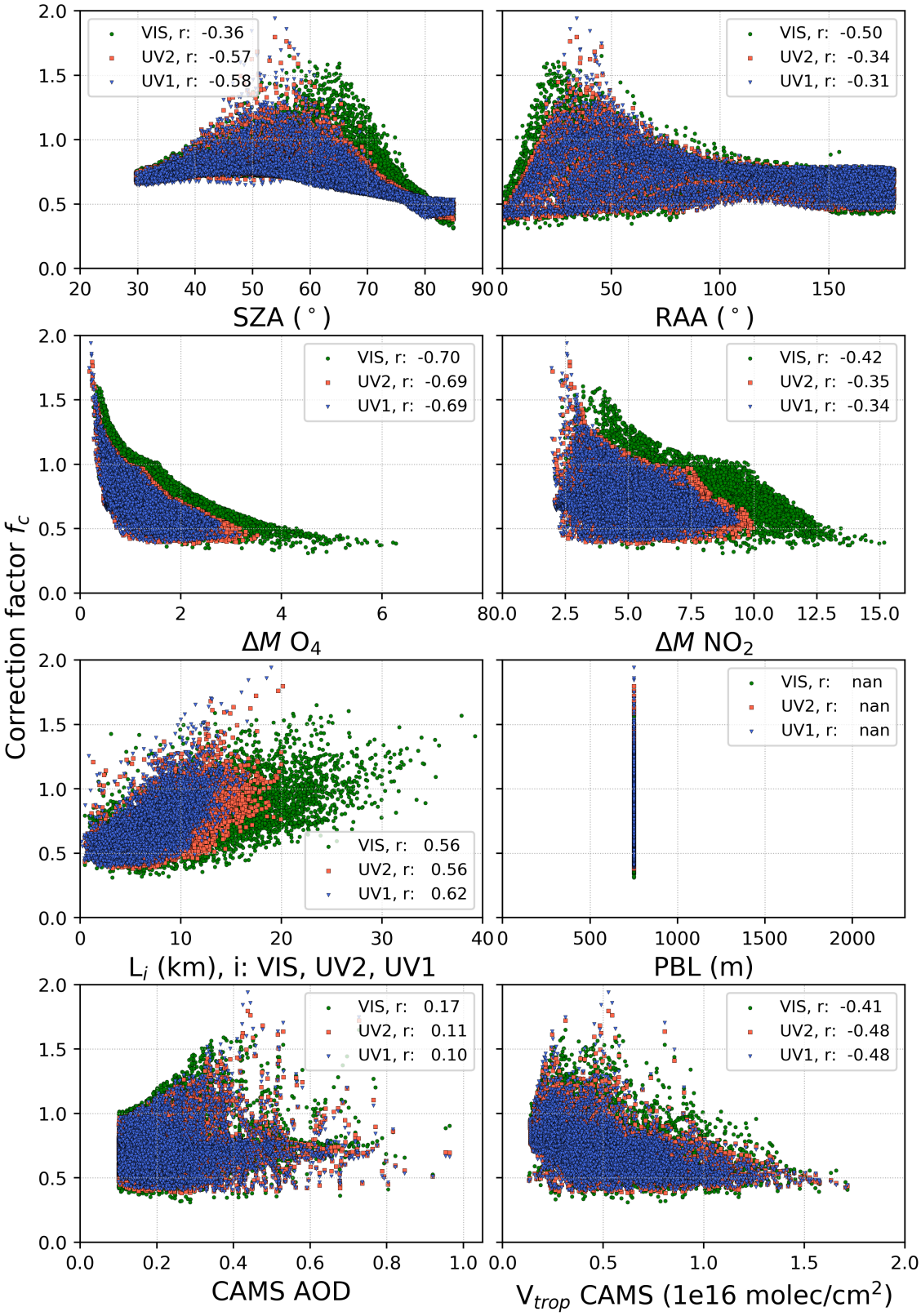


Fig. AF.16: *Scatter plot of f_c with various parameters for three different fitting windows and direction c (colour-coded). Pearson's correlation coefficient is given in the legend of each subplot. PBL was kept constant, V_{O_4} was scaled.*

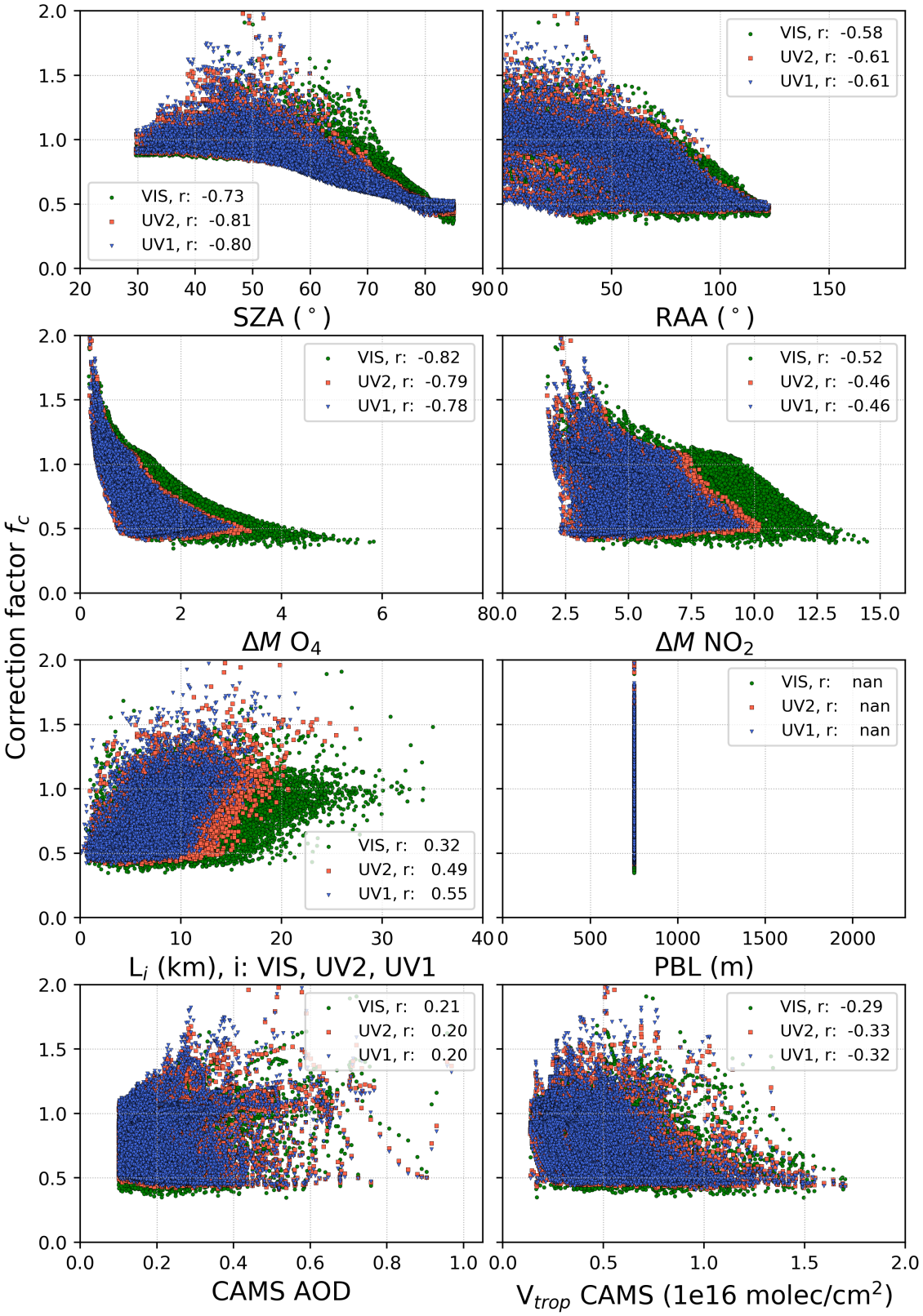


Fig. AF.17: *Scatter plot of f_c with various parameters for three different fitting windows and direction s (colour-coded). Pearson's correlation coefficient is given in the legend of each subplot. PBL was kept constant, V_{O_4} was scaled.*

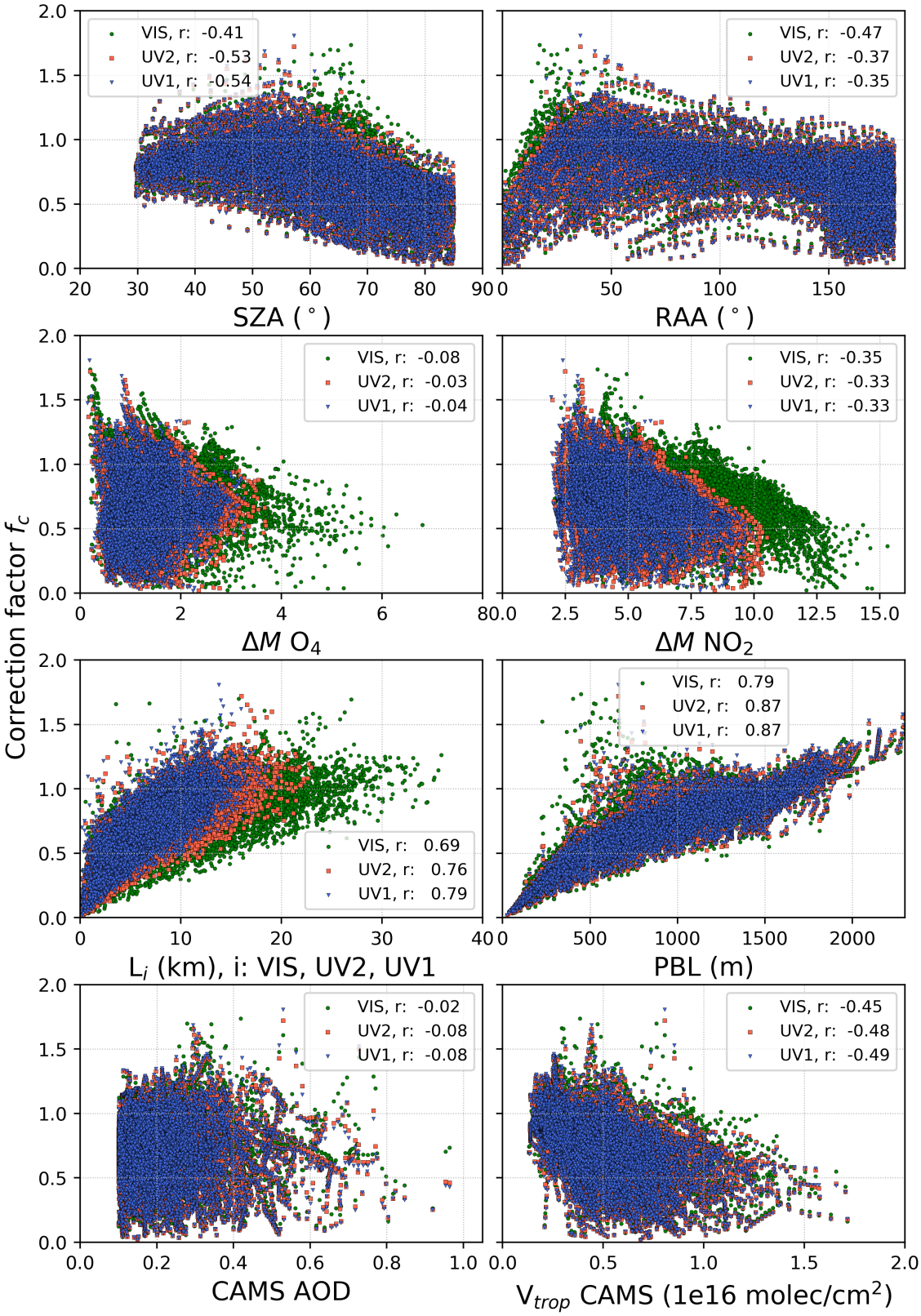


Fig. AF.18: *Scatter plot of f_c with various parameters for three different fitting windows and direction c (colour-coded). Pearson's correlation coefficient is given in the legend of each subplot. PBL was applied from ECMWF, V_{O_4} was kept constant.*

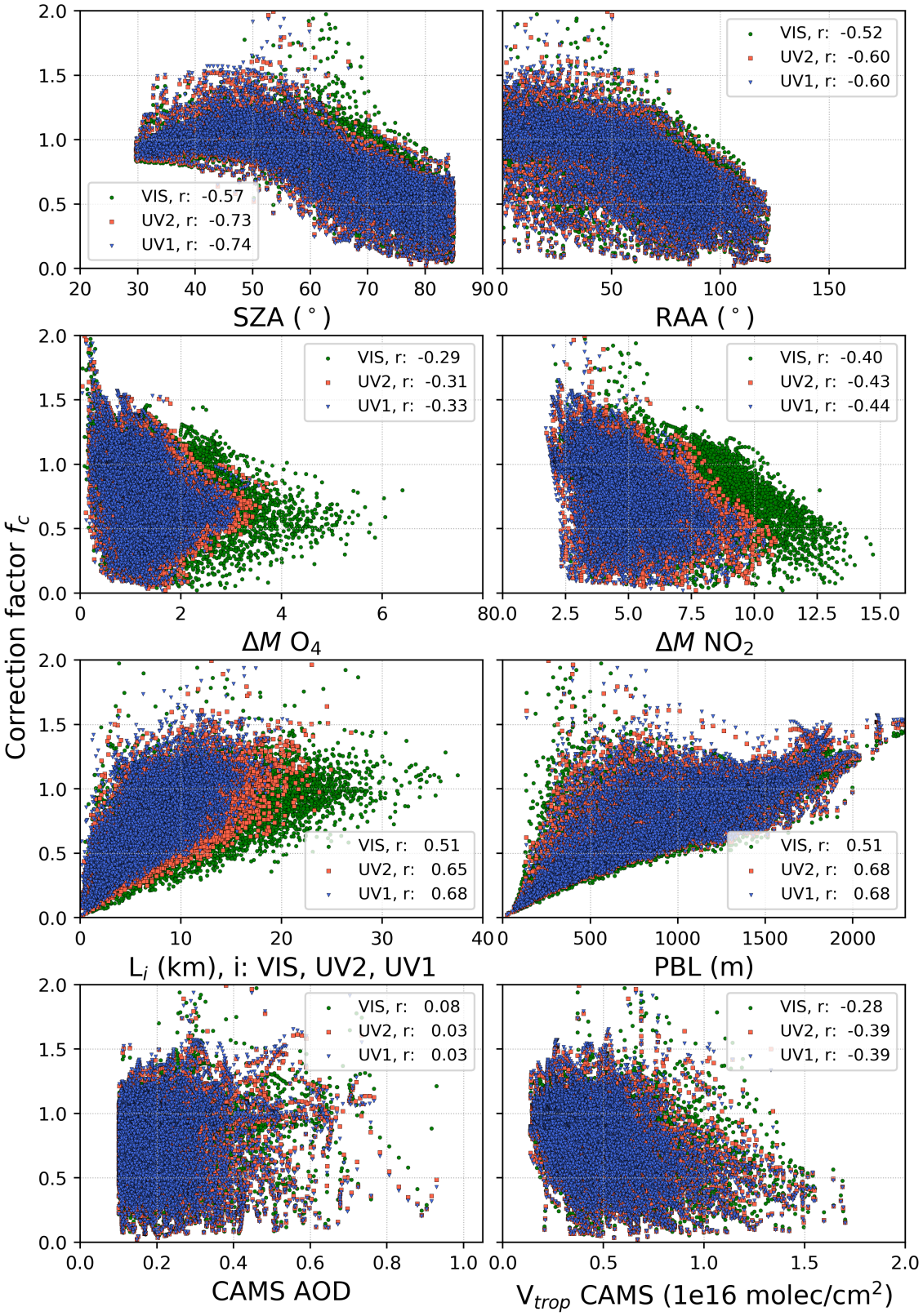


Fig. AF.19: *Scatter plot of f_c with various parameters for three different fitting windows and direction s (colour-coded). Pearson's correlation coefficient is given in the legend of each subplot. PBL was applied from ECMWF, V_{O_4} was kept constant.*

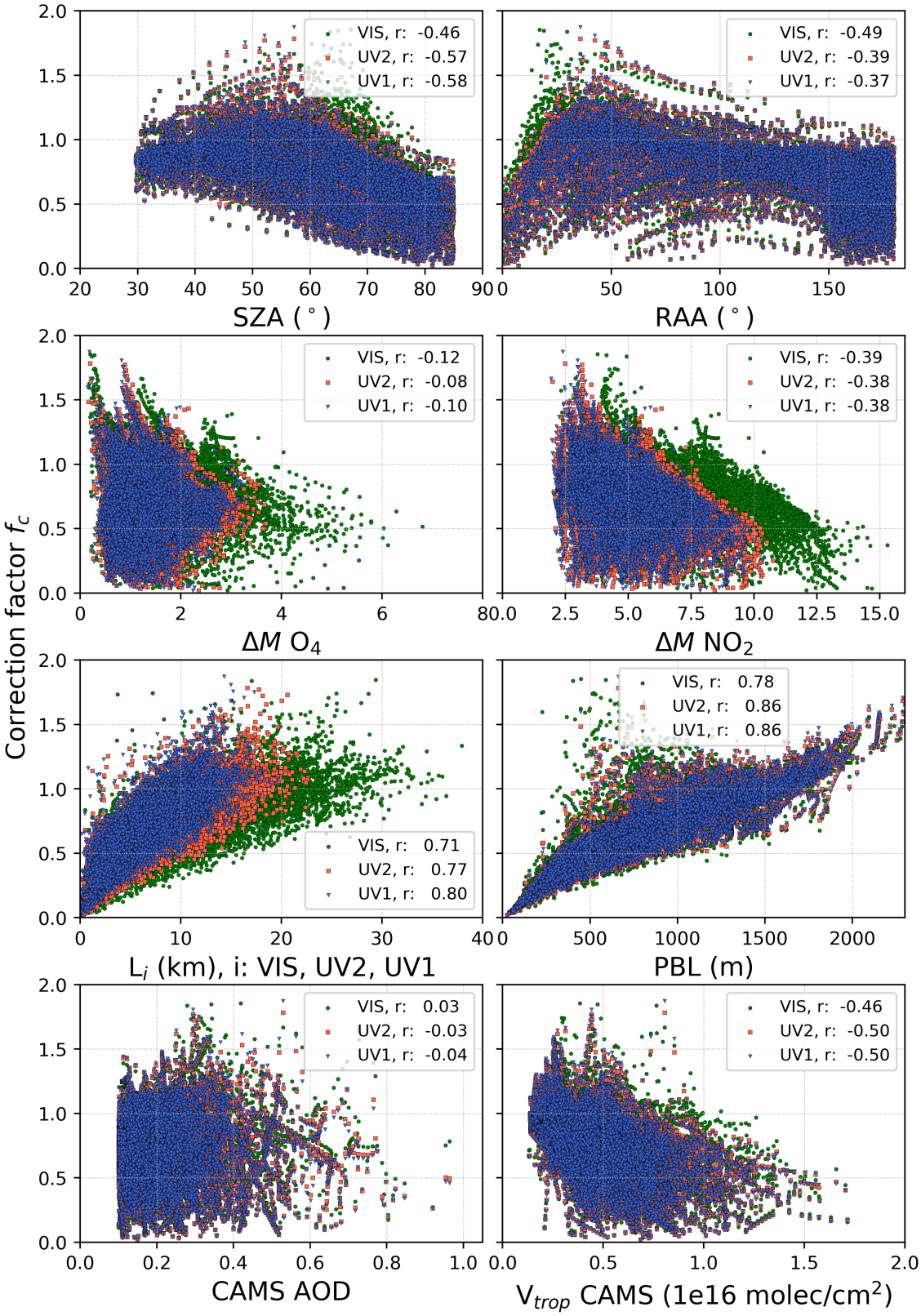


Fig. AF.20: *Scatter plot of f_c with various parameters for three different fitting windows and direction c (colour-coded). Pearson's correlation coefficient is given in the legend of each subplot. PBL was applied from ECMWF, V_{O_4} was scaled.*

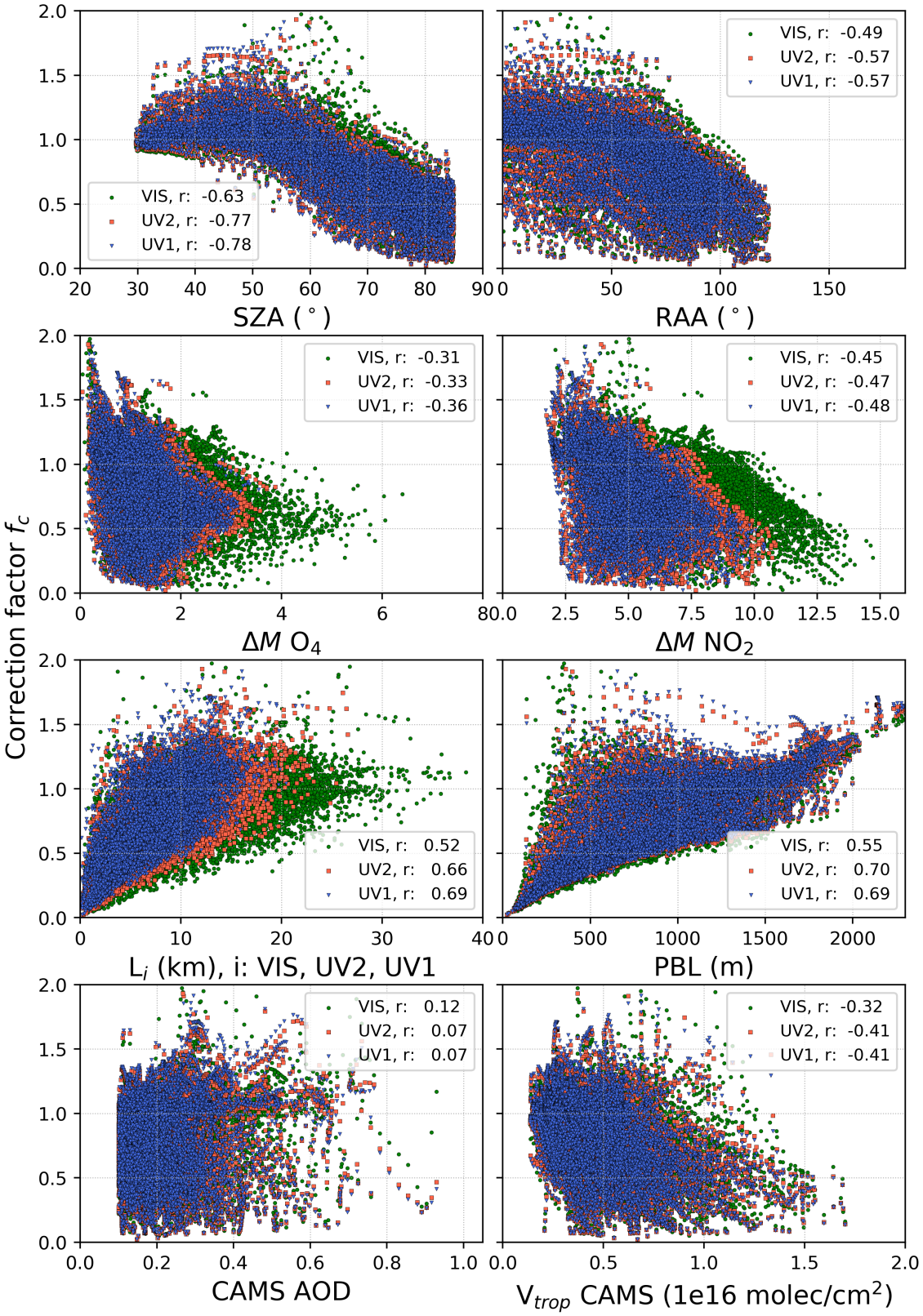


Fig. AF.21: *Scatter plot of f_c with various parameters for three different fitting windows and direction s (colour-coded). Pearson's correlation coefficient is given in the legend of each subplot. PBL was applied from ECMWF, V_{O_4} was scaled.*

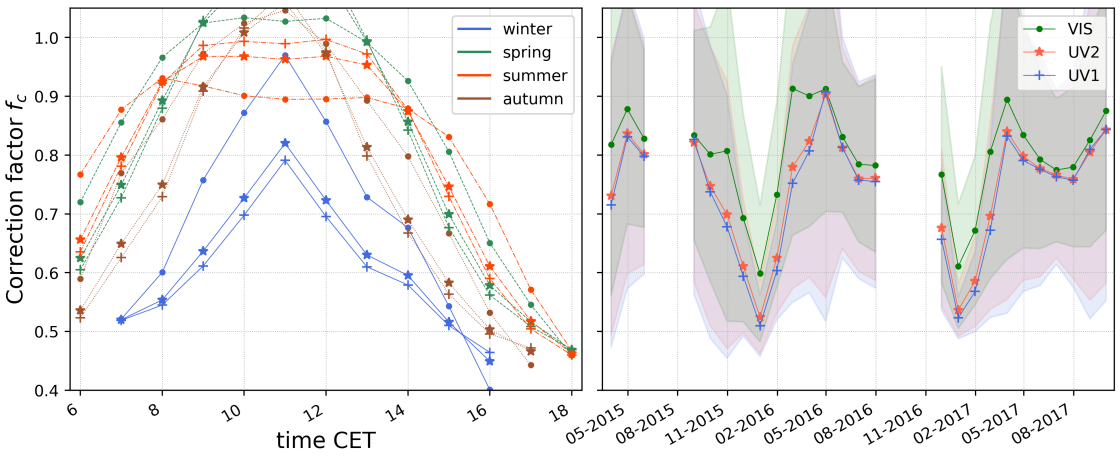


Fig. AF.22: Left: *Mean f_c values for different seasons of the year colour-coded and for direction s. Different marker styles depict the different fitting windows (see right legend).* ***Right:*** *Time series of monthly averaged f_c with different fitting windows colour-coded. PBL and V_{O_4} were kept constant.*

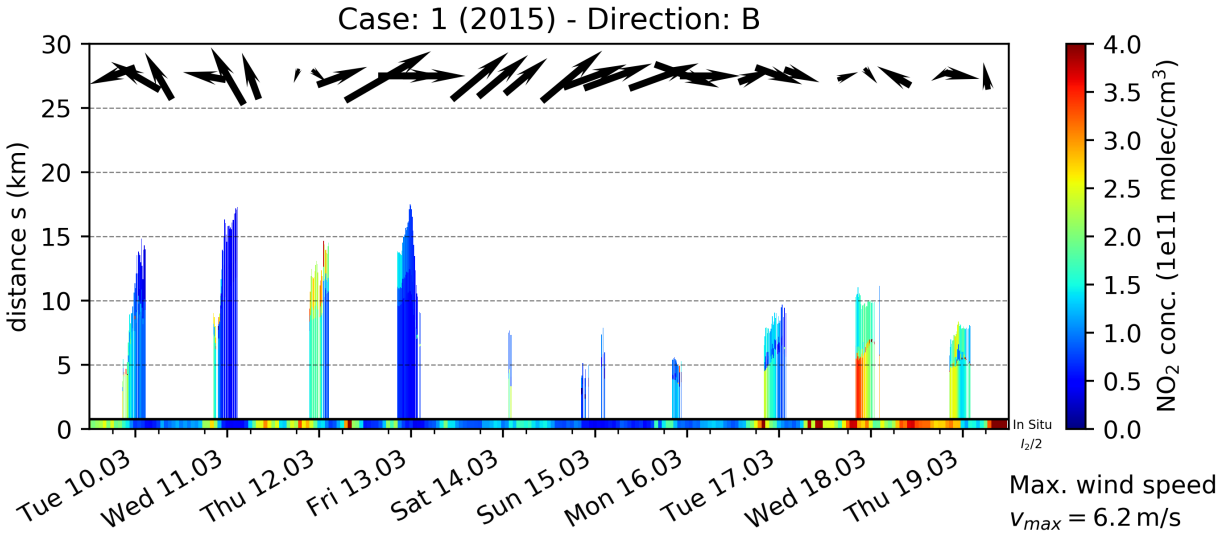


Fig. AF.23: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ concentrations divided by 2 are shown at the smallest distance.*

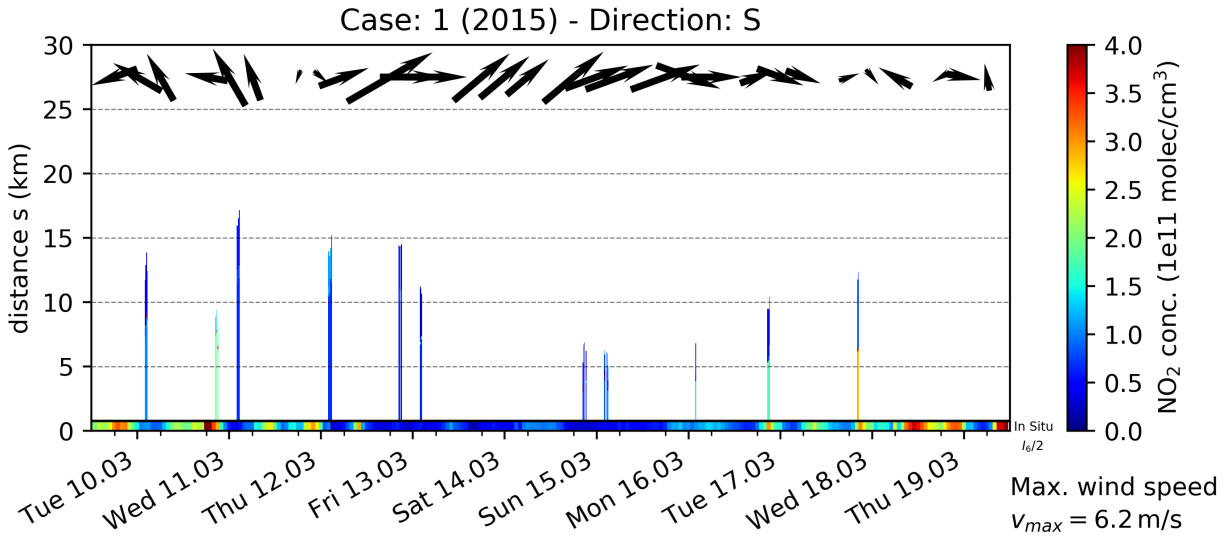


Fig. AF.24: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ concentrations divided by 2 are shown at the smallest distance.*

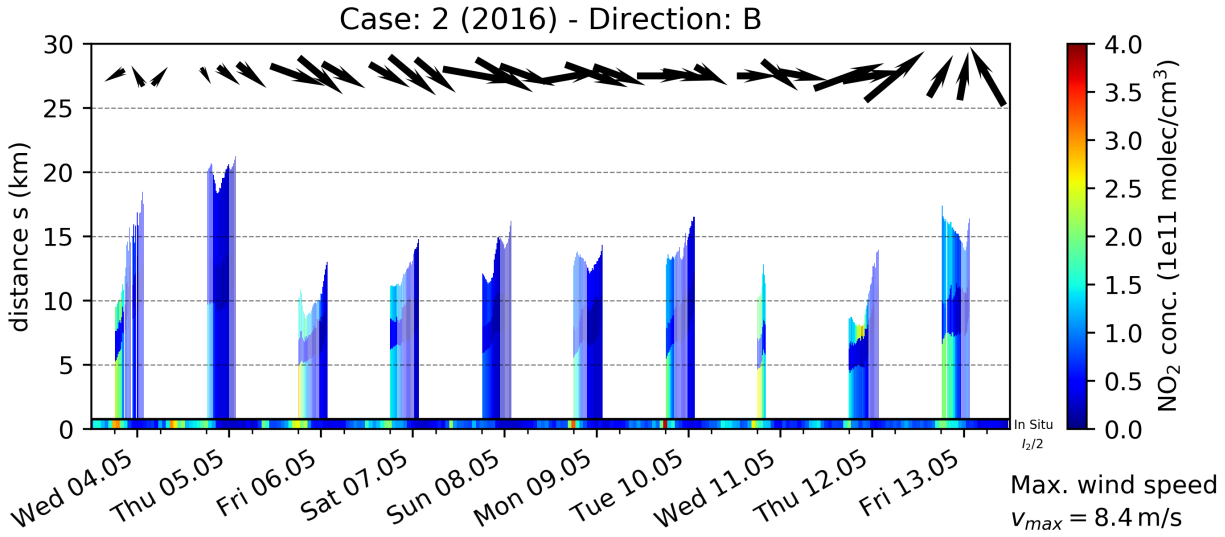


Fig. AF.25: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ concentrations divided by 2 are shown at the smallest distance.*

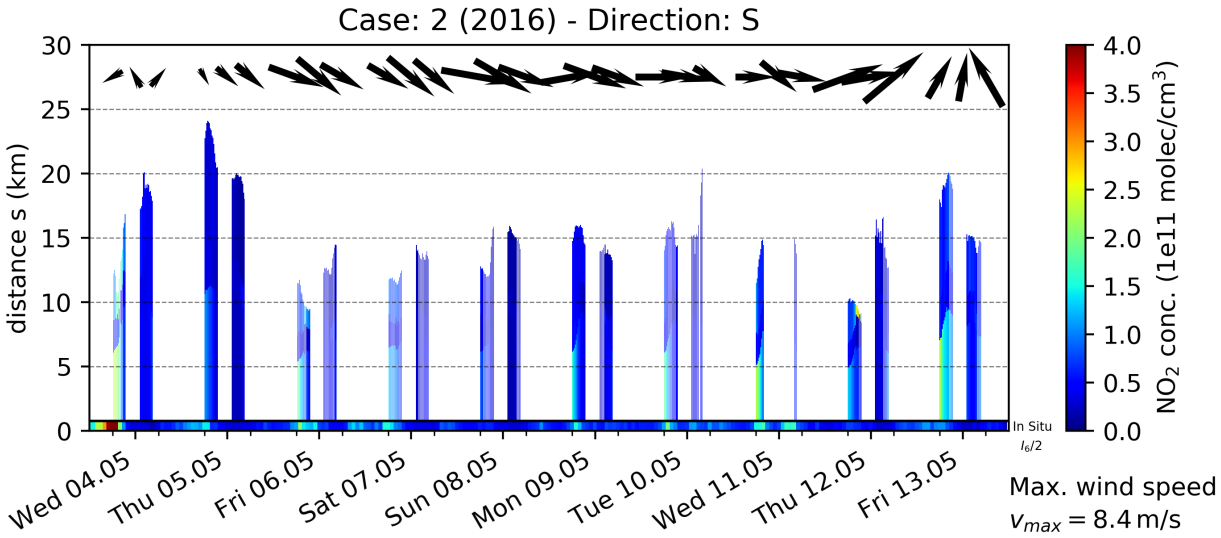


Fig. AF.26: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ concentrations divided by 2 are shown at the smallest distance.*

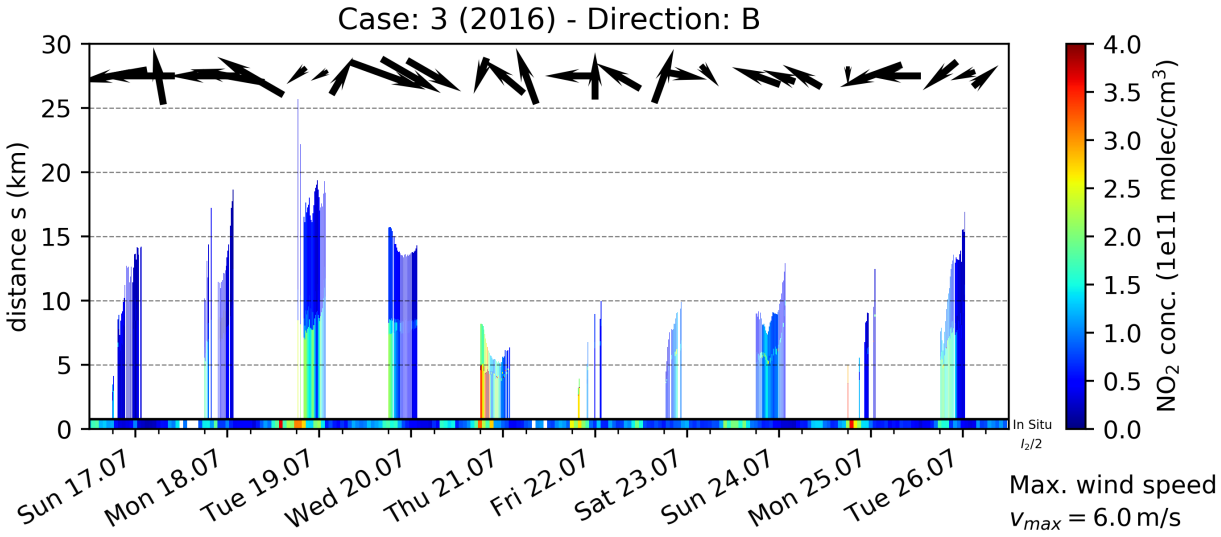


Fig. AF.27: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ concentrations divided by 2 are shown at the smallest distance.*

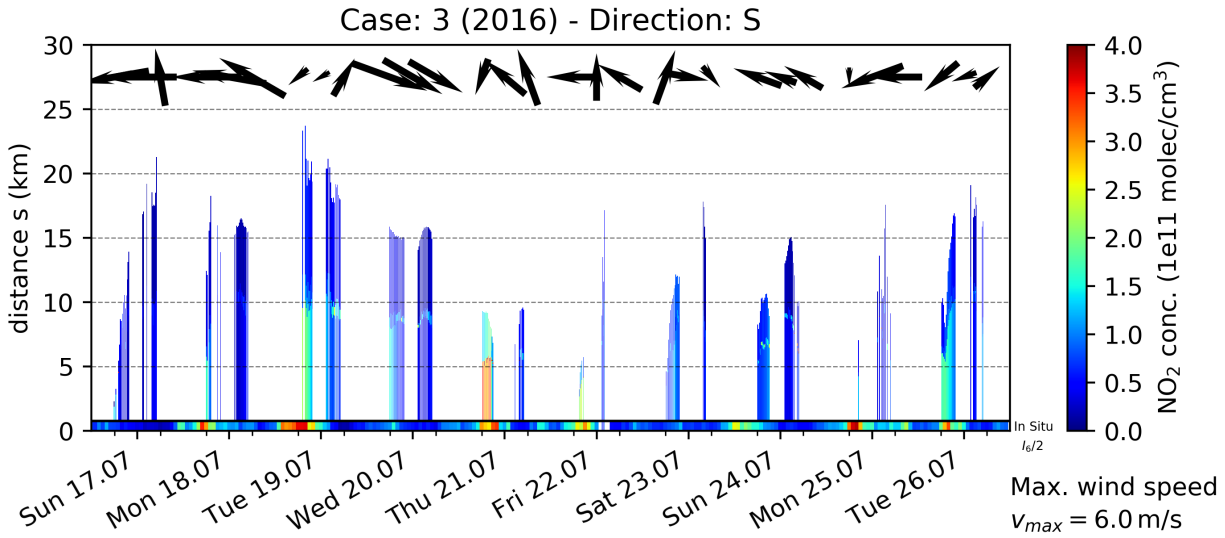


Fig. AF.28: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ concentrations divided by 2 are shown at the smallest distance.*

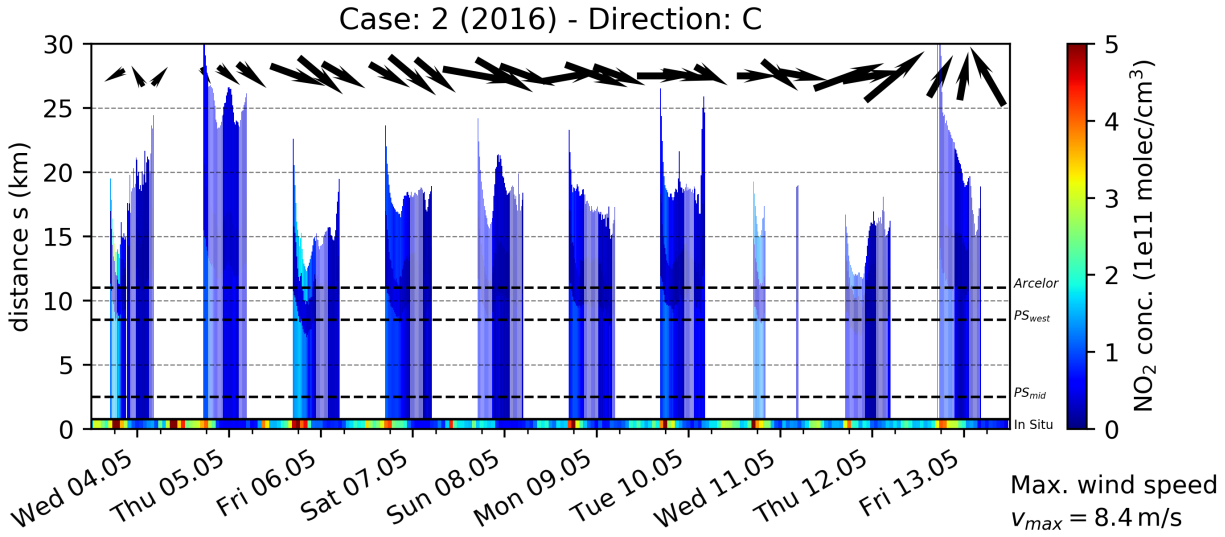


Fig. AF.29: *Distance depending NO_2 concentrations, colour-coded. The black arrows depict the prevalent wind speeds and directions at 06:00, 12:00 and 18:00 LT with the maximal wind speed as a value given in the lower right corner. Continuous in situ concentrations divided by 2 are shown at the smallest distance. The black dashed lines show the distance to strong emitters.*

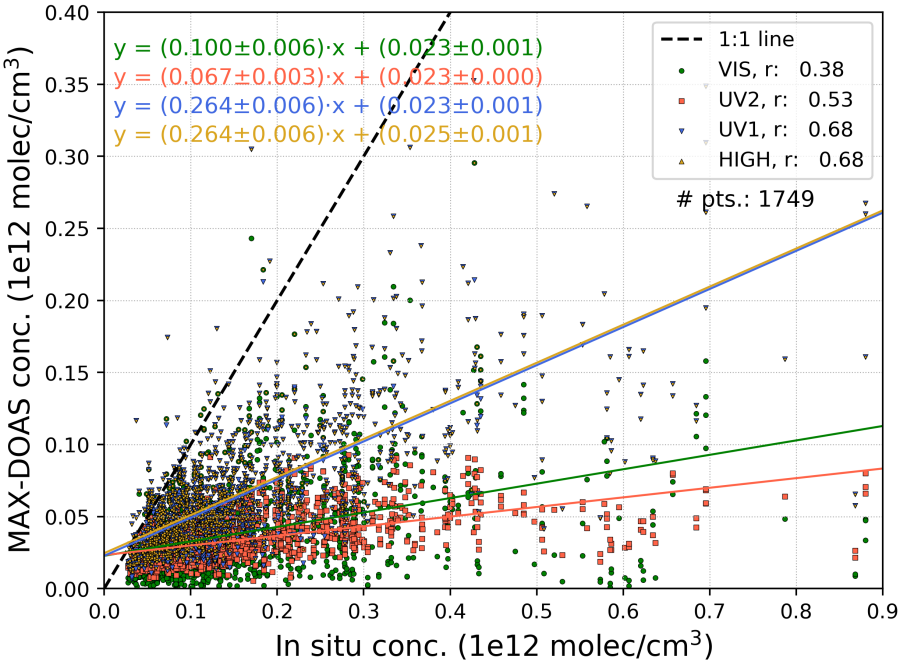


Fig. AF.30: *Correlation plot of onion peeling NO_2 concentration datasets with in situ measurements for direction b. The dataset HIGH summarizes the highest concentrations of the three onion peeling values per time step.*

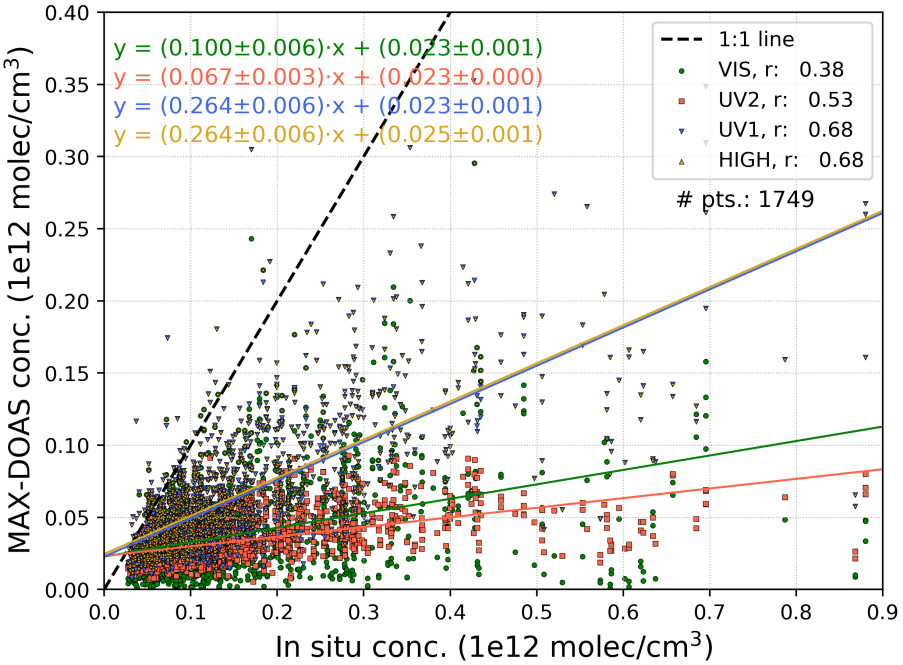


Fig. AF.31: *Correlation plot of onion peeling NO_2 concentration datasets with in situ measurements for direction s. The dataset HIGH summarizes the highest concentrations of the three onion peeling values per time step.*

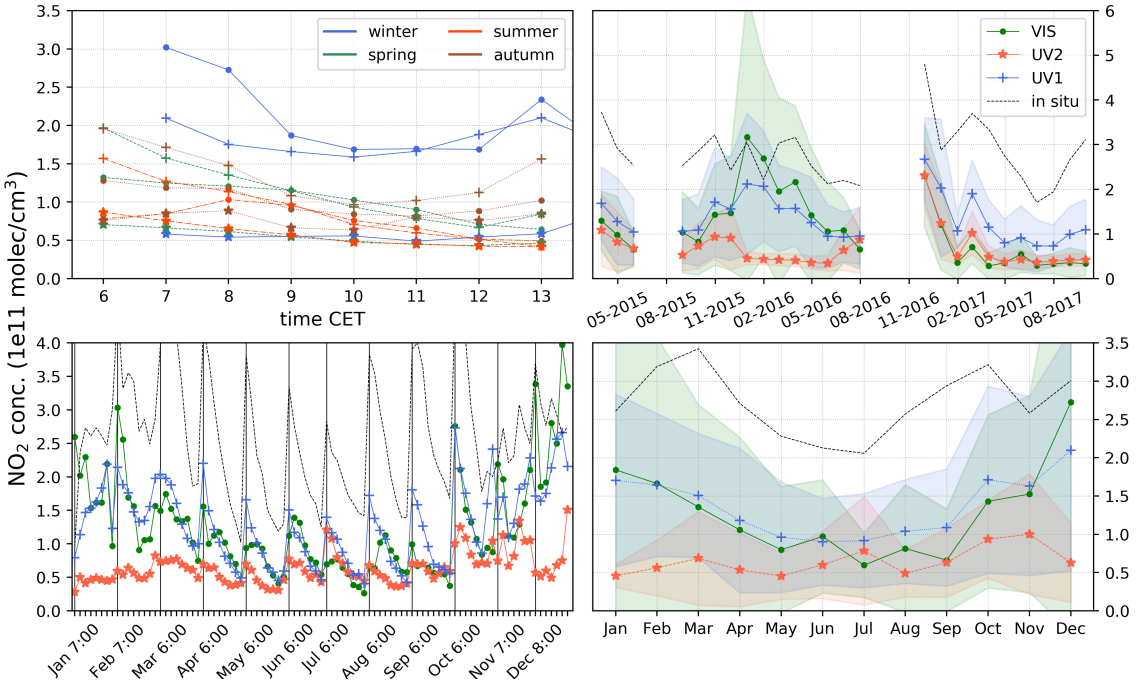


Fig. AF.32: *NO_2 concentrations for the three fitting windows as temporal mean values and for direction b.* ***Top left****: Daily variation for the four seasons of the year.* ***Top right****: Time series of monthly mean values.* ***Bottom left****: Mean daily variations for each month of the year.* ***Bottom right****: Monthly mean values averaged over the full time period. Different marker styles indicate the three fitting windows while the shaded area shows the standard deviations. Black, dashed lines represent in situ concentrations.*

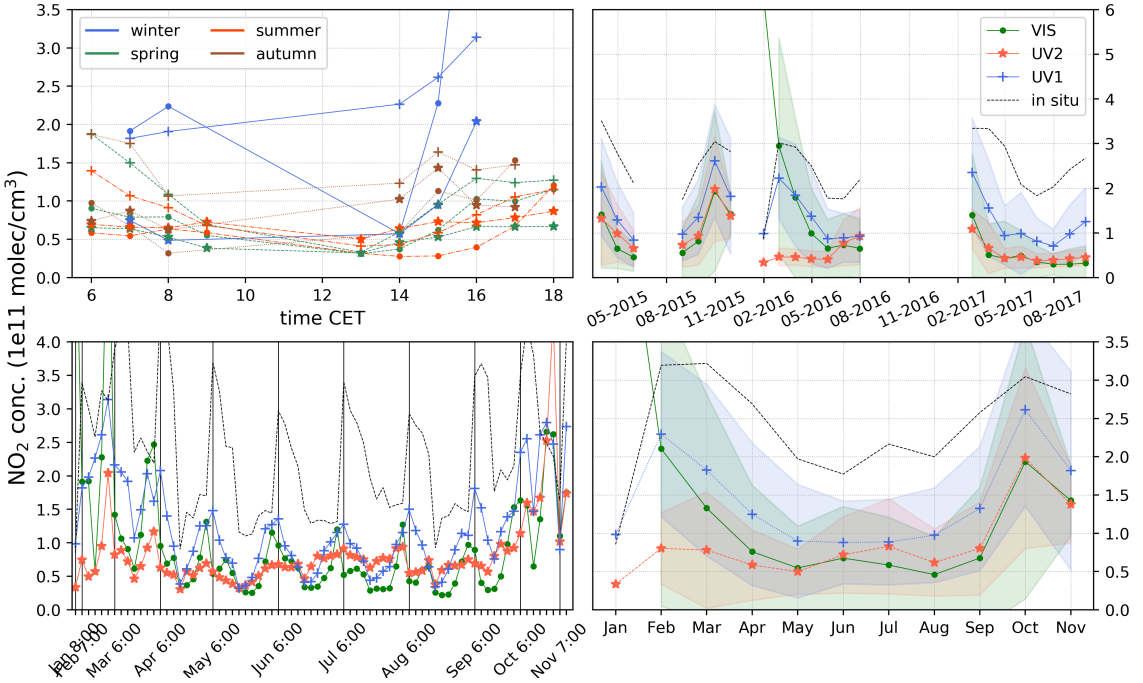


Fig. AF.33: *As above but for direction s.*

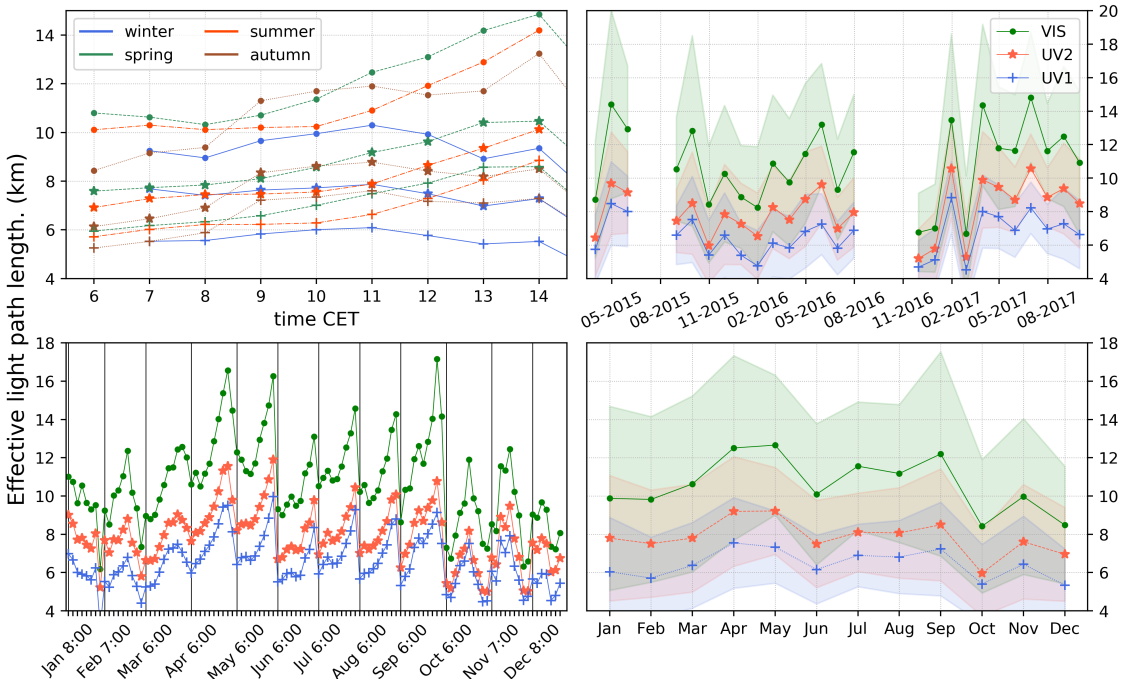


Fig. AF.34: *Effective light paths for the three fitting windows as temporal mean values and for direction c.* ***Top left****: Daily variation for the four seasons of the year.* ***Top right****: Time series of monthly mean values.* ***Bottom left****: Mean daily variations for each month of the year.* ***Bottom right****: Monthly mean values averaged over the full time period. Different marker styles indicate the three fitting windows while the shaded area shows the standard deviations.*

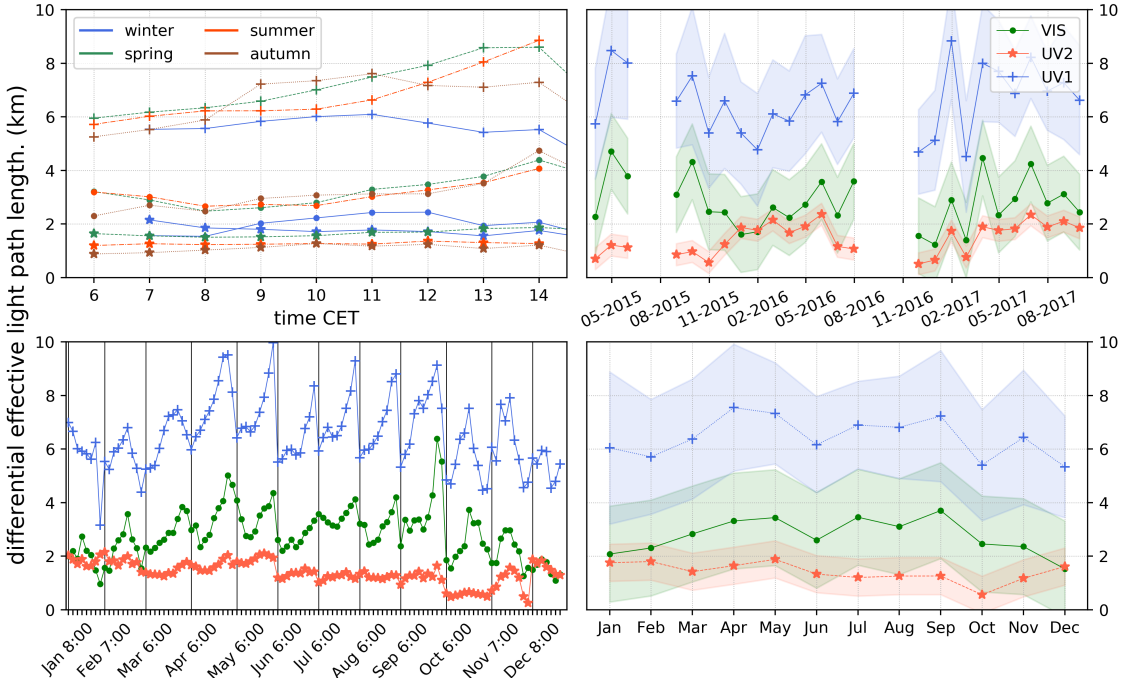


Fig. AF.35: *As above but for the differential effective light path and direction c.*

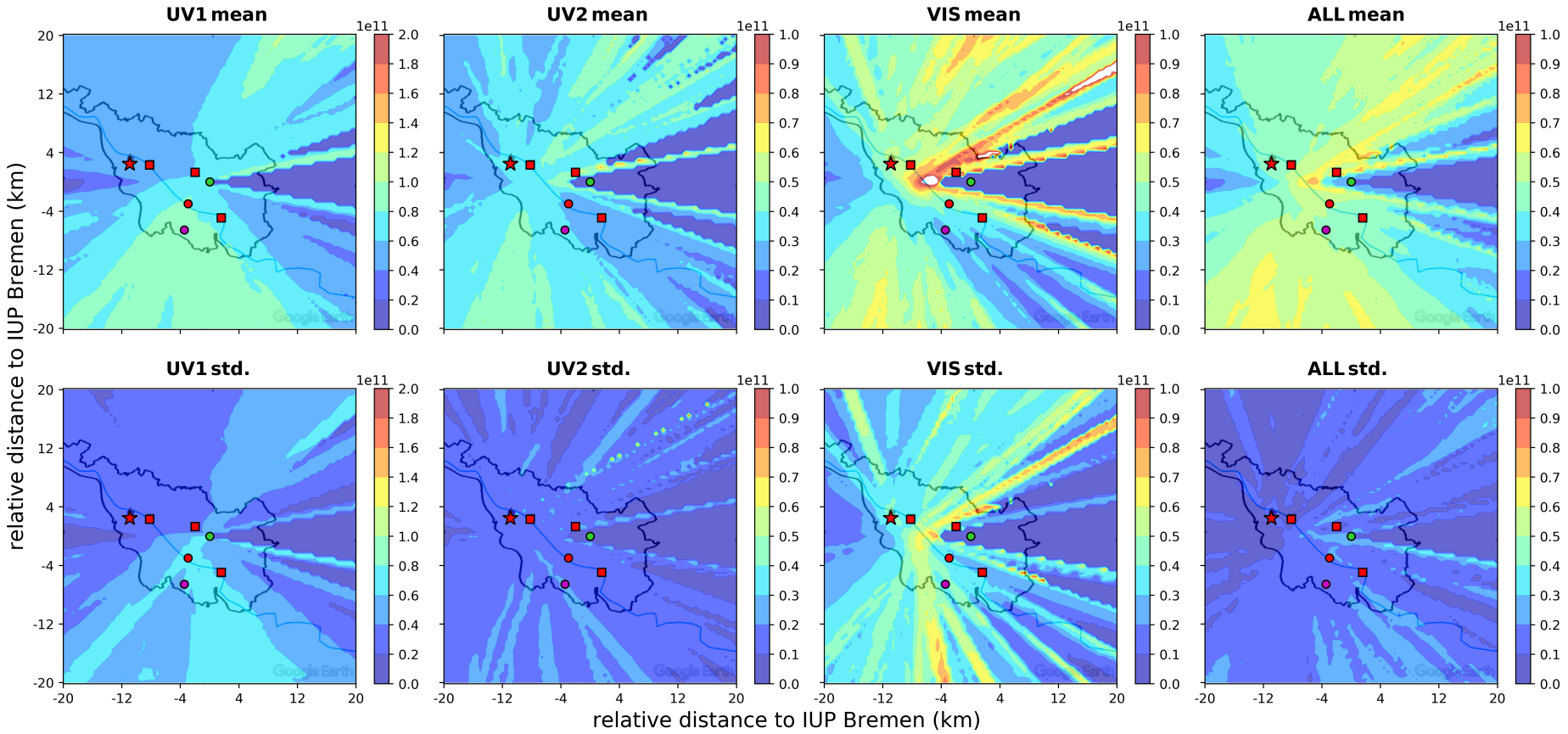


Fig. AF.36: *Mean NO_2 concentration maps of Bremen, with the concentration colour-coded in molec/cm^3. In the top row, three different fitting windows and the average of these three maps for is shown, for direction b. Matching standard deviations can be seen in the bottom row. Red squares depict power plants, the red star shows the steel manufacturing facility. The dots in green, red and purple represent the IUP, the city centre, and the DWD airport meteorology station, respectively.*

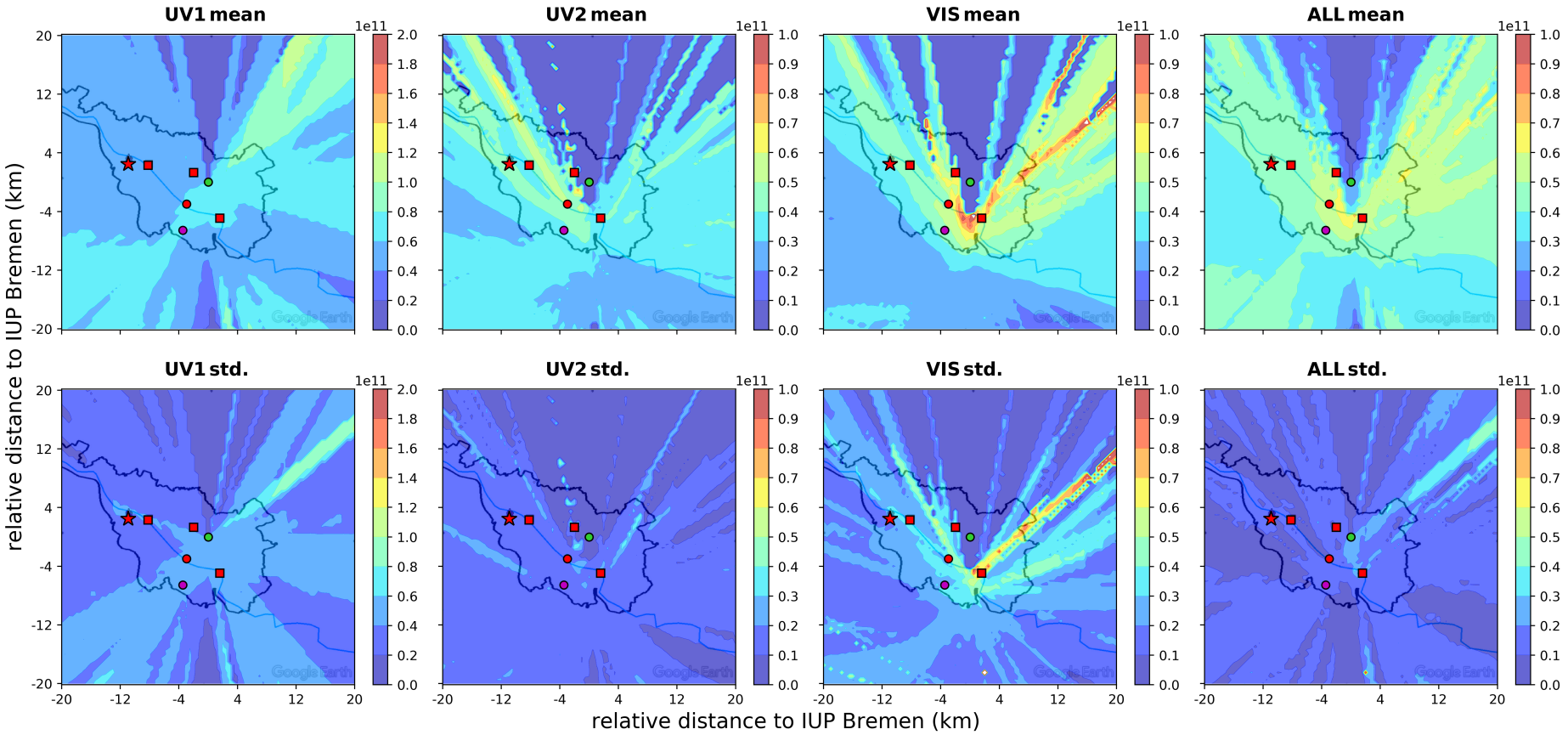


Fig. AF.37: *Mean NO_2 concentration maps of Bremen, with the concentration colour-coded in molec/cm^3. In the top row, three different fitting windows and the average of these three maps for is shown, for direction s. Matching standard deviations can be seen in the bottom row. Red squares depict power plants, the red star shows the steel manufacturing facility. The dots in green, red and purple represent the IUP, the city centre, and the DWD airport meteorology station, respectively.*

A.2 Additional material to Chapter 4

	abbrev.	AOT	scale height (km)	max. height (km)
Exponential	E1	0.2	1.0	6.0
	E2	0.6	1.0	6.0
	E3	1.0	1.0	6.0
	E4	0.6	0.6	6.0
	E5	0.6	1.4	6.0
Box	B1	0.4	-	0.5
	B2	0.4	-	1.0
	B3	0.4	-	2.0

Tab. AT.2: *Profile shape describing parameters for the aerosol extinction coefficient profiles (Exp, Box) used within the calculation of synthetic test datasets.*

	abbrev.	AOT	max. ext. (1/km)	width (km)	max. height (km)
max ext.	GA1	0.05	0.1	0.2	1.0
	GA2	0.10	0.2	0.2	1.0
	GA3	0.15	0.3	0.2	1.0
	GA4	0.20	0.4	0.2	1.0
width	GB1	0.05	0.2	0.1	1.0
	GB2	0.10	0.2	0.2	1.0
	GB3	0.15	0.2	0.3	1.0
	GB4	0.20	0.2	0.4	1.0
max. height	GC00	0.10	0.2	0.2	0.0
	GC05	0.18	0.2	0.2	0.5
	GC10	0.20	0.2	0.2	1.0
	GC15	0.20	0.2	0.2	1.5
	GC20	0.20	0.2	0.2	2.0
	GC25	0.20	0.2	0.2	2.5
	GC30	0.20	0.2	0.2	3.0
	GC35	0.20	0.2	0.2	3.5
	GC40	0.20	0.2	0.2	4.0

Tab. AT.3: *Profile shape describing parameters for the aerosol extinction coefficient profiles (Gaussian) used within the calculation of synthetic test datasets. Vertically written words describe the test quantity of this dataset.*

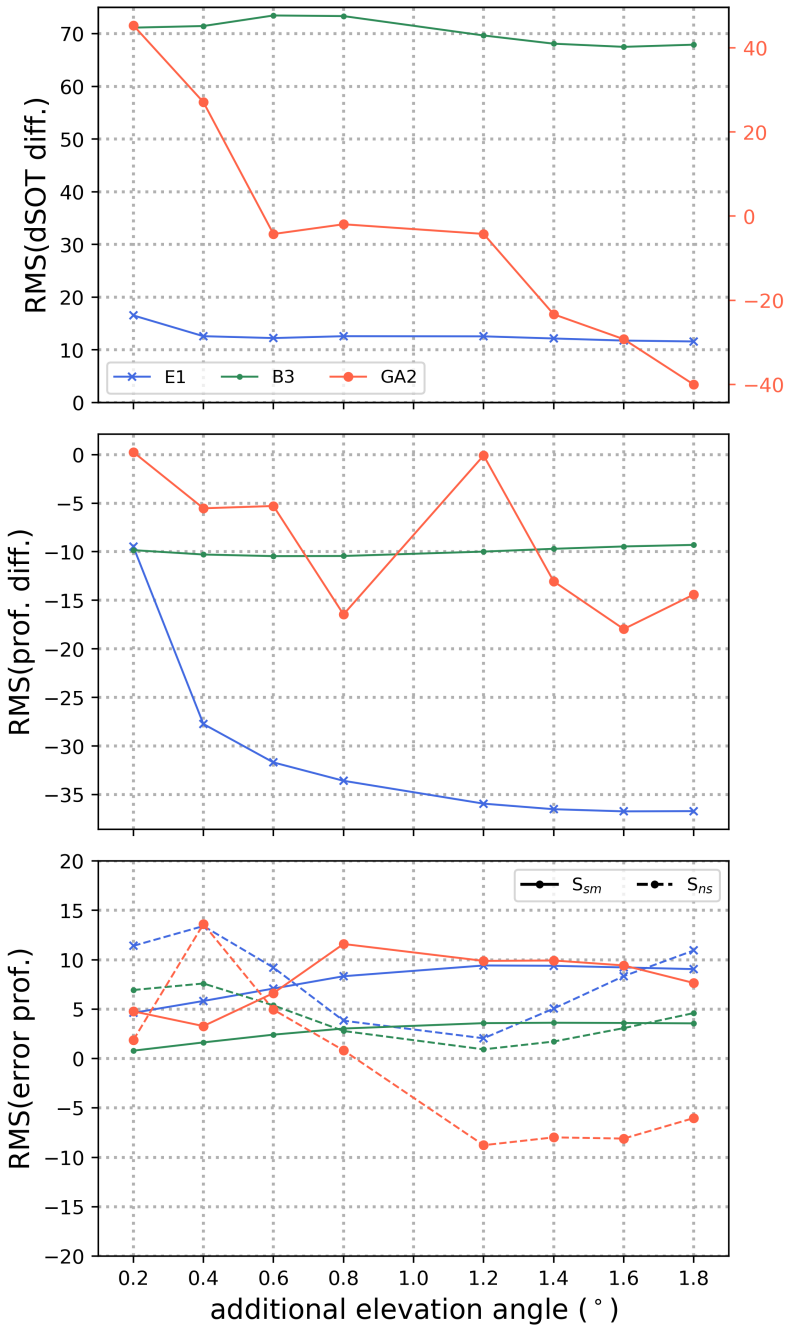


Fig. AF.38: *Relative differences of the specific quantity without 0° but with an additional LOS (additional elevation angle) to the quantity with neglection of 0° only, for three scenarios and for LOS11. Relative difference (in %) of, Top: RMS of the difference of retrieved and simulated dSOT; Mid row: RMS of the difference of the true and retrieved profile; Bottom row: RMS of the error profiles.*

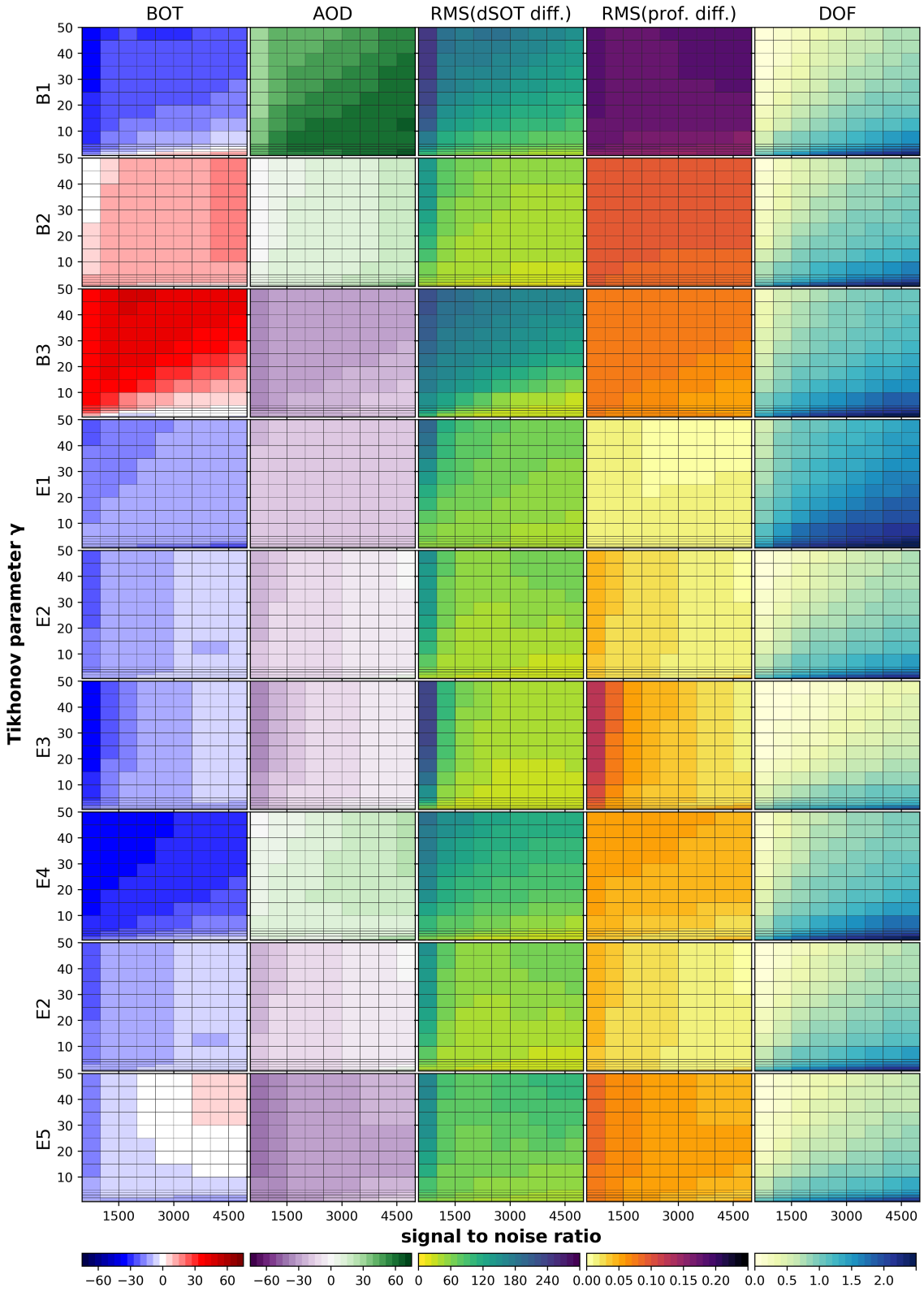


Fig. AF.39: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOD (in %).* ***3rd, 4th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: degrees of freedom.* ***Retrieval mode****: apri, variance(z), pre scaling (see Tab. 4.4)*

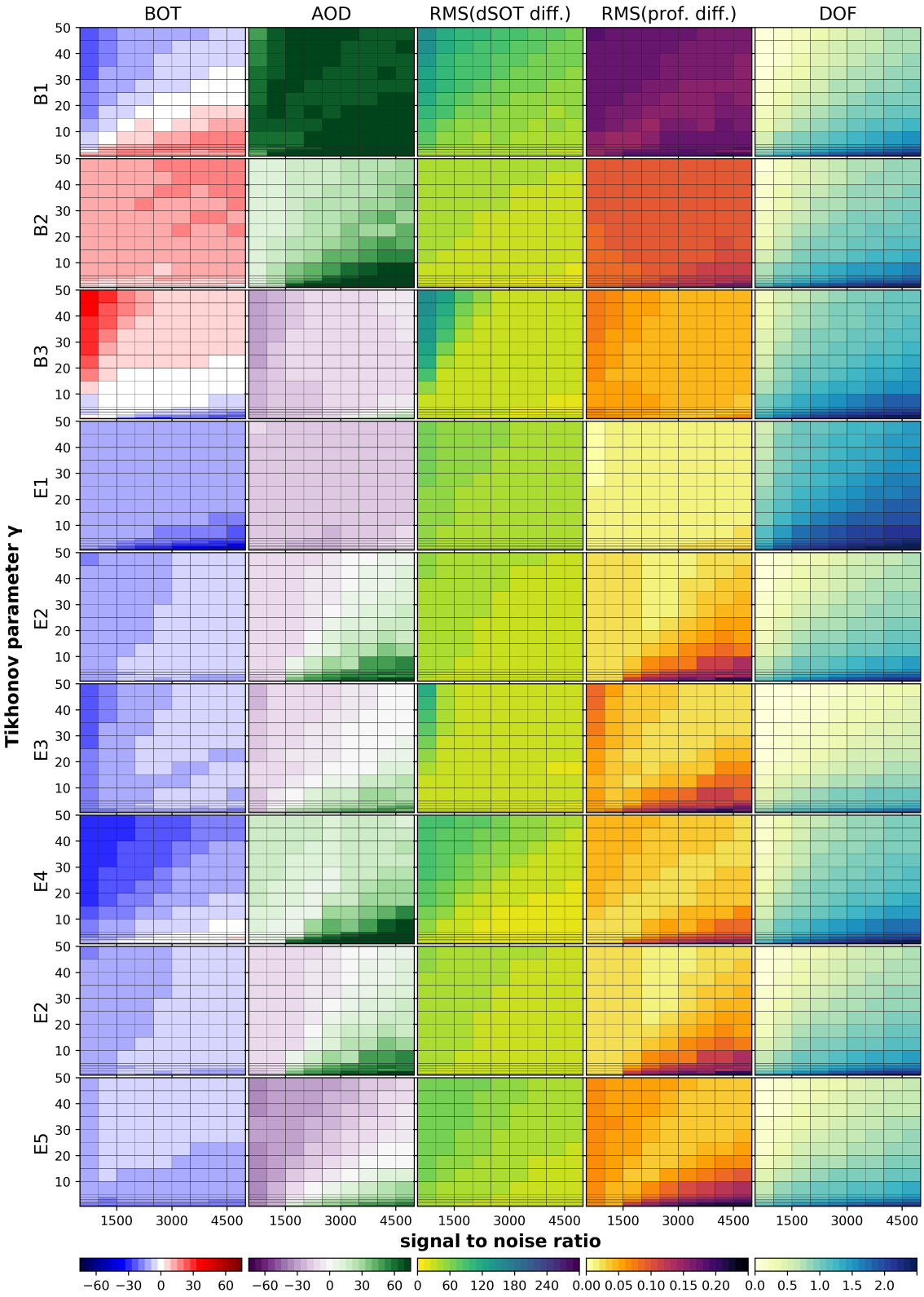


Fig. AF.40: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOD (in %).* ***3rd, 4th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: degrees of freedom.* ***Retrieval mode****: prev., variance(z)., pre scaling (see Tab. 4.4)*

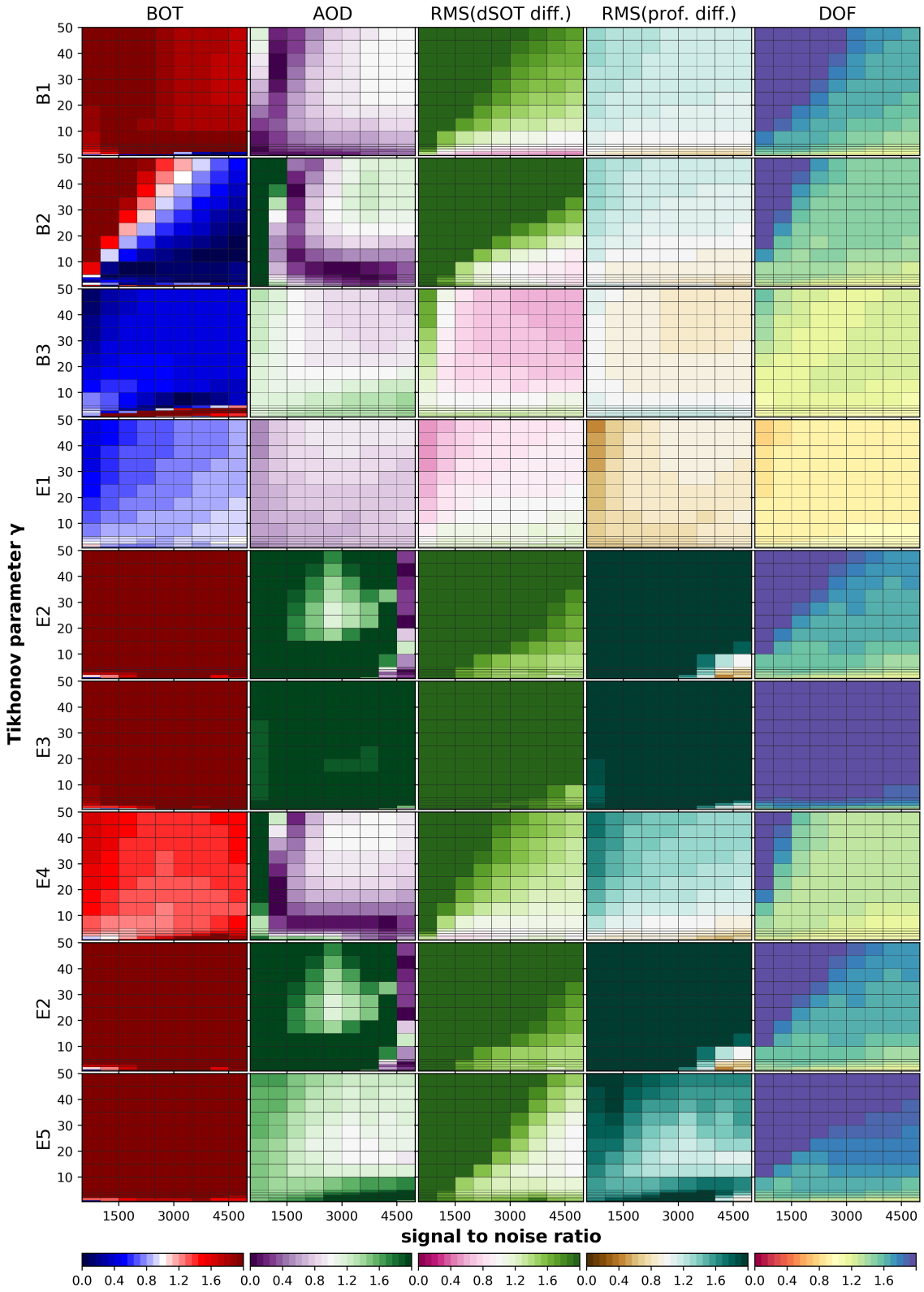


Fig. AF.41: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Ratio of absolutes of the relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOT (in %).* ***3rd, 4th column****: Ratio of the RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: Ratio of the degrees of freedom.* ***Ratios for retrieval modes****: apri, const., fixed a priori profile to apri, const., pre scaling (see Tab. 4.4)*

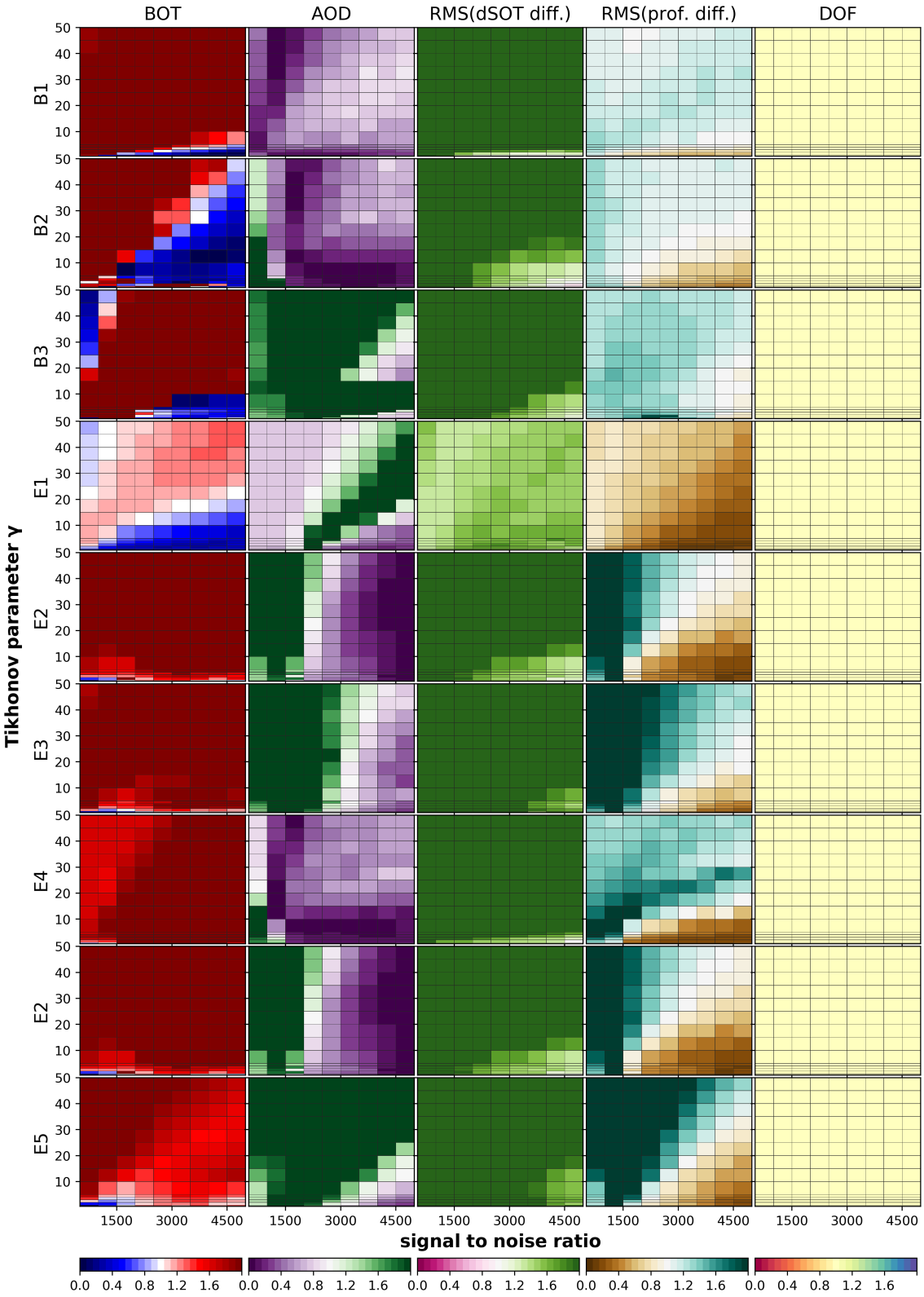


Fig. AF.42: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Ratio of absolutes of the relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOT (in %).* ***3rd, 4th column****: Ratio of the RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: Ratio of the degrees of freedom.* ***Ratios for retrieval modes****: apri, const., fixed a priori profile to prev, const., fixed a priori profile (see Tab. 4.4)*

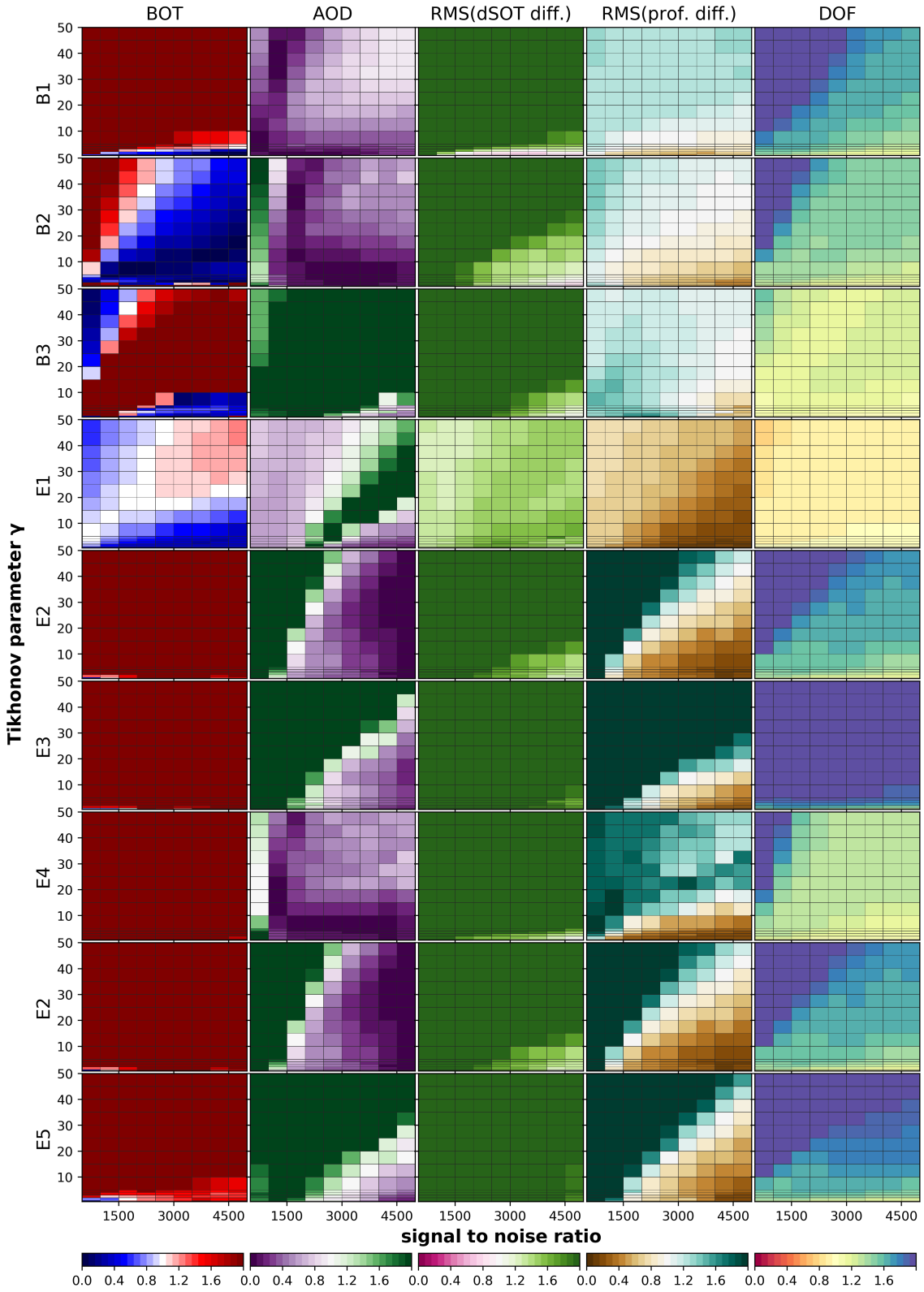


Fig. AF.43: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Ratio of absolutes of the relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOT (in %).* ***3rd, 4th column****: Ratio of the RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: Ratio of the degrees of freedom.* ***Ratios for retrieval modes****: apri, const., fixed a priori profile to prev, const., pre scaling (see Tab. 4.4)*

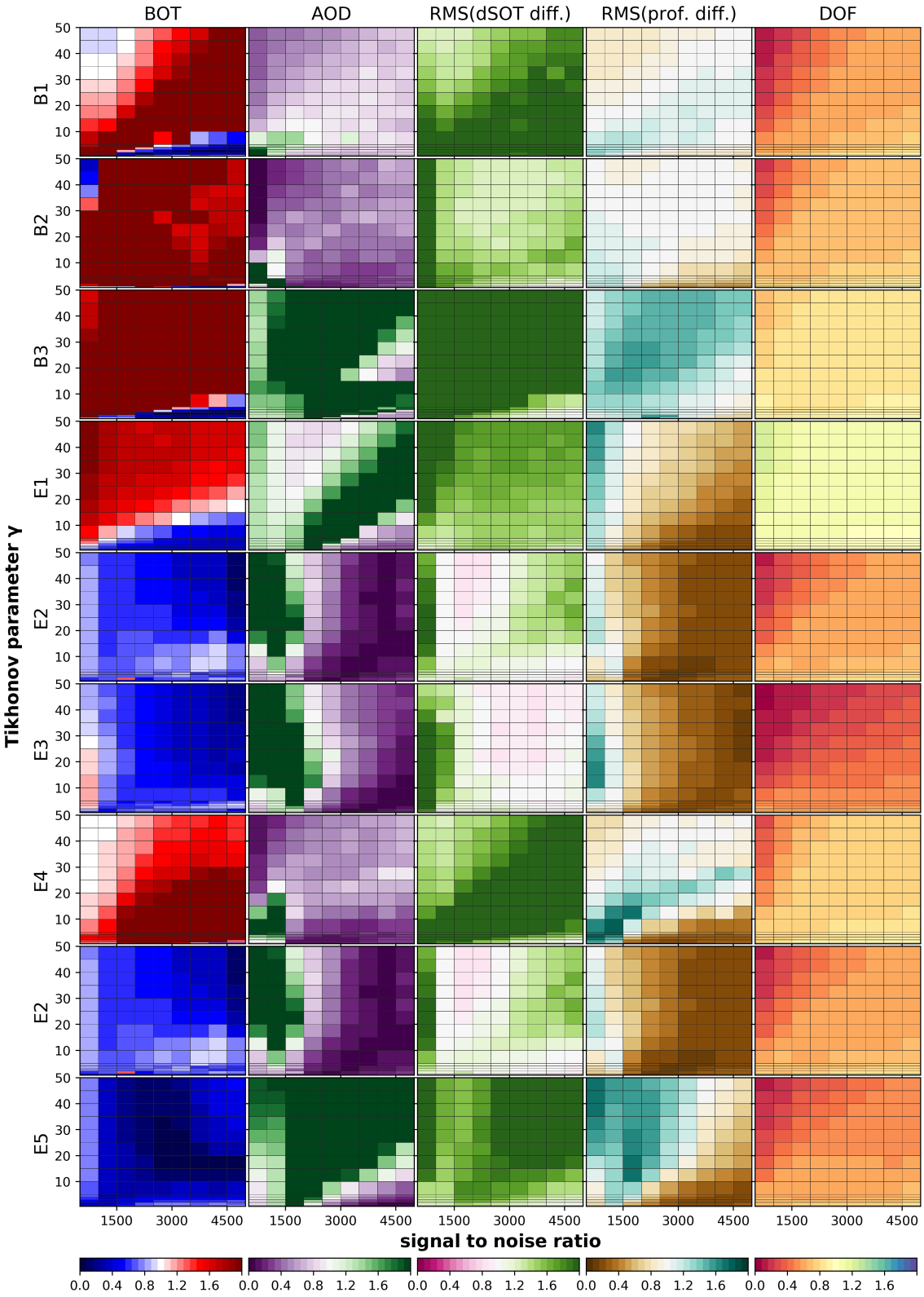


Fig. AF.44: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Ratio of absolutes of the relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOT (in %).* ***3rd, 4th column****: Ratio of the RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: Ratio of the degrees of freedom.* ***Ratios for retrieval modes****: apri, const., pre scaling to prev, const., fixed a priori profile (see Tab 4.4)*

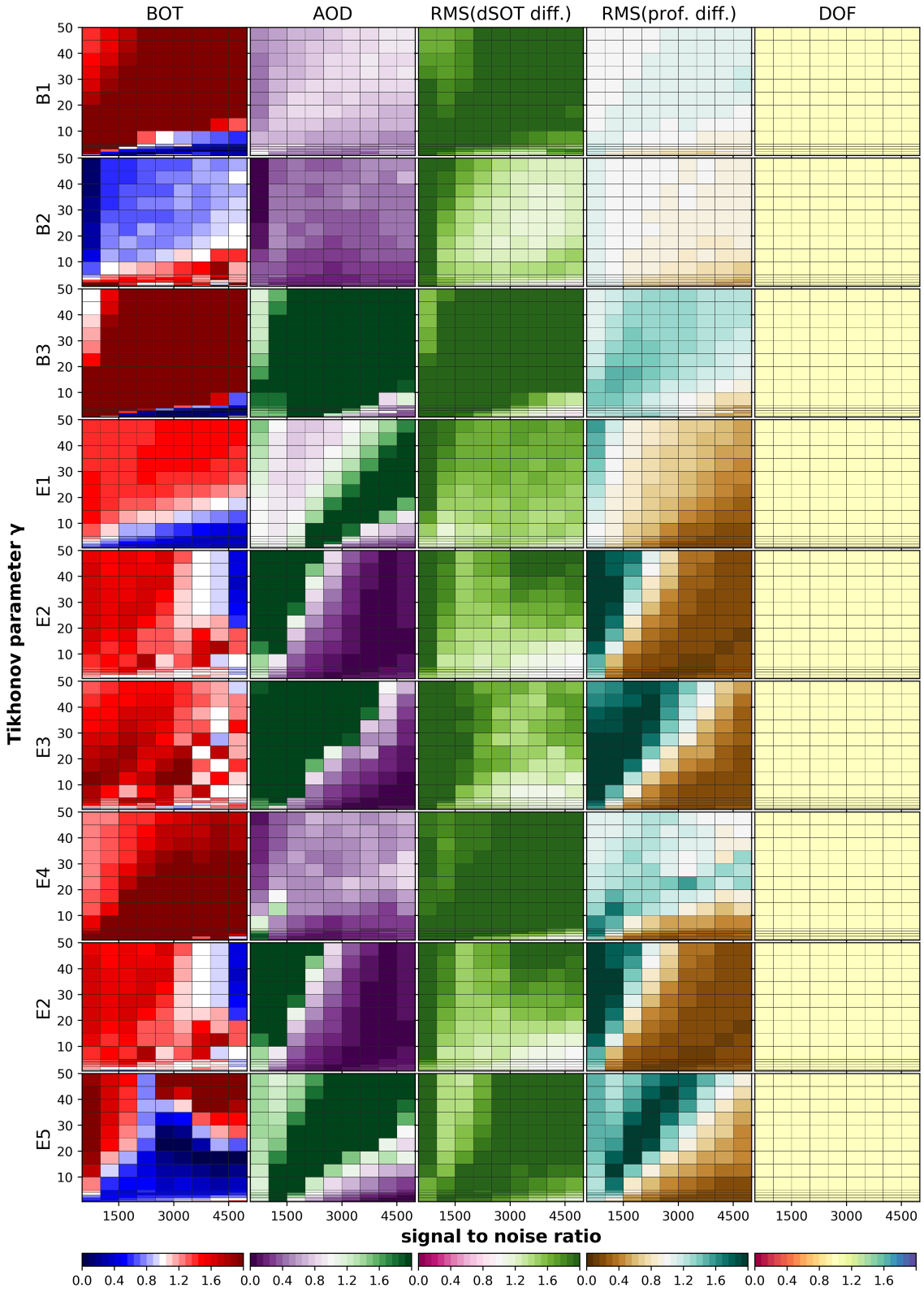


Fig. AF.45: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Ratio of absolutes of the relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOT (in %).* ***3rd, 4th column****: Ratio of the RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: Ratio of the degrees of freedom.* ***Ratios for retrieval modes****: apri, const., pre scaling to prev, const., pre scaling (see Tab 4.4)*

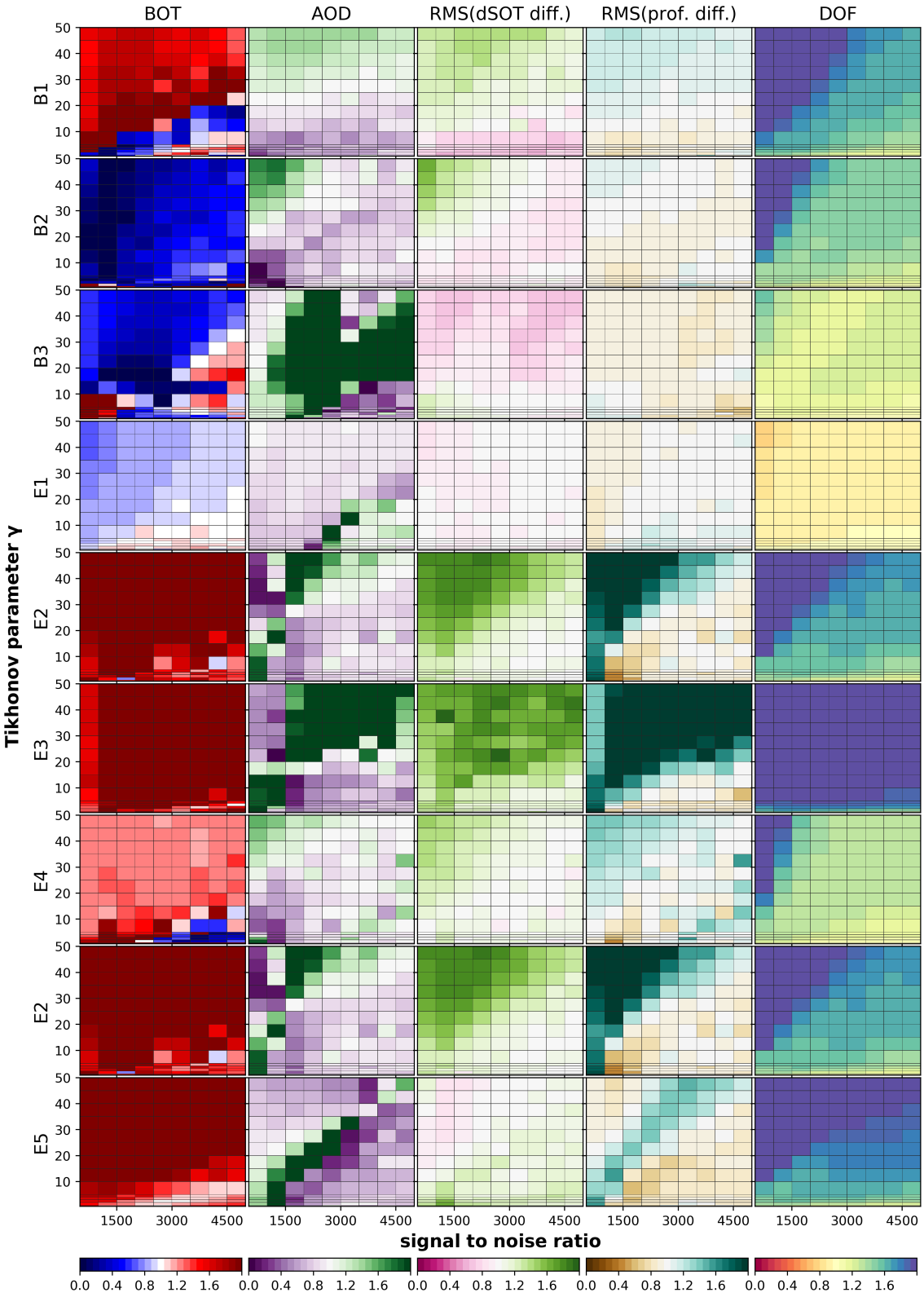


Fig. AF.46: *Variation of SNR and γ for three box (B1 - B3) and 5 exp. profiles (E1 - E5), with resulting parameters colour-coded.* ***1st, 2nd column****: Ratio of absolutes of the relative difference of, retrieved to true bottom extinction coefficient value (in %), retrieved to true AOT (in %).* ***3rd, 4th column****: Ratio of the RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***5th column****: Ratio of the degrees of freedom.* ***Ratios for retrieval modes****: prev, const., fixed a priori profile to prev, const., pre scaling (see Tab. 4.4)*

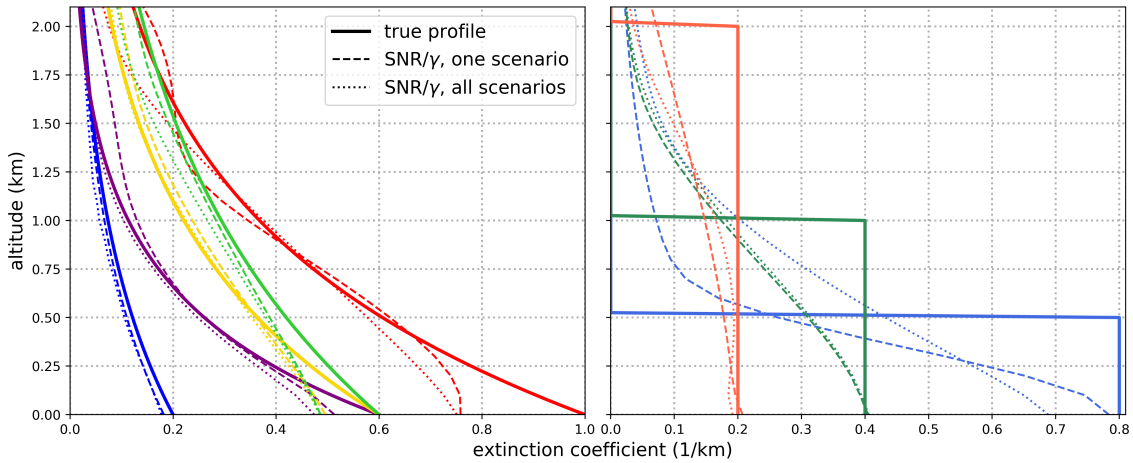


Fig. AF.47: *Profiles retrieved with the optimal regularization parameters for each individual scenario (dashed lines) and for the mean optimal SNR and γ values of all scenarios (pointed lines).* ***Retrieval mode****: apri, const., fixed a priori profile (see Table 4.4)*

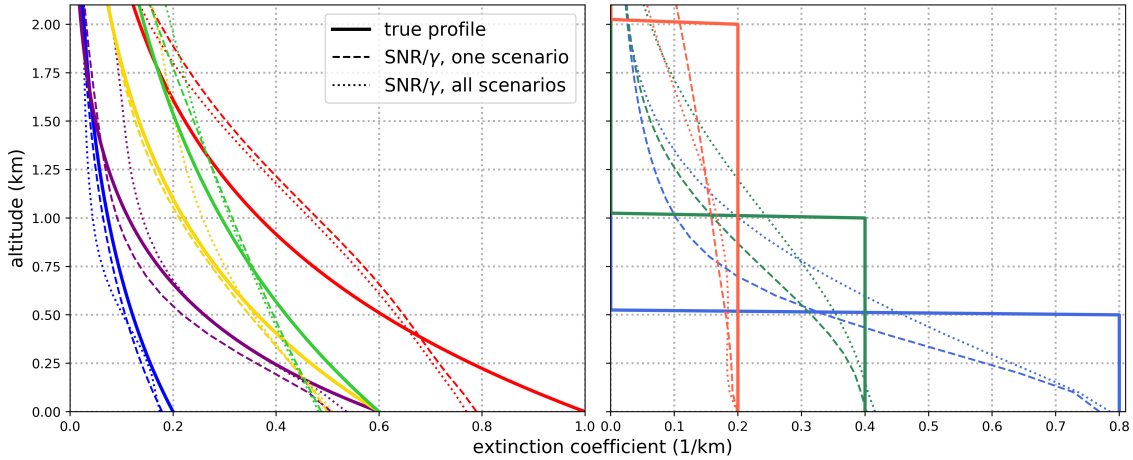


Fig. AF.48: *Profiles retrieved with the optimal regularization parameters for each individual scenario (dashed lines) and for the mean optimal SNR and γ values of all scenarios (pointed lines).* ***Retrieval mode****: prev, const., fixed a priori profile (see Table 4.4)*

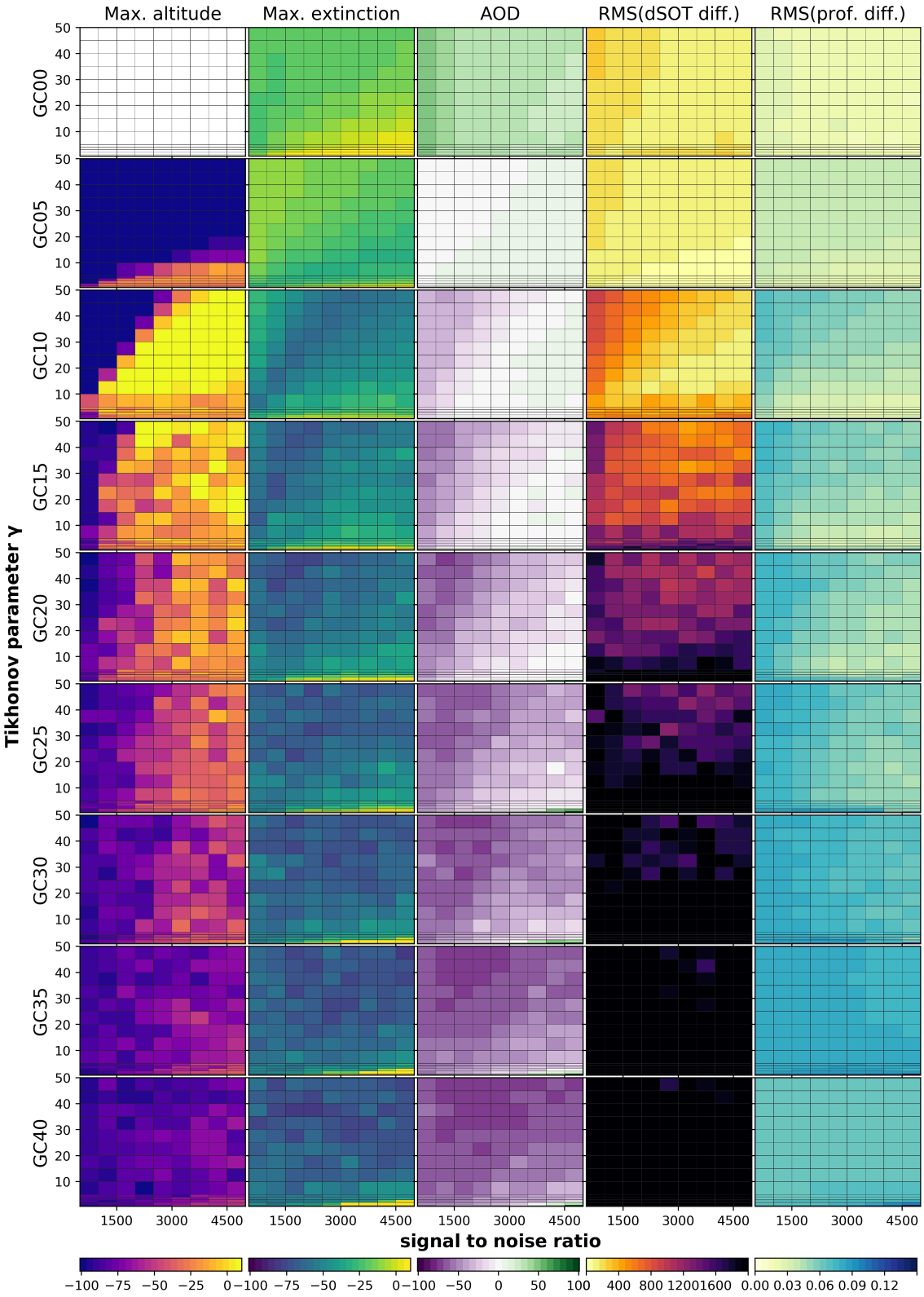


Fig. AF.49: *Variation of SNR and γ for the Gaussian shaped profiles (GC00 - GC40), with resulting parameters colour-coded.* ***1st, 2nd, 3rd column****: Relative difference of, retrieved to true maximum peak altitude (in %), retrieved to true maximum extinction (in %), retrieved to true AOD (in %).* ***4th, 5th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***Retrieval mode****: apri, const., fixed a priori profile (see Tab. 4.4)*

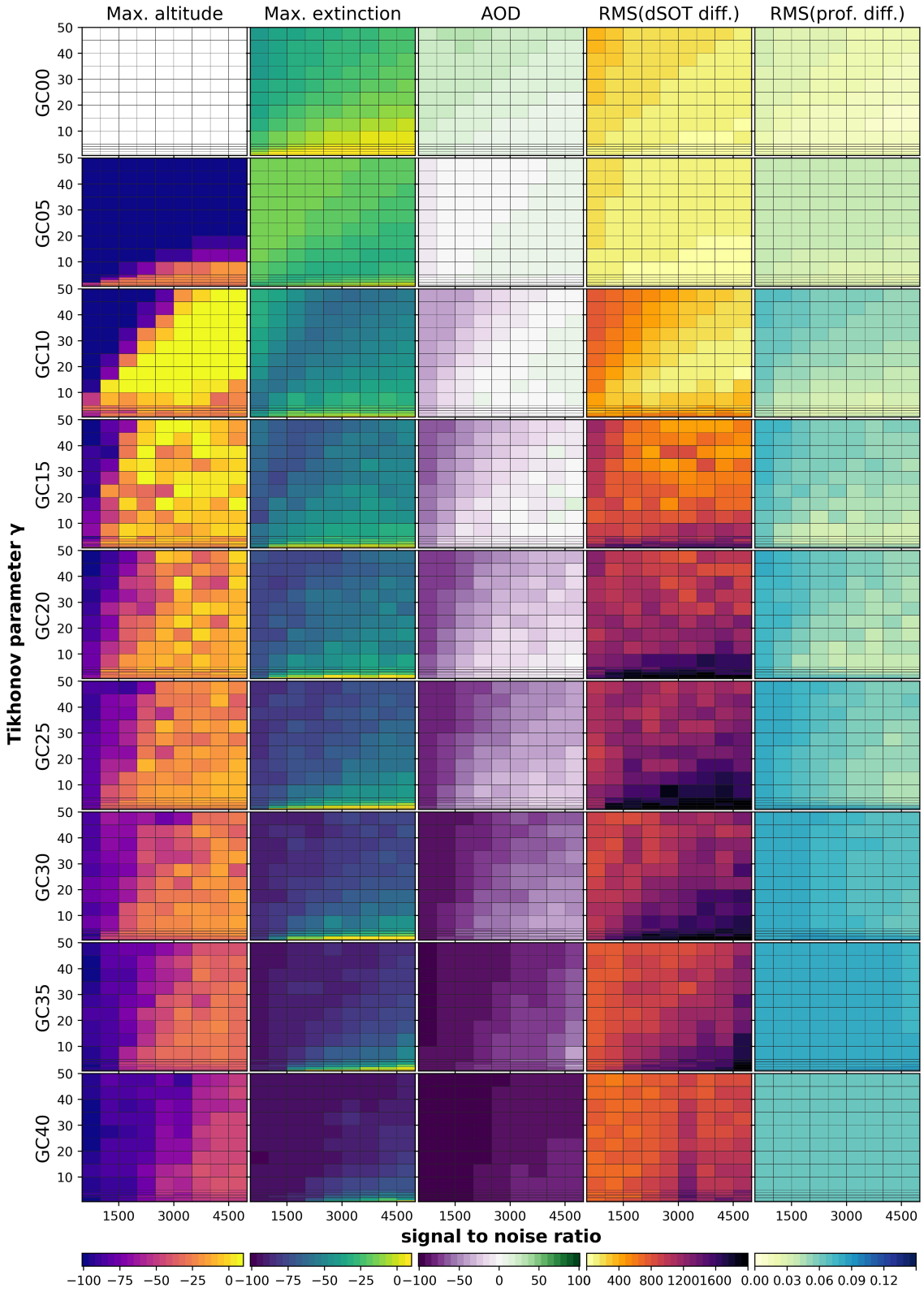


Fig. AF.50: *Variation of SNR and γ for the Gaussian shaped profiles (GC00 - GC40), with resulting parameters colour-coded.* ***1st, 2nd, 3rd column****: Relative difference of, retrieved to true maximum peak altitude (in %), retrieved to true maximum extinction (in %), retrieved to true AOD (in %).* ***4th, 5th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***Retrieval mode****: apri, const., pre scaling (see Tab. 4.4)*

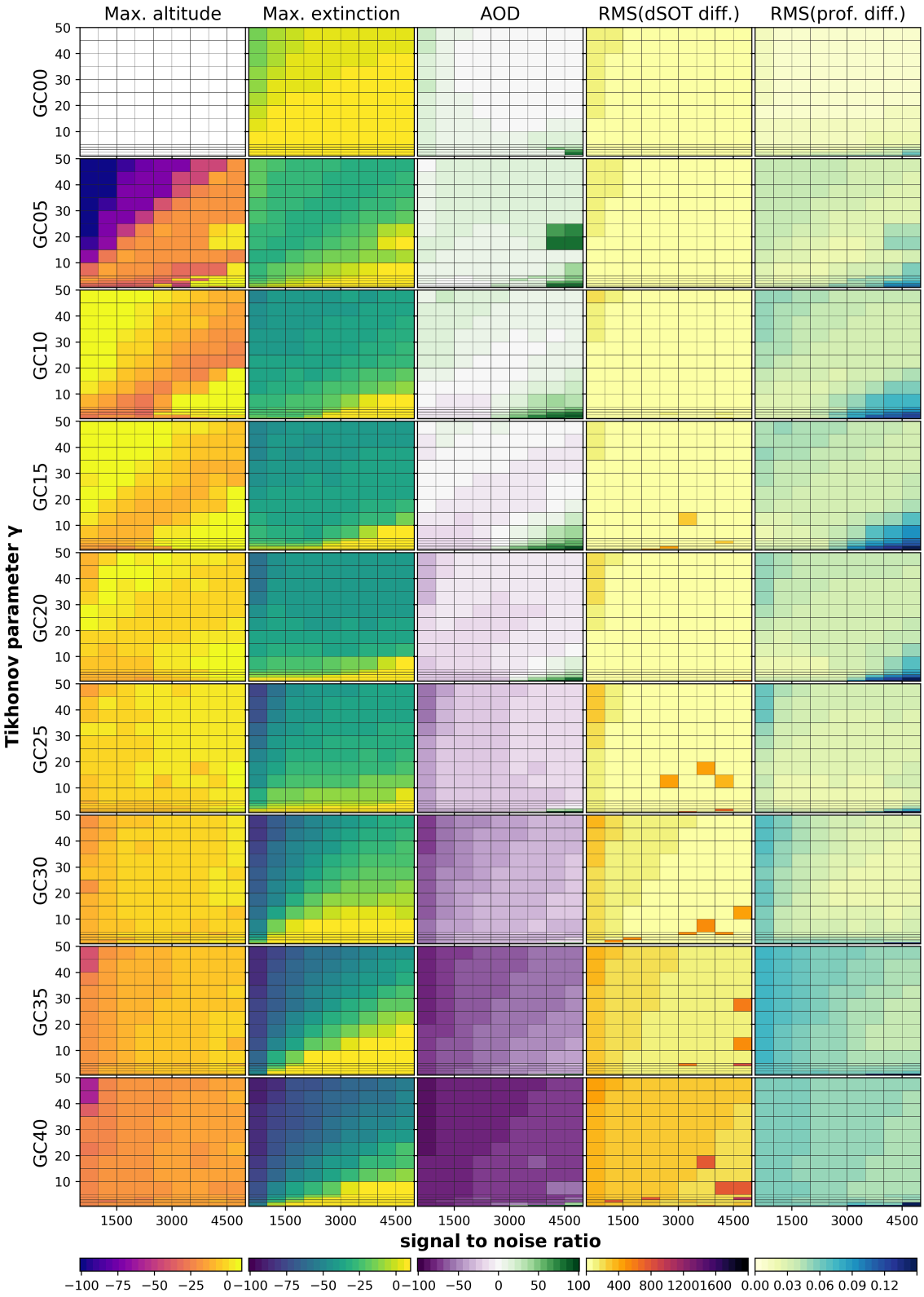


Fig. AF.51: *Variation of SNR and γ for the Gaussian shaped profiles (GC00 - GC40), with resulting parameters colour-coded.* ***1st, 2nd, 3rd column****: Relative difference of, retrieved to true maximum peak altitude (in %), retrieved to true maximum extinction (in %), retrieved to true AOD (in %).* ***4th, 5th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***Retrieval mode****: prev., const., fixed a priori profile (see Tab. 4.4)*

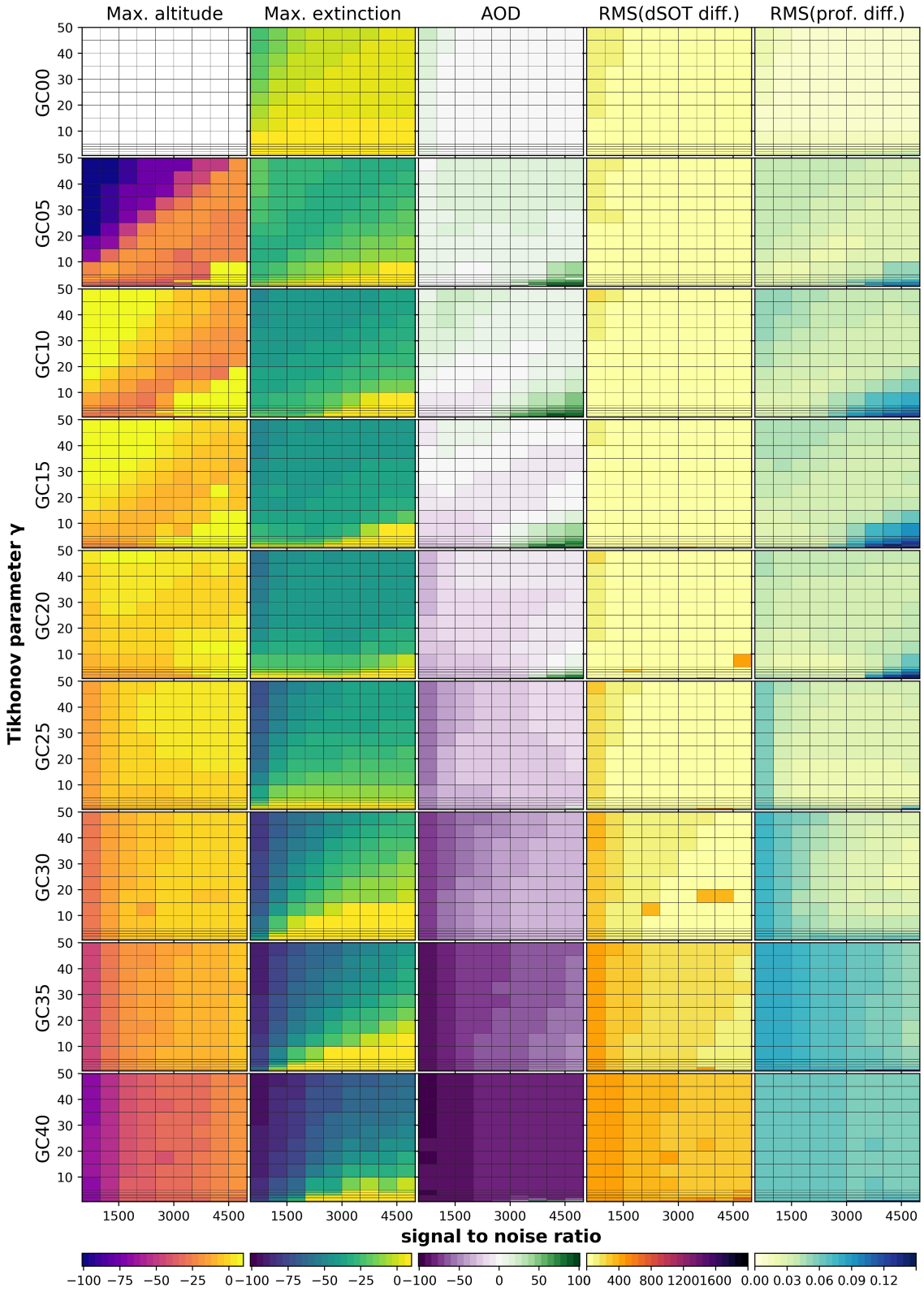


Fig. AF.52: *Variation of SNR and γ for the Gaussian shaped profiles (GC00 - GC40), with resulting parameters colour-coded.* ***1st, 2nd, 3rd column****: Relative difference of, retrieved to true maximum peak altitude (in %), retrieved to true maximum extinction (in %), retrieved to true AOD (in %).* ***4th, 5th column****: RMS of, the difference of retrieved and true dSOT, the difference of retrieved and true profile (in 1/km).* ***Retrieval mode****: prev., const., pre scaling (see Tab. 4.4)*

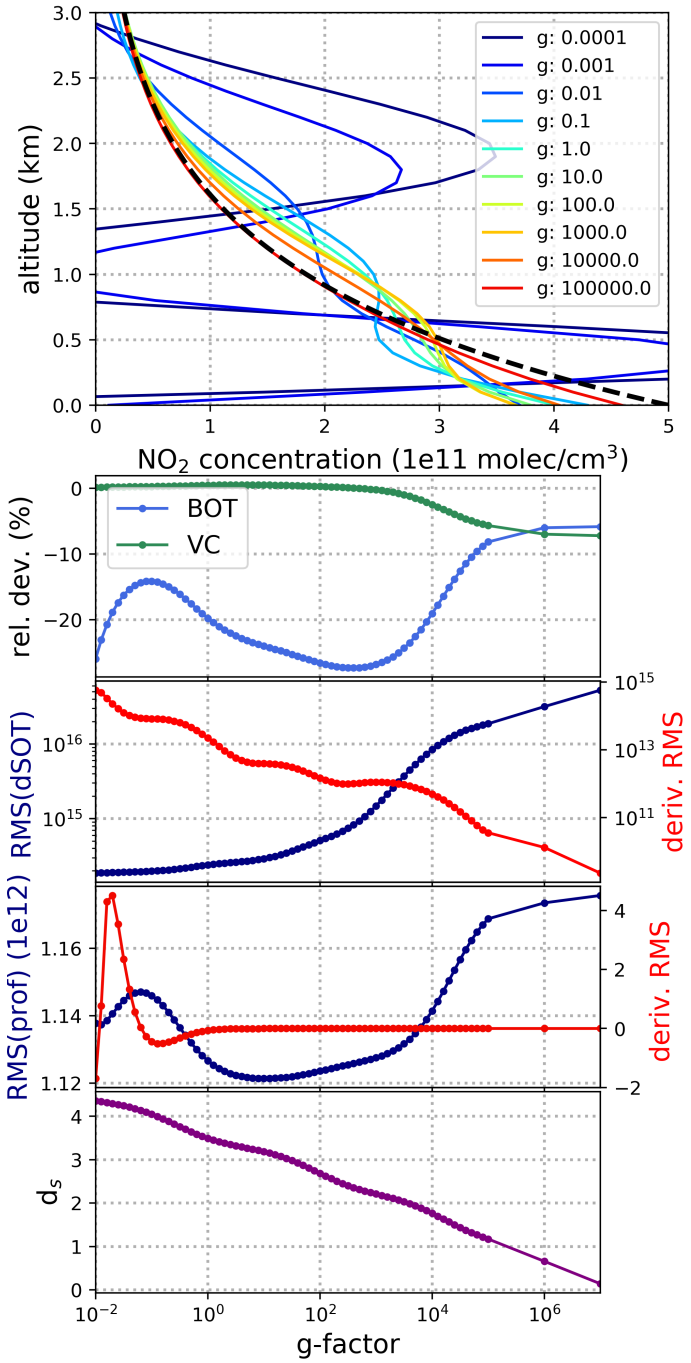


Fig. AF.53: *Different quantities retrieved with varying g-factors.* ***Top****: profiles retrieved with pre-scaled a priori profile.* ***2nd row****: relative deviations of retrieved to true bottom concentration (blue) and vertical column density (green), in %.* ***3rd row****: RMS of retrieved and true dSOT differences (blue) and its derivative (red).* ***4th row****: RMS of retrieved and true profile (blue) and its derivative (red).* ***Bottom row****: degrees of freedom.*

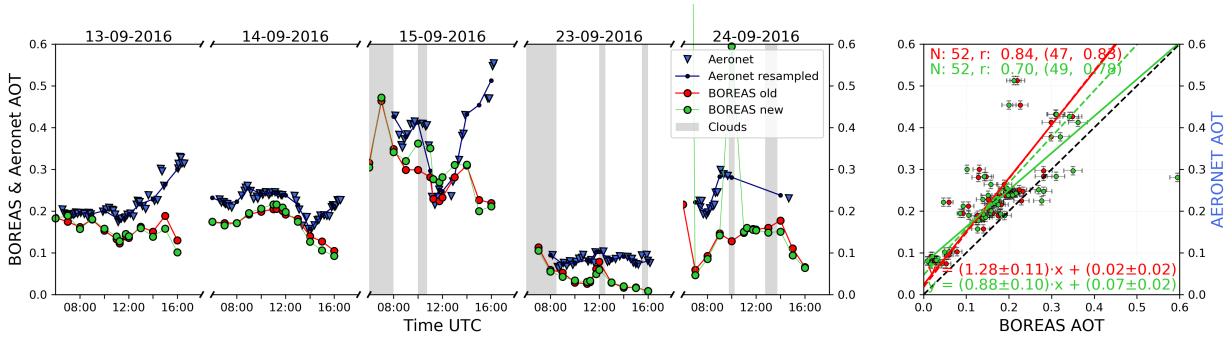


Fig. AF.54: *Left: time series of AERONET (blue) and BOREAS AOT (red: paper data, green: new data with AERONET phase fct and SSA). Small triangles show the original AERONET measurement, small dots with connecting lines depict the resampled data. Grey areas indicate clouds. Right: scatter plot of both datasets including parameters of the orthogonal regression and Pearsons correlation coefficient. The numbers outside (inside) of the brackets show correlations for unfiltered (filtered) data. The solid (dashed) lines show the regression lines of the unfiltered (filtered) datasets.*

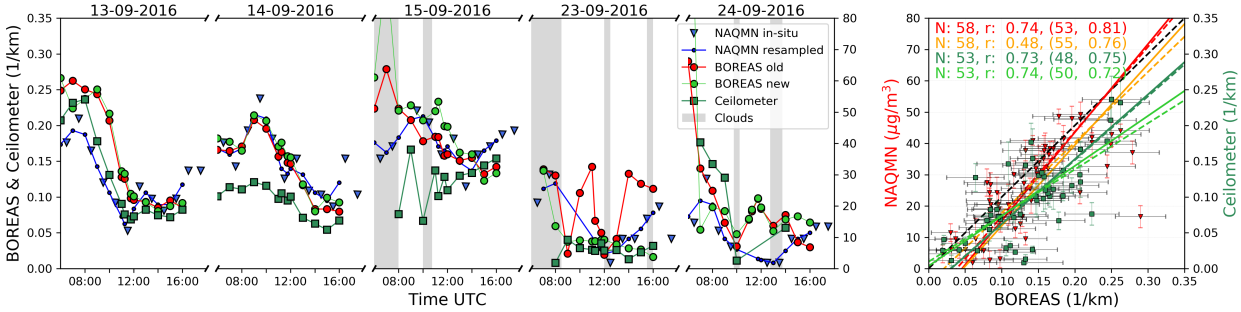


Fig. AF.55: *Left: time series of BOREAS (red: paper data, bright green dots: new data with AERONET phase fct and SSA), ceilometer (dark green squares) and NAQMN in situ (blue) and near surface aerosol parameters. Small triangles show the original hourly NAQMN measurements, small dots with connecting lines depict the resampled data. Green squares show ceilometer near-surface extinction values evaluated by averaging the 10 - 50 m coefficients. Right: as in Figure above. Regression parameters can be found in Tab. AT.4.*

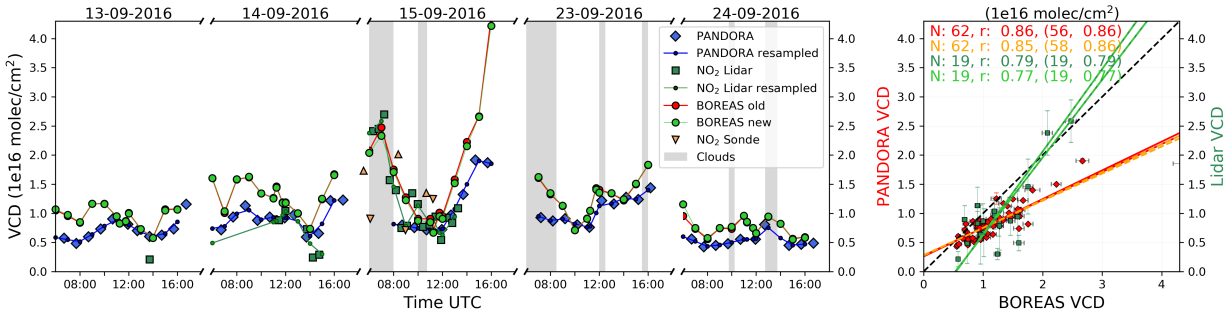


Fig. AF.56: *Left: time series of Pandora (blue diamonds), LIDAR (green squares) and BOREAS NO_2-VCD (red: paper data, bright green dots: new data with AERONET phase fct and SSA). Small dots with connecting lines depict the resampled datasets of Pandora and LIDAR. Orange triangles are the representation of the integrated NO_2 sondes measurements with the ascend and descend separated in triangles with the edge to the top or to the bottom, respectively. Right: as in Figure above.*

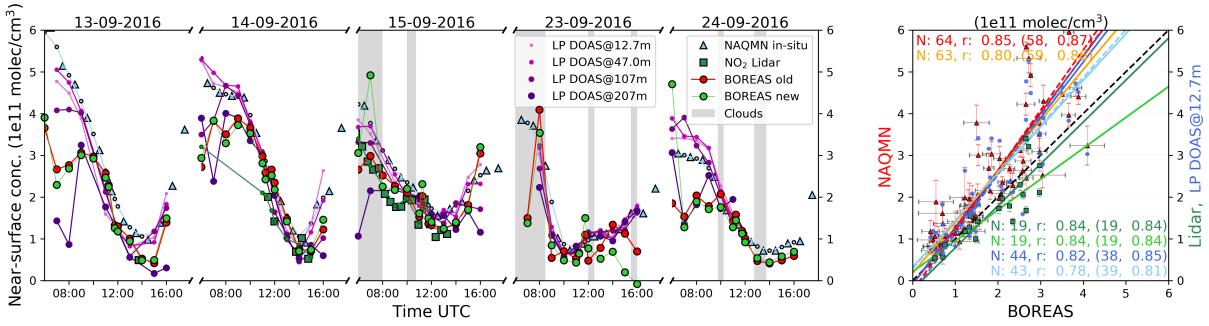


Fig. AF.57: *Left: time series of in situ and remote sensing NO_2 near surface concentrations. Triangles show in situ instruments for NAQMN as bright blue triangles. The NO_2 LIDAR is plotted as green squares and the LP DOAS as different sized circles in shades of magenta (lowest altitude as smallest circle). BOREAS is shown as red (paper) and bright green dots (new data with AERONET phase fct and SSA). Right: as in Figure above.*

AOT

	N	r	m	b
Paper - AERONET (filtered)	47	0.83	1.295 ± 0.119	0.018 ± 0.022
New settings - AERONET (filtered)	49	0.78	1.103 ± 0.112	0.047 ± 0.021
Paper - AERONET (unfiltered)	52	0.84	1.276 ± 0.106	0.034 ± 0.018
New settings - AERONET (unfiltered)	52	0.70	0.883 ± 0.104	0.074 ± 0.021

surface extinction

	N	r	m	b
Paper - NAQMN (filtered)	53	0.81	270.7 ± 27.9	-10.4 ± 4.1
New settings - NAQMN (filtered)	55	0.76	225.2 ± 26.1	-4.7 ± 3.9
Paper - NAQMN (unfiltered)	58	0.74	282.0 ± 34.2	-12.9 ± 5.0
New settings - NAQMN (unfiltered)	58	0.48	258.1 ± 63.5	-12.2 ± 10.0
Paper - Ceilometer (filtered)	48	0.75	0.890 ± 0.099	-0.027 ± 0.015
New settings - Ceilometer (filtered)	50	0.72	0.645 ± 0.076	0.010 ± 0.011
Paper - Ceilometer (unfiltered)	53	0.73	0.904 ± 0.100	-0.028 ± 0.015
New settings - Ceilometer (unfiltered)	53	0.74	0.699 ± 0.076	0.004 ± 0.011

Tab. AT.4: *Regression parameters and Pearsons correlations coefficients for the paper results and the profiles retrieved with AERONET phase function and SSA shown in the Figures AF.54 - AF.55. The regression equation follows: $y = mx + b$.*

VCD				
	N	**r**	**m**	**b**
Paper - Pandora (filtered)	56	0.86	0.466 ± 0.036	0.284 ± 0.047
New settings - Pandora (filtered)	58	0.86	0.479 ± 0.036	0.273 ± 0.046
Paper - Pandora (unfiltered)	62	0.86	0.479 ± 0.036	0.273 ± 0.046
New settings - Pandora (unfiltered)	62	0.85	0.493 ± 0.037	0.258 ± 0.048
Paper - Lidar (filtered)	19	0.79	1.335 ± 0.223	-0.711 ± 0.297
New settings - Lidar (filtered)	19	0.77	1.416 ± 0.250	-0.771 ± 0.323
Paper - Lidar (unfiltered)	19	0.79	1.335 ± 0.223	-0.711 ± 0.297
New settings - Lidar (unfiltered)	19	0.77	1.416 ± 0.250	-0.771 ± 0.323
surface concentration				
	N	**r**	**m**	**b**
Paper - NAQMN (filtered)	58	0.87	1.442 ± 0.101	-0.245 ± 0.206
New settings - NAQMN (filtered)	59	0.86	1.219 ± 0.090	0.175 ± 0.182
Paper - NAQMN (unfiltered)	64	0.85	1.367 ± 0.100	-0.093 ± 0.202
New settings - NAQMN (unfiltered)	63	0.80	1.219 ± 0.103	0.178 ± 0.213
Paper - Lidar (filtered)	19	0.84	1.004 ± 0.142	-0.220 ± 0.294
New settings - Lidar (filtered)	19	0.84	0.739 ± 0.106	0.210 ± 0.241
Paper - Lidar (unfiltered)	19	0.84	1.004 ± 0.142	-0.220 ± 0.294
New settings - Lidar (unfiltered)	19	0.84	0.739 ± 0.106	0.210 ± 0.241
Paper - LP DOAS (filtered)	38	0.82	1.425 ± 0.135	-0.312 ± 0.290
New settings - LP DOAS (filtered)	39	0.78	1.145 ± 0.121	0.284 ± 0.256
Paper - LP DOAS (unfiltered)	44	0.85	1.333 ± 0.132	-0.102 ± 0.280
New settings - LP DOAS (unfiltered)	43	0.81	1.118 ± 0.122	0.323 ± 0.266

Tab. AT.5: *Regression parameters and Pearsons correlations coefficients for the paper results and the profiles retrieved with AERONET phase function and SSA shown in the Figures AF.56 - AF.57. The regression equation follows:* $y = mx + b$.

AOT				
	N	**r**	**m**	**b**
Paper (filtered)	47	0.83	1.295 ± 0.119	0.018 ± 0.022
New settings (filtered)	49	0.78	0.924 ± 0.092	0.058 ± 0.019
Paper (unfiltered)	52	0.84	1.276 ± 0.106	0.034 ± 0.018
New settings (unfiltered)	52	0.70	0.920 ± 0.086	0.060 ± 0.018
surface extinction				
	N	**r**	**m**	**b**
Paper - NAQMN (filtered)	53	0.81	270.7 ± 27.9	-10.4 ± 4.1
New settings - NAQMN (filtered)	55	0.79	229.4 ± 24.1	-5.4 ± 3.7
Paper - NAQMN (unfiltered)	58	0.74	282.0 ± 34.2	-12.9 ± 5.0
New settings - NAQMN (unfiltered)	58	0.78	214.8 ± 22.8	-3.7 ± 3.6
Paper - Ceilometer (filtered)	48	0.75	0.890 ± 0.099	-0.027 ± 0.015
New settings - Ceilometer (filtered)	51	0.75	0.613 ± 0.068	0.014 ± 0.011
Paper - Ceilometer (unfiltered)	53	0.73	0.904 ± 0.100	-0.028 ± 0.015
New settings - Ceilometer (unfiltered)	53	0.77	0.623 ± 0.065	0.012 ± 0.010

Tab. AT.6: *Regression parameters and Pearsons correlations coefficients for the paper results and the profiles retrieved with new settings and the AERONET phase function and SSA shown in the Figures 4.44 - 4.45. The regression equation follows:* $y = mx + b$.

VCD				
	N	**r**	**m**	**b**
Paper - Pandora (filtered)	56	0.86	0.466 ± 0.036	0.284 ± 0.047
New settings - Pandora (filtered)	60	0.87	0.479 ± 0.036	0.273 ± 0.046
Paper - Pandora (unfiltered)	62	0.86	0.479 ± 0.036	0.273 ± 0.046
New settings - Pandora (unfiltered)	62	0.84	0.507 ± 0.040	0.237 ± 0.051
Paper - Lidar (filtered)	19	0.79	1.335 ± 0.223	-0.711 ± 0.297
New settings - Lidar (filtered)	18	0.68	1.424 ± 0.318	-0.798 ± 0.387
Paper - Lidar (unfiltered)	19	0.79	1.335 ± 0.223	-0.711 ± 0.297
New settings - Lidar (unfiltered)	19	0.79	1.495 ± 0.251	-0.876 ± 0.325
surface concentration				
	N	**r**	**m**	**b**
Paper - NAQMN (filtered)	58	0.87	1.442 ± 0.101	-0.245 ± 0.206
New settings - NAQMN (filtered)	60	0.82	1.374 ± 0.113	0.000 ± 0.222
Paper - NAQMN (unfiltered)	64	0.85	1.367 ± 0.100	-0.093 ± 0.202
New settings - NAQMN (unfiltered)	63	0.82	1.334 ± 0.108	0.029 ± 0.213
Paper - Lidar (filtered)	19	0.84	1.004 ± 0.142	-0.220 ± 0.294
New settings - Lidar (filtered)	19	0.90	0.962 ± 0.108	-0.288 ± 0.235
Paper - Lidar (unfiltered)	19	0.84	1.004 ± 0.142	-0.220 ± 0.294
New settings - Lidar (unfiltered)	18	0.90	0.874 ± 0.097	-0.142 ± 0.222
Paper - LP DOAS (filtered)	38	0.85	1.425 ± 0.135	-0.312 ± 0.290
New settings - LP DOAS (filtered)	40	0.79	1.315 ± 0.149	0.058 ± 0.306
Paper - LP DOAS (unfiltered)	44	0.82	1.333 ± 0.132	-0.102 ± 0.280
New settings - LP DOAS (unfiltered)	43	0.78	1.271 ± 0.139	0.078 ± 0.290

Tab. AT.7: *Regression parameters and Pearsons correlations coefficients for the paper results and the profiles retrieved with new settings and the AERONET phase function and SSA shown in the Figures 4.46 - 4.47. The regression equation follows:* $y = mx + b$.

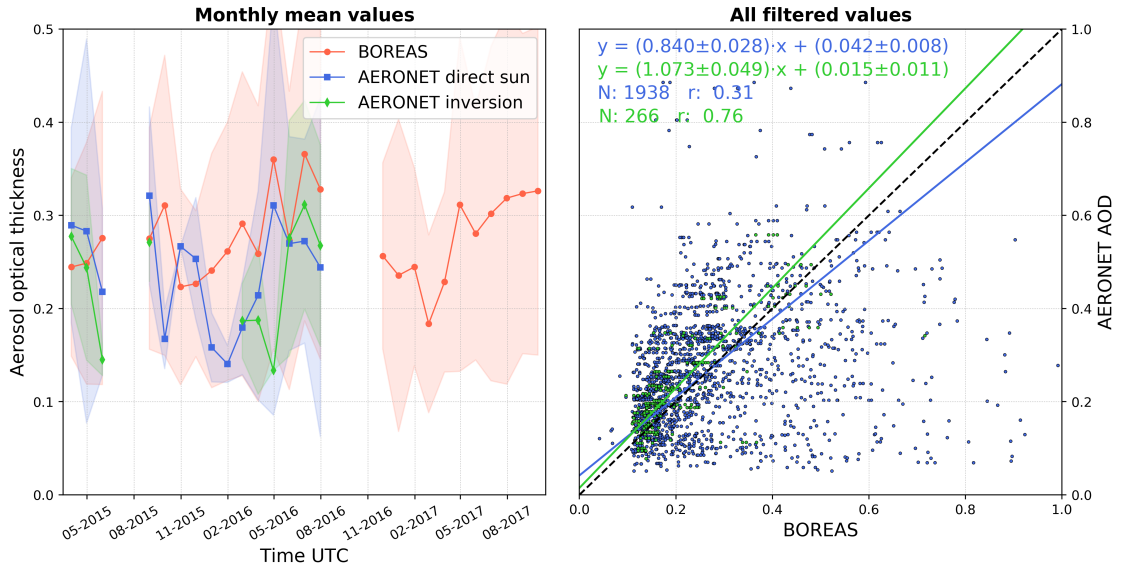


Fig. AF.58: ***Left****: Time series of monthly mean BOREAS AOT and AERONET direct sun and inversion AODs.* ***Right****: Scatter plot including correlation coefficients and regression lines for all filtered values in the UV spectral range.*

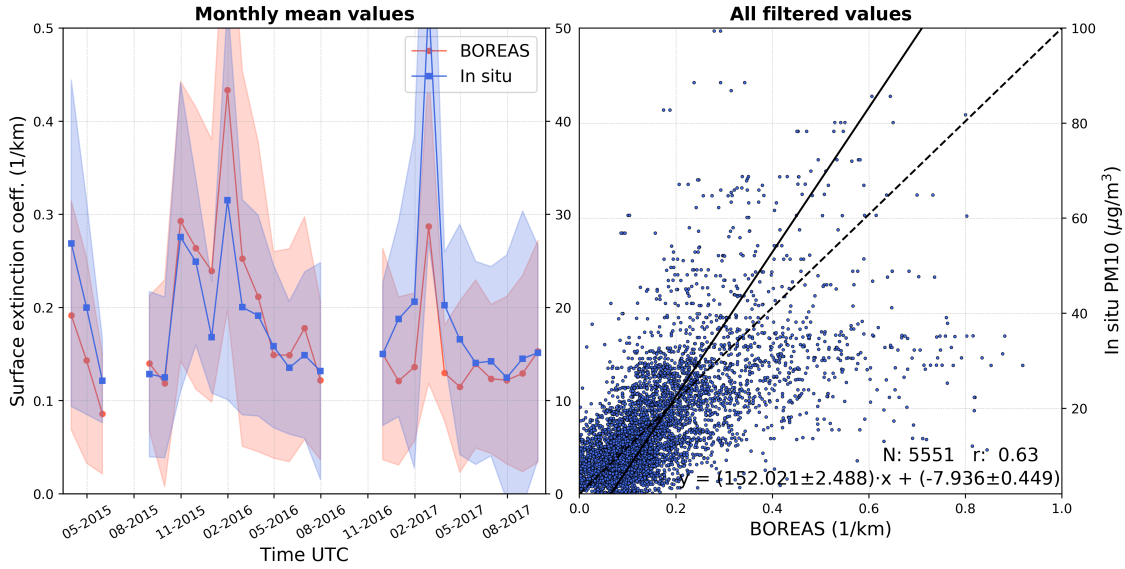


Fig. AF.59: ***Left****: Time series of monthly mean BOREAS surface extinction values and in situ PM10 concentrations from the in situ station I_2.* ***Right****: Scatter plot including correlation coefficients and regression lines for all filtered values in the UV spectral range.*

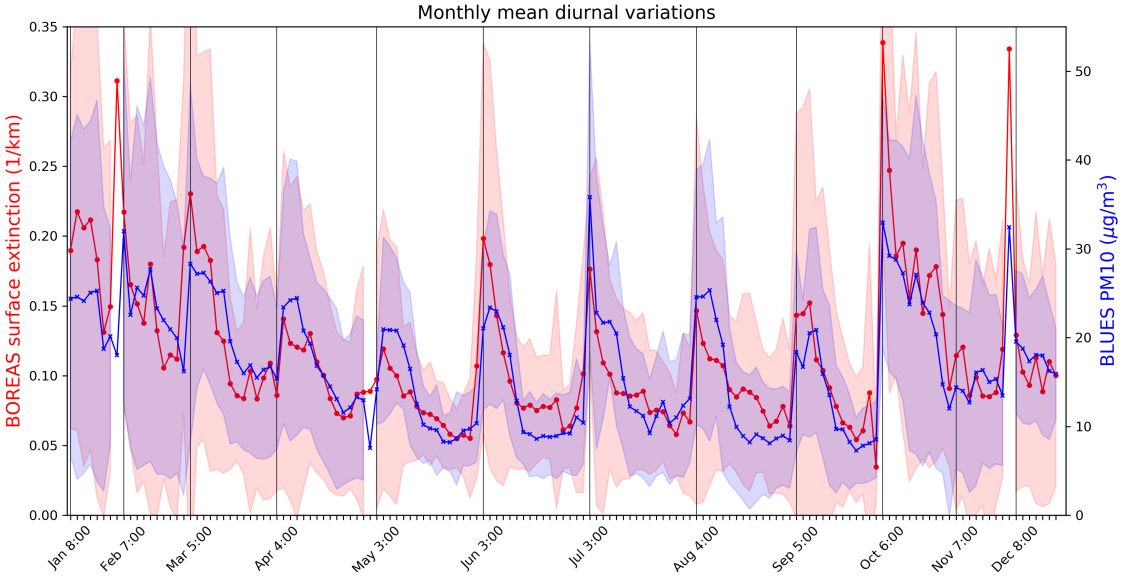


Fig. AF.60: *Comparison of monthly mean diurnal surface extinction and PM10 concentrations for UV. The shaded areas indicate the standard deviations.*

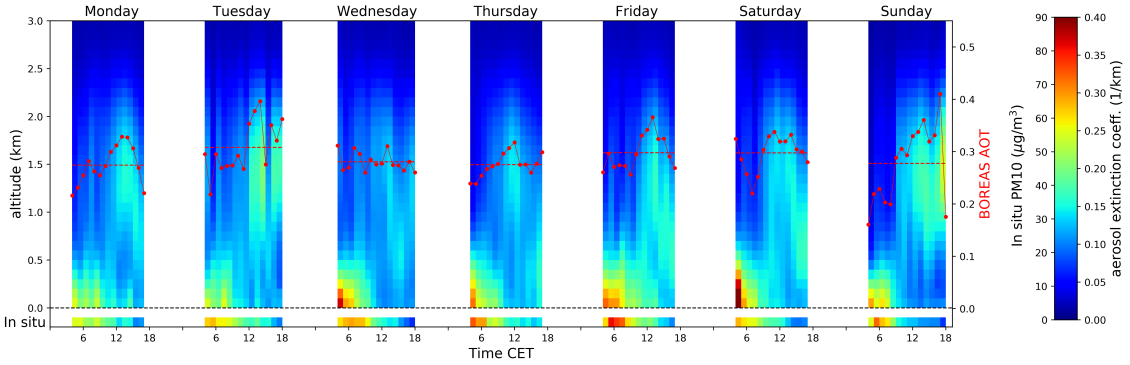


Fig. AF.61: *Average extinction coefficient profiles (colour-coded) for the days of the week retrieved in the ultra-violet spectral range. The lowest colours indicate the in situ PM10 concentrations for BLUES site I_2. The red lines show BOREAS AOT's with the mean level depicted as dashed line.*

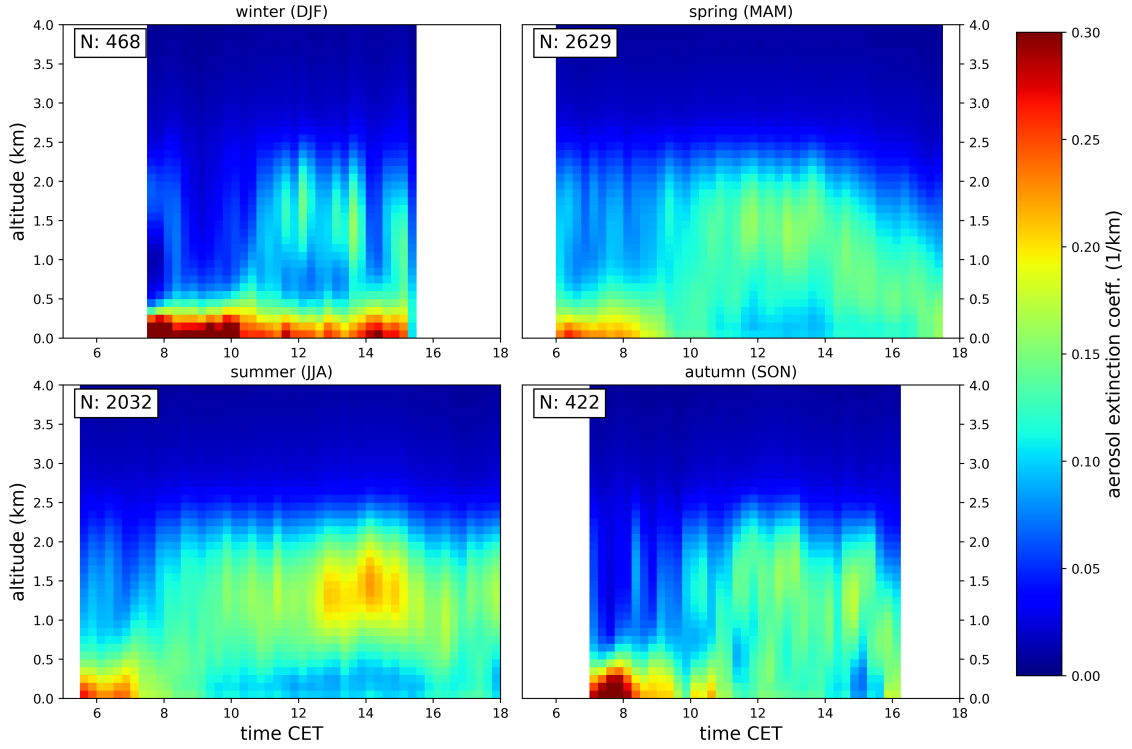


Fig. AF.62: *Average extinction coefficient profiles (colour-coded) for the seasons of the year retrieved in the ultra-violet spectral range. The numbers in the corners state the number of data points used for the averaging.*

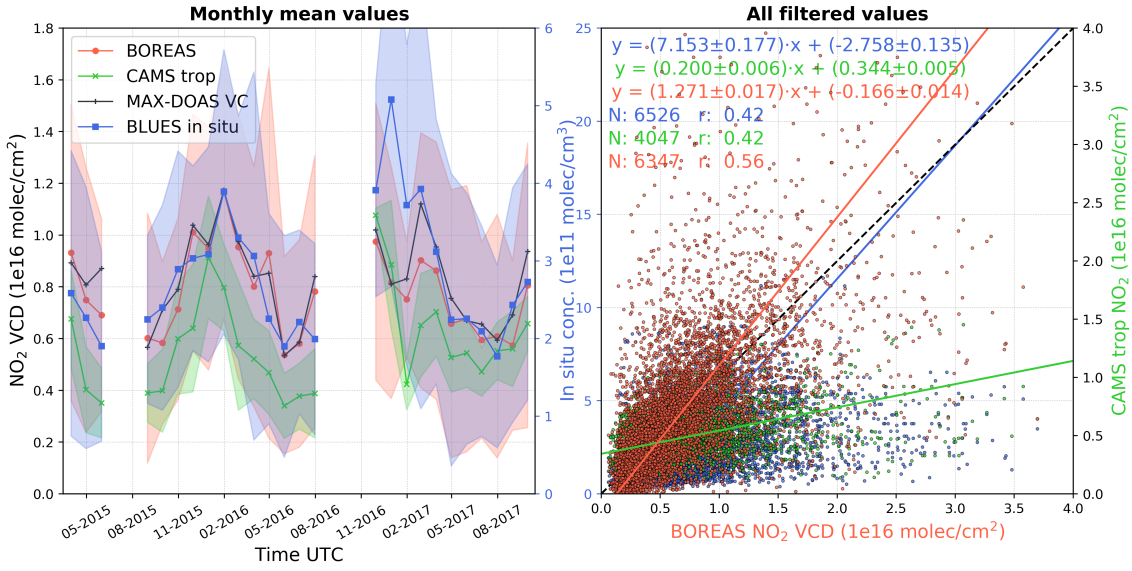


Fig. AF.63: ***Left****: Time series of monthly mean BOREAS, MAX-DOAS (geom. corrected VC), and CAMS VCD as well as BLUES in situ* NO_2*concentrations.* ***Right****: Scatter plot including correlation coefficients and regression lines for all filtered values in the ultra-violet spectral range.*

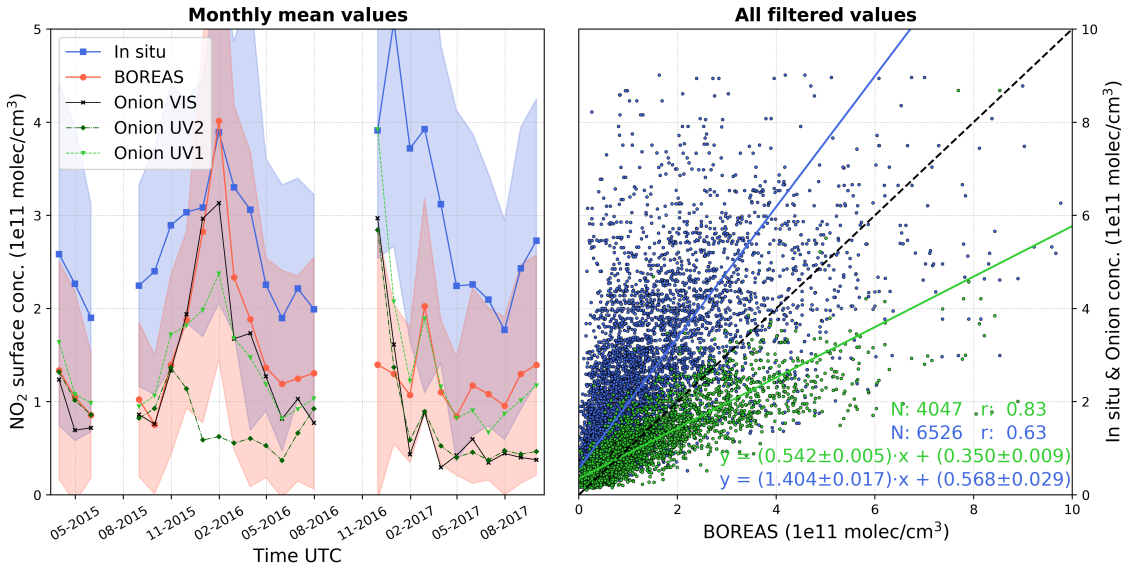


Fig. AF.64: ***Left****: Time series of monthly mean BOREAS* NO_2 *and BLUES in situ surface concentrations (in situ station* I_2*) in addition to the* NO_2 *concentrations retrieved with the onion peeling approach.* ***Right****: Scatter plot including correlation coefficients and regression lines for all filtered values in the ultra-violet spectral range.*

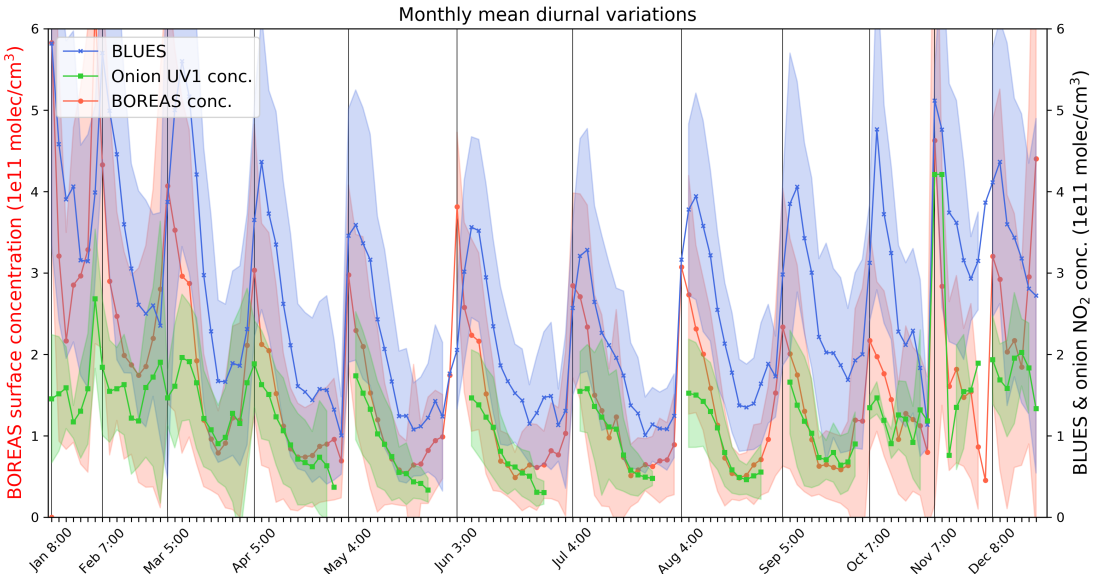


Fig. AF.65: *Comparison of monthly mean diurnal* NO_2 *surface concentrations retrieved with BOREAS and the onion peeling approach (for UV1) as well as data from the BLUES in situ site* I_2 *(UV). The shaded areas indicate the standard deviations.*

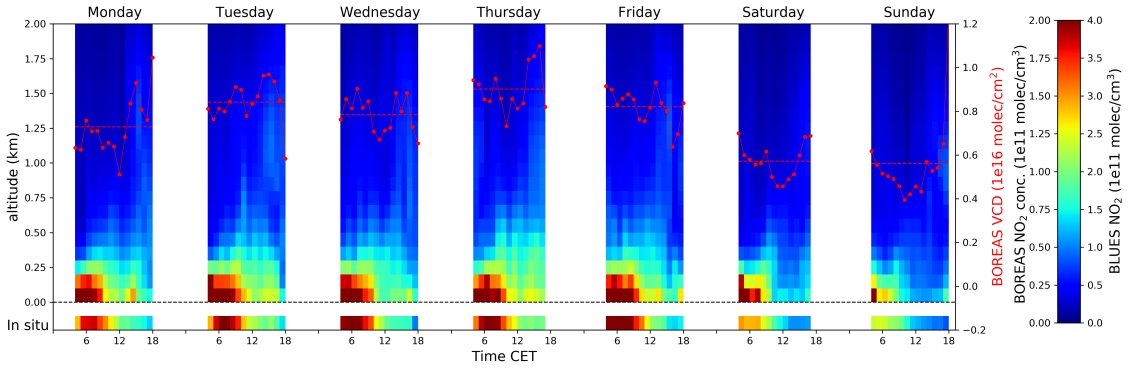


Fig. AF.66: *Average NO_2 concentration profiles (colour-coded) for the days of the week retrieved in the ultra-violet spectral range. The lowest colours indicate the in situ concentrations for BLUES site I_2. The red lines show BOREAS VCD's with the mean level depicted as dashed line.*

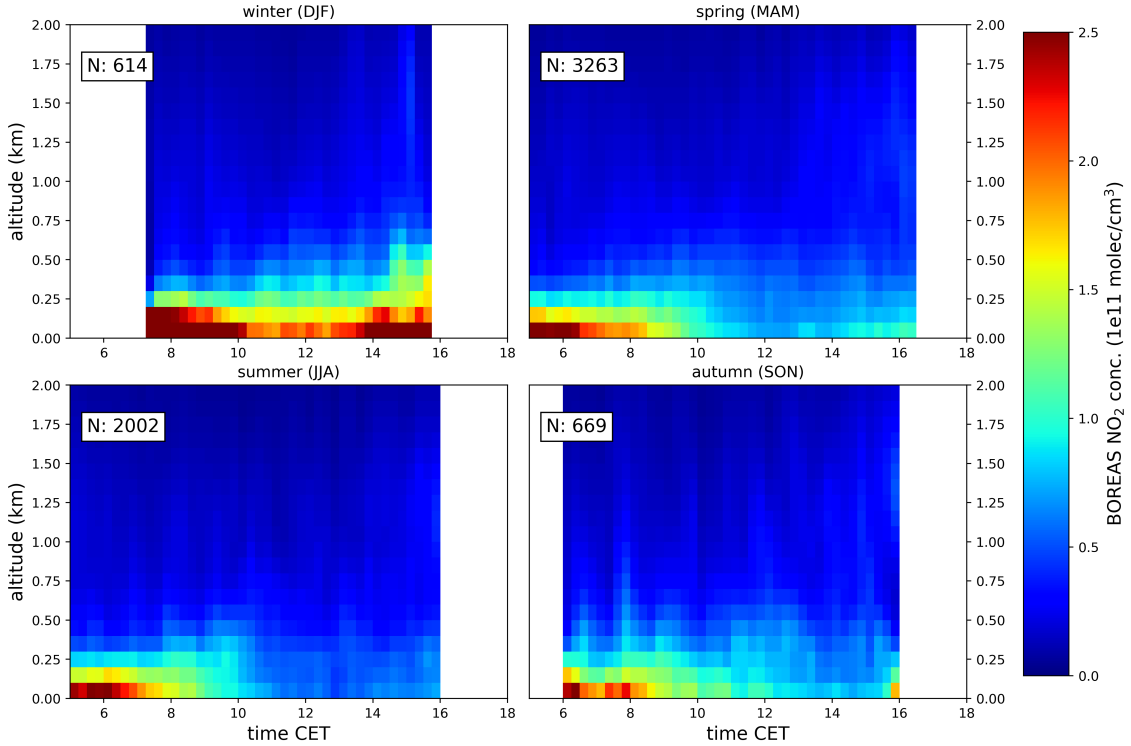


Fig. AF.67: *Average NO_2 concentration profiles (colour-coded) for the seasons of the year retrieved in the ultra-violet spectral range. The numbers in the corners state the number of data points used for the averaging.*

A.3 Theoretical supplements

A.3.1 Empirical detection limit

For the definition of a simple detection limit, Equation 3.1 is considered for one absorber only. The criterion is, that the product of $\Delta S \cdot \sigma(\lambda)$ should be larger or equal to the residual term:

$$\Delta S \cdot \sigma(\lambda) \geq r(\lambda) \Leftrightarrow \widetilde{D} \approx \frac{r(\lambda)}{\sigma(\lambda)} \tag{1.1}$$

Unfortunately, $\widetilde{D}$ is usually inconvenient to calculate because the residual is saved in different files for each fit (at IUP Bremen) which means a large computational effort for long time series. In addition, the detection limit calculated with the above equation is independent from the number of pixels because $r(\lambda)$ and $\sigma(\lambda)$ are on the same spectral grid. To satisfy the additional needs, the application of the root-sum-of-squares (RSS) to the above equation enables an additional estimation:

$$\Delta S \cdot \sigma(\lambda) \geq r(\lambda) \Rightarrow \Delta S \cdot \sqrt{\sum_{\lambda} \sigma^2(\lambda)} \geq \sqrt{\sum_{\lambda} r^2(\lambda)} \geq 3\sqrt{\frac{\sum_{\lambda} r^2(\lambda)}{N}} \tag{1.2}$$

The last step requires that the standard deviation (3σ) of the residual is smaller than the RSS of the residual, which is usually true. The factor 3 is somehow arbitrary but the estimation is valid as long as the multiplication factor is $x \leq \sqrt{N}$ (e.g. $x = 3 \leq \sqrt{500}$). With 3σ, it is expected that 99.8% of all residual values are within the interval of the standard deviation.
Now, the new detection limit can be found as follows:

$$D' = \frac{3 \cdot RMS}{\sqrt{\sum_{\lambda} \sigma^2(\lambda)}} \tag{1.3}$$

Here, the definition of the RMS (Eq. 3.2) was used and the differential cross sections are considered. Note that, with consideration of the order of the cross section instead of $\sqrt{\sum_{\lambda} \sigma^2(\lambda)}$, one can derive a rough detection limit as introduced by PETERS et al. [2012]. However, this detection limit is not suggested as it is independent of N and the order the cross sections differs strongly for various species and within the fitting window.

In Figure AF.68, this detection limit for NO_2 in the visible fitting window is depicted for different numbers of pixels N and for RMS = 1e-4. Colour coded are $\overline{D}$ derived with convolved NO_2 cross sections with Gaussian distributions (FWHM from 0 to $1\,\mathrm{nm}$). Note, that the relative difference between the largest FWHM and a not convolved cross section

is at maximum around 24%. The dependence on N can be neglected as long as N is larger than 100.

In contrast to the FWHM, Figure AF.69 shows that the RMS is the more important dependency. The grey vertical lines indicate the pixel numbers which match the properties of the Bremen MAX-DOAS instrument. For this figure, the fitting windows, detection limits and pixel numbers of the specific fitting windows can be seen in Table AT.8 (for RMS = 1e-4).

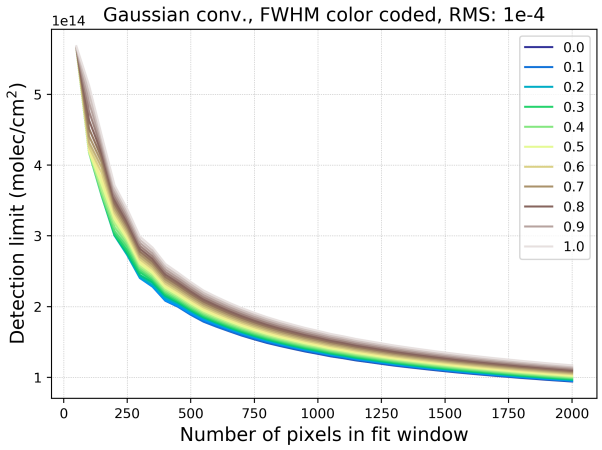


Fig. AF.68: *Detection limit for a different number of pixels within the fitting window. The cross section was convolved with a Gaussian with different FWHM values colour-coded (0 - 1.0; 0 means no convolution).* $425 - 490\,nm$*, RMS = 1e-4*

Species	fit window (nm)	No. of Pixel	Det. limit
NO_2	425 to 490	503	1.9e14 (molec/cm^2)
NO_2	338 to 370	504	4.3e14 (molec/cm^2)
O_4	425 to 490	503	8.6e40 (molec^2/cm^5)
O_4	338 to 370	504	1.1e41 (molec^2/cm^5)
HCHO	336.5 to 359	354	1.4e15 (molec/cm^2)
BrO	336 to 359	354	4.8e12 (molec/cm^2)
CHOCHO	420 to 465	348	2.0e14 (molec/cm^2)
SO_2	310 to 328	282	3.2e14 (molec/cm^2)

Tab. AT.8: *Detection limits for the MAX-DOAS instrument in Bremen (RMS of 1e-4).*

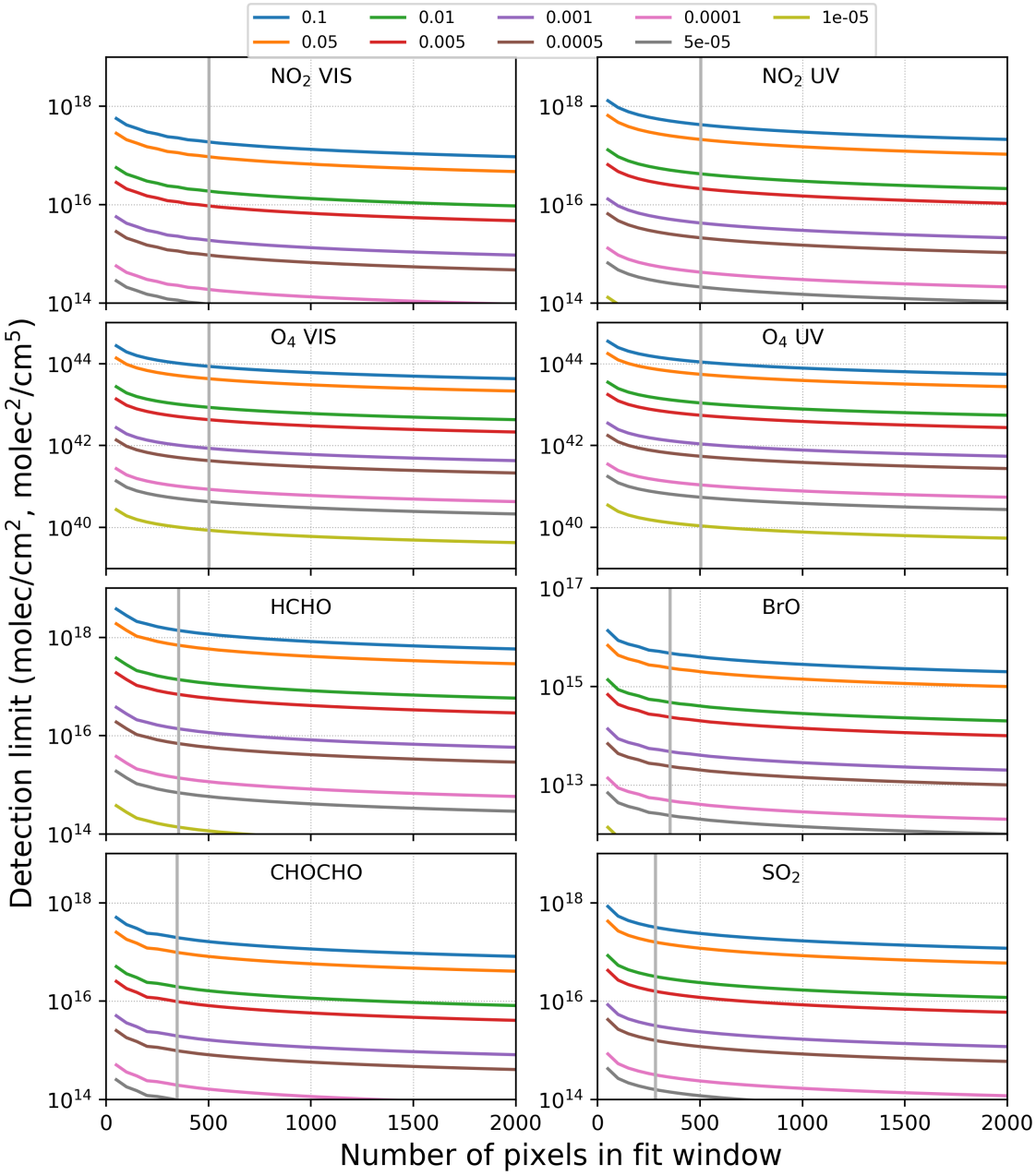


Fig. AF.69: *The detection limit plotted for various species with different RMS values colour-coded. The grey vertical lines indicate the specific number of pixels of the fitting windows in Table AT.8 for the MAX-DOAS instrument in Bremen. The cross sections are noted in Table 3.3 with the additional CHOCHO and SO_2 cross sections by* VOLKAMER et al. [2005] *and* BOGUMIL et al. [2003] *respectively.*

List of Figures

List of Tables

List of Acronyms

AERONET	AErosol RObotic NETwork
AOD	Aerosol Optical Depth
AOT	Aerosol Optical Thickness
AK	Averaging Kernel
AMF	Air Mass Factor
BAMF	Block (or Box) Air Mass Factor
BIRA	Belgian Institute for Space Aeronomy
BLUES	Bremer LuftUEberwachungsSystem, air quality network Bremen
BOREAS	Bremen Optimal estimation REtrieval for Aerosols and trace gaseS
BOT	BOTtom (extinction or concentration)
BREDOM	BREmian DOas Network for atmospheric Measurements
CAMS	Copernicus Atmosphere Monitoring Service
CCD	Charged Coupled Device
CI	Colour Index
CET	Central European Time
CINDI	Cabauw Intercomparison of Nitrogen Dioxide measuring Instruments
CME	Coronal Mass Ejection
CTM	Chemistry Transfer Model
dAMF	differential Air Mass Factor
DOAS	Differential Optical Absorption Spectroscopy
DOF	Degrees Of Freedom
dop	Degree Of Polarization
dSCD	differential Slant Column Density (also SC or SCD)
dSOT	differential Slant Optical Thickness
DWD	Deutscher Wetter Dienst
ECMWF	European Centre for Medium-Range Weather Forecast
E-PRTR	European Pollutant Release and Transfer Register
ESA	European Space Agency
FOV	Field Of View
FRM4DOAS	Fiducial Reference Measurements 4 ground-based DOAS
FWHM	Full Width at Half Maximum

IUP	Institut für UmweltPhysik, Institude for environmental physics
LIDAR	LIght Detection And Ranging
LM	Levenberg - Marquardt (algorithm)
LP DOAS	Long Path DOAS
LOS	Line Of Sight
LUT	Look Up Table
MAD-CAT	Multi Axis Doas - Comparison campaign for Aerosols and Trace gases
MAX-DOAS	Mult-AXis Differential Optical Absorption Spectroscopy
MLH	Mixing Layer Height
NAQMN	Dutch National Air Quality Monitoring Network
NASA	National Aeronautics and Space Adminsitration
NMVOC	Non-Methane Volatile Organic Compound
OE	Optimal Estimation
OVP	OVerPass
PBL	Planetary Boundary Layer (height)
PCA	Principle Component Analysis
PM	Particulate Matter
ppm	parts per million
ppb	parts per billion
PS	Power Station
QA4ECV	Quality Assurance 4 Essential Climate Variables
RAA	Relative Azimuth Angle
RIVM	Dutch National Institute for Public Health and the Environment (RijksInstituut voor Volksgezondheid en Milieu)
RF	Radiative Forcing
RMS	Root Mean Square
RRS	Rotational Raman Scattering
RTM	Radiative Transfer Model
SCIATRAN	Radiative Transfer Model for SCIAMACHY
SH	Scale Height
SNR	Signal to Noise Ratio
SSA	Single Scattering Albedo
SZA	Solar Zenith Angle
TOA	Top Of the Atmosphere
UV	Ultra Violet
VAA	Viewing Azimuth Angle
VCD	Vertical Column Density (also VC or VCD)
VIS	visible
VMR	Volume Mixing Ratio

VRC	Vibrational Raman Scattering
WHO	World Health Organisation
WMO	World Meteorological Organization

ACKNOWLEDGEMENT

Within the last years, several people have been of great importance to the success of this thesis. Their valuable advices and support helped me through tough times during my PhD life and enabled me to progress and finish this work.

I thank Prof. Dr. J. P. Burrows for giving me the opportunity of writing this PhD thesis at IUP Bremen and for his good advices which helped me to improve my thesis and my paper.
Furthermore, Prof. Dr. T. Wagner kindly agreed on being my second examiner for this thesis. Additionally, we had some great discussions and I thank you for your time and efforts in supporting me with this work:
Very special thanks goes to Andreas Richter for supervising my work and for giving me the most valuable advices and support in the framework of my PhD. There was not even one moment in which I did not receive an answer to my questions, even though I could sometimes see that you didn't have time for me back then (but I asked anyway, because of my tenacious nature...). Many thanks to you!
I thank Enno Peters for his supervision during the time he worked at IUP but also for his advices afterwards. Besides his scientific support, I thank Enno as well as Lisa Behrens and André Seyler for their time offside IUP during different leisure activities which helped to come down in stressful times.
Additionally, my thanks go to Lisa, Enno, and Andreas for their advices and corrections which helped to improve my thesis. It wouldn't be the same without your valuable input.
I thank my former and recent office colleagues Stefan Schreier and Ilias Bougoudis (and Lisa again). It was always fun to share the office with you.
Of course, I thank the whole DOAS group (past and present) for the pleasant atmosphere in the last years. I received a lot of support from all of you and I enjoyed the several group activities a lot.
Furthermore, people from other groups at IUP Bremen were important for the success of this thesis. I thank Mihalis Vrekoussis and Andreas Hilboll from the LAMOS group for helping me to understand and improve several aspects of my work.
I thank Vladimir and Alexei Rozanov for their numerous advices and their persistent sup-

port regarding the use of SCIATRAN.
Naturally, many thanks to the whole IUP Bremen and all of its members as I felt always comfortable here and everyone was eager to help me when I had a question.

Besides colleagues from work, I would like to thank several other people who were of importance for me. Unfortunately, I can not name them all individually so here is a brief summary. I thank all of my friends, former fellow students, and football players who helped me to relax offside from the IUP and to clear my mind in stressful times.

Finally, I thank my family for their unlimited support. This work as well as my study before were not possible without you. Many thanks!

Of course, I would like to apologize to all people who were not acknowledged here. Consider yourself as most important but my end-of-PhD-time-mental-condition does not enable me to remember all of you... Thanks!

CHAPTER

B

BIBLIOGRAPHY

ALIWELL, S. R., M. Van ROOZENDAEL, P. V. JOHNSTON, A. RICHTER, T. WAGNER, D. W. ARLANDER, J. P. BURROWS, D. J. FISH, R. L. JONES, K. K. TØRNKVIST, J.-C. LAMBERT, K. PFEILSTICKER, and I. PUNDT (**2002**). "Analysis for BrO in Zenith-Sky Spectra: An Intercomparison Exercise for Analysis Improvement". In: *Journal of Geophysical Research: Atmospheres* 107.D14, ACH 10-1-ACH 10-20. ISSN: 2156-2202. DOI: 10/fqh92g. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001JD000329 (visited on 07/22/2018) (cit. on p. 60).

ASCHWANDEN, M. J. (**2006**). *Physics of the Solar Corona: An Introduction with Problems and Solutions*. 2nd ed. Springer-praxis books in astrophysics and astronomy. Berlin ; New York: Springer. 892 pp. ISBN: 978-3-540-30765-5 (cit. on pp. 6, 8).

BACKUS, G. and F. GILBERT (**1970**). "Uniqueness in the Inversion of Inaccurate Gross Earth Data". In: *Philosophical Transactions of the Royal Society of London Series A* 266, pp. 123–192. ISSN: 0962-8436. DOI: 10/fr5443. URL: http://adsabs.harvard.edu/abs/1970RSPTA.266..123B (visited on 08/14/2018) (cit. on p. 133).

BARRET, B., M. DE MAZIÈRE, and P. DEMOULIN (**2002**). "Retrieval and Characterization of Ozone Profiles from Solar Infrared Spectra at the Jungfraujoch". In: *Journal of Geophysical Research: Atmospheres* 107.D24, p. 4788. ISSN: 2156-2202. DOI: 10.1029/2001JD001298. URL: http://onlinelibrary.wiley.com/doi/10.1029/2001JD001298/abstract (visited on 02/10/2018) (cit. on pp. 113, 168).

BEIRLE, S., U. PLATT, M. WENIG, and T. WAGNER (**2003**). "Weekly Cycle of NO2 by GOME Measurements: A Signature of Anthropogenic Sources". In: *Atmospheric Chemistry and Physics* 3.6, pp. 2225–2232. ISSN: 1680-7324. DOI: 10/fg79vw. URL: https://www.atmos-chem-phys.net/3/2225/2003/ (visited on 06/06/2018) (cit. on p. 81).

BLECHSCHMIDT, Anne-Marlene, Joaquim ARTETA, Adriana COMAN, Lyana CURIER, Henk ESKES, Gilles FORET, Clio GIELEN, Francois HENDRICK, Virginie MARÉCAL, Frédérik MELEUX, Jonathan PARMENTIER, Enno PETERS, Gaia PINARDI, Ankie J. M. PITERS, Matthieu PLU, Andreas RICHTER, Mikhail SOFIEV, Álvaro M. VALDEBENITO, Michel Van ROOZENDAEL, Julius VIRA, Tim VLEMMIX, and John P. BURROWS (**2017**). "Comparison of Tropospheric NO_2 Columns from MAX-DOAS Retrievals and Regional Air Quality Model Simulations". In: *Atmospheric Chemistry and Physics Discussions*, pp. 1–50. ISSN: 1680-7316. DOI: 10/gcdk9s. URL: https://www.atmos-chem-phys-discuss.net/acp-2016-1003/ (visited on 11/23/2018) (cit. on p. 202).

BLUES (**2017**). *BLUES - Das Bremer Luftüberwachungssystem*. URL: https://www.bauumwelt.bremen.de/umwelt/luft/luftqualitaet-24505 (visited on 03/15/2018) (cit. on pp. 52, 54).

BOBROWSKI, N., G. HÖNNINGER, B. GALLE, and U. PLATT (**2003**). "Detection of Bromine Monoxide in a Volcanic Plume". In: *Nature* 423.6937, pp. 273–276. ISSN: 1476-4687. DOI: 10/bv8hq8. URL: https://www.nature.com/articles/nature01625 (visited on 05/06/2018) (cit. on pp. 2, 105).

BOGUMIL, K, J ORPHAL, T HOMANN, S VOIGT, P SPIETZ, O. C FLEISCHMANN, A VOGEL, M HARTMANN, H KROMMINGA, H BOVENSMANN, J FRERICK, and J. P BURROWS (**2003**). "Measurements of Molecular Absorption Spectra with the SCIAMACHY Pre-Flight Model: Instrument Characterization and Reference Data for Atmospheric Remote-Sensing in the 230–2380 Nm Region". In: *Journal of Photochemistry and Photobiology A: Chemistry*. Atmospheric Photochemistry 157.2, pp. 167–184. ISSN: 1010-6030. DOI: 10/bhxtm9. URL: http://www.sciencedirect.com/science/article/pii/S1010603003000625 (visited on 06/05/2018) (cit. on p. 275).

BÖSCH, T., V. ROZANOV, A. RICHTER, E. PETERS, A. ROZANOV, F. WITTROCK, A. MERLAUD, J. LAMPEL, S. SCHMITT, M. de HAIJ, S. BERKHOUT, B. HENZING, A. APITULEY, M. den HOED, J. VONK, M. TIEFENGRABER, M. MÜLLER, and J. P. BURROWS (**2018**). "BOREAS - a New MAX-DOAS Profile Retrieval Algorithm for Aerosols and Trace Gases". In: *Atmospheric Measurement Techniques Discussions* 2018, pp. 1–39. ISSN: 1867-8610. DOI: 10/gdqbdd. URL: https://www.atmos-meas-tech-discuss.net/amt-2018-153/ (visited on 06/25/2018) (cit. on pp. 38, 105, 107–109, 112, 167, 177, 179, 181, 183, 187, 213).

BOTHMER, V. and I. A. DAGLIS (**2007**). *Space Weather: Physics and Effects*. Springer-Praxis books in environmental sciences. Berlin ; New York: Springer. 438 pp. ISBN: 978-3-540-23907-9 (cit. on pp. 5, 7).

BOVENSMANN, H., J. P. BURROWS, M. BUCHWITZ, J. FRERICK, S. NOËL, V. V. ROZANOV, K. V. CHANCE, and A. P. H. GOEDE (**1999**). "SCIAMACHY: Mission Objectives and Measurement Modes". In: *Journal of the atmospheric sciences* 56.2, pp. 127–150. DOI: 10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1999)056%3C0127:SMOAMM%3E2.0.CO;2 (visited on 09/12/2017) (cit. on pp. 2, 34).

BREWER, A. W., C. T. MCELROY, and J. B. KERR (**1973**). "Nitrogen Dioxide Concentrations in the Atmosphere". In: *Nature* 246.5429, pp. 129–133. ISSN: 1476-4687. DOI: 10/bccbns. URL: https://www.nature.com/articles/246129a0 (visited on 09/14/2018) (cit. on p. 2).

BRINKMANN, R. T. (**1968**). "Rotational Raman Scattering in Planetary Atmospheres". In: *The Astrophysical Journal* 154, p. 1087. ISSN: 0004-637X. DOI: 10.1086/149827. URL: http://adsabs.harvard.edu/abs/1968ApJ...154.1087B (visited on 03/04/2018) (cit. on p. 31).

BRINKSMA, E. J., G. PINARDI, H. VOLTEN, R. BRAAK, A. RICHTER, A. SCHÖNHARDT, M. van ROOZENDAEL, C. FAYT, C. HERMANS, R. J. DIRKSEN, T. VLEMMIX, A. J. C. BERKHOUT, D. P. J. SWART, H. OETJEN, F. WITTROCK, T. WAGNER, O. W. IBRAHIM, G. de LEEUW, M. MOERMAN, R. L. CURIER, E. A. CELARIER, A. CEDE, W. H. KNAP, J. P. VEEFKIND, H. J. ESKES, M. ALLAART, R. ROTHE, A. J. M. PITERS, and P. F. LEVELT (**2008**). "The 2005 and 2006 DANDELIONS NO_2 and Aerosol Intercomparison Campaigns". In: *Journal of Geophysical Research* 113.D16. ISSN: 0148-0227. DOI: 10.1029/2007JD008808. URL: http://doi.wiley.com/10.1029/2007JD008808 (visited on 09/12/2017) (cit. on p. 173).

BRUNS, Marco (**2004**). "NO_2 Profile Retrieval Using Airborne Multiaxis DifferentialOptical Absorption Spectrometer (AMAXDOAS) Data". Phd thesis. Bremen: University of Bremen. URL: http://nbn-resolving.de/urn:nbn:de:gbv:46-diss000011240 (cit. on p. 126).

BURROWS, J. P., E. HÖLZLE, A. P. H. GOEDE, H. VISSER, and W. FRICKE (**Apr. 1, 1995**). "SCIAMACHY—Scanning Imaging Absorption Spectrometer for Atmospheric Chartography". In: *Acta Astronautica*. Earth Observation 35.7, pp. 445–451. ISSN: 0094-5765. DOI: 10/fvbbw9. URL: http://www.sciencedirect.com/science/article/pii/009457659400278T (visited on 12/14/2018) (cit. on p. 2).

BURROWS, J.P., P. BORRELL, and U. PLATT (**2011**). *The Remote Sensing of Tropospheric Composition from Space*. Physics of Earth and Space Environments. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-642-14791-3. DOI: 10.1007/978-3-642-14791-3. URL: http://link.springer.com/10.1007/978-3-642-14791-3 (visited on 09/12/2017) (cit. on pp. 19, 34, 48).

BURROWS, J.P., M. WEBER, M. BUCHWITZ, V. ROZANOV, Annette LADSTÄTTER-WEI\S SENMAYER, A. RICHTER, R. DEBEEK, R. HOOGEN, K. BRAMSTEDT, K.-U. EICHMANN, et al. (**1999**). "The Global Ozone Monitoring Experiment (GOME): Mission Concept and First Scientific Results". In: *Journal of the Atmospheric Sciences* 56.2, pp. 151–175. DOI: 10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1999)056%3C0151:TGOMEG%3E2.0.CO%3B2 (visited on 09/12/2017) (cit. on pp. 2, 34).

CALLIES, J, E CORPACCIOLI, M EISINGER, A HAHNE, and A LEFEBVRE (**2000**). "GOME-2 – Metop's Second-Generation Sensor for Operational Ozone Monitoring". In: *ESA Bulletin* 102, p. 9. URL: https://esamultimedia.esa.int/docs/metop/GOME-2-102.pdf (cit. on p. 2).

CESAR (**2018**). *Cesar Observatory*. URL: http://www.cesar-observatory.nl/index.php (visited on 02/13/2018) (cit. on pp. 168, 174, 184).

CHAN, K. L., A. HARTL, Y. F. LAM, P. H. XIE, W. Q. LIU, H. M. CHEUNG, J. LAMPEL, D. PÖHLER, A. LI, J. XU, H. J. ZHOU, Z. NING, and M. O. WENIG (**2015**). "Observations of Tropospheric NO2 Using Ground Based MAX-DOAS and OMI Measurements during the Shanghai World Expo 2010". In: *Atmospheric Environment* 119, pp. 45–58. ISSN: 1352-2310. DOI: 10/f7wbh5. URL: http://www.sciencedirect.com/science/article/pii/S1352231015302867 (visited on 06/04/2018) (cit. on p. 78).

CHAN, K. L., M. WIEGNER, M. WENIG, and D. PÖHLER (**2018**). "Observations of Tropospheric Aerosols and NO2 in Hong Kong over 5years Using Ground Based MAX-DOAS". In: *Science of The Total Environment* 619-620, pp. 1545–1556. ISSN: 0048-9697. DOI: 10/gffng8. URL: http://www.sciencedirect.com/science/article/pii/S0048969717328656 (visited on 10/26/2018) (cit. on p. 205).

CHANDRASEKHAR, Subrahmanyan (**1960**). *Radiative Transfer*. OCLC: ocm00335528. New York: Dover Publications. 393 pp. ISBN: 978-0-486-60590-6 (cit. on p. 31).

CLÉMER, K., M. VAN ROOZENDAEL, C. FAYT, F. HENDRICK, C. HERMANS, G. PINARDI, R. SPURR, P. WANG, and M. DE MAZIÈRE (**2010**). "Multiple Wavelength Retrieval of Tropospheric Aerosol Optical Properties from MAXDOAS Measurements in Beijing". In: *Atmospheric Measurement Techniques* 3.4, pp. 863–878. ISSN: 1867-8548. DOI: 10.5194/amt-3-863-2010. URL: http://www.atmos-meas-tech.net/3/863/2010/ (visited on 09/12/2017) (cit. on pp. 2, 63, 106, 116, 153, 196).

CRUTZEN, P. J. (**1970**). "The Influence of Nitrogen Oxides on the Atmospheric Ozone Content". In: *Quarterly Journal of the Royal Meteorological Society* 96.408, pp. 320–325. ISSN: 1477-870X. DOI: 10/bp9fkd. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49709640815 (visited on 09/14/2018) (cit. on p. 1).

DASKALAKIS, Nikos, Kostas TSIGARIDIS, Stelios MYRIOKEFALITAKIS, George S. FANOURGAKIS, and Maria KANAKIDOU (**2016**). "Large Gain in Air Quality Compared to an Alternative Anthropogenic Emissions Scenario". In: *Atmospheric Chemistry and Physics* 16.15, pp. 9771–9784. ISSN: 1680-7316. DOI: 10/f83kxn. URL: https://www.atmos-chem-phys.net/16/9771/2016/acp-16-9771-2016.html (visited on 08/13/2018) (cit. on p. 100).

DAVIDSON, P. A. (**2002**). "An Introduction to Magnetohydrodynamics". In: *American Journal of Physics* 70.7, pp. 781–781. ISSN: 0002-9505. DOI: 10.1119/1.1482065. URL: http://aapt.scitation.org/doi/abs/10.1119/1.1482065 (visited on 03/01/2018) (cit. on p. 5).

DEE, D.P., S. M. UPPALA, A. J. SIMMONS, P. BERRISFORD, P. POLI, KOBAYASHI S., ANDRAE U., BALMASEDA M. A., BALSAMO G., BAUER P., BECHTOLD P., BELJAARS A. C. M., VAN DE BERG L., BIDLOT J., BORMANN N., DELSOL C., DRAGANI R., FUENTES M., GEER A. J., HAIMBERGER L., HEALY S. B., HERSBACH H., HÓLM E. V., ISAKSEN L., KÅLLBERG P., KÖHLER M., MATRICARDI M., MCNALLY A. P., MONGE-SANZ B. M., MORCRETTE J.-J., PARK B.-K., PEUBEY C., DE ROSNAY P., TAVOLATO C., THÉPAUT J.-N., and VITART F. (**2011**). "The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System". In: *Quarterly Journal of the Royal Meteorological Society* 137.656, pp. 553–597. ISSN: 0035-9009. DOI: 10/cz2w58. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.828 (visited on 03/22/2018) (cit. on p. 63).

DEMTRÖDER, W. (**2006**). *Atoms, Molecules, and Photons: An Introduction to Atomic-, Molecular-, and Quantum-Physics*. Berlin ; New York: Springer. 571 pp. ISBN: 978-3-540-20631-6 (cit. on p. 15).

– (**2010**). *Experimentalphysik 4*. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 978-3-642-01598-4. DOI: 10.1007/978-3-642-01598-4. URL: http://link.springer.com/10.1007/978-3-642-01598-4 (visited on 09/12/2017) (cit. on pp. 15, 19).

DUBOVIK, O., A. SMIRNOV, B. N. HOLBEN, M. D. KING, Y. J. KAUFMAN, T. F. ECK, and I. SLUTSKER (**2000**). "Accuracy Assessments of Aerosol Optical Properties Retrieved from Aerosol Robotic Network (AERONET) Sun and Sky Radiance Measurements". In: *Journal of Geophysical Research: Atmospheres* 105.D8, pp. 9791–9806. ISSN: 2156-2202. DOI: 10.1029/2000JD900040. URL: http://onlinelibrary.wiley.com/doi/10.1029/2000JD900040/abstract (visited on 09/12/2017) (cit. on pp. 63, 169, 225).

DUBOVIK, Oleg, Brent HOLBEN, Thomas F. ECK, Alexander SMIRNOV, Yoram J. KAUFMAN, Michael D. KING, Didier TANRÉ, and Ilya SLUTSKER (**Feb. 1, 2002**). "Variability of Absorption and Optical Properties of Key Aerosol Types Observed in Worldwide Locations". In: *Journal of the Atmospheric Sciences* 59.3, pp. 590–608. ISSN: 0022-4928. DOI: 10/cdvdp3. URL: https://journals.ametsoc.org/doi/abs/10.1175/1520-0469(2002)059%3C0590:VOAAOP%3E2.0.CO%3B2 (visited on 05/06/2018) (cit. on p. 109).

DWD (**Mar. 15, 2018**). *Deutscher Wetterdienst - FTP Server*. URL: ftp://ftp-cdc.dwd.de/pub/CDC/observations_germany/climate/hourly/ (visited on 03/15/2018) (cit. on pp. 51, 88).

E-PRTR (**2018**). *European Pollutant Release and Transfer Register - European Environment Agency*. URL: http://prtr.ec.europa.eu/#/facilitylevels (visited on 03/15/2018) (cit. on p. 53).

EISINGER, M., A. RICHTER, A. LADSTÄTTER-WEISSENMAYER, and J. P. BURROWS (**1997**). "DOAS Zenith Sky Observations: 1. BrO Measurements over Bremen (53°N) 1993–1994". In: *Journal of Atmospheric Chemistry* 26.1, pp. 93–108. ISSN: 1573-0662. DOI: 10/bqnbf5. URL: https://doi.org/10.1023/A:1005776629692 (visited on 09/14/2018) (cit. on p. 2).

FINLAYSON-PITTS, Barbara J. and James N. PITTS JR (**1999**). *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*. Academic press. ISBN: 978-0-08-052907-3 (cit. on p. 48).

FITSCHEN, Arnd and Hartwig NORDMANN (**2015**). *Verkehrsentwicklung Auf Bundesfernstraßen 2015*. URL: http://bast.opus.hbz-nrw.de/frontdoor.php?source_opus=1904&la=de (visited on 05/25/2018) (cit. on p. 81).

FLAGAN, Richard C. and John H. SEINFELD (**1988**). *Fundamentals of Air Pollution Engineering*. Prentice Hall. 568 pp. ISBN: 978-0-13-332537-9 (cit. on pp. 46, 47).

FLEISCHMANN, Oliver C., Matthias HARTMANN, John P. BURROWS, and Johannes ORPHAL (**2004**). "New Ultraviolet Absorption Cross-Sections of BrO at Atmospheric Temperatures Measured by Time-Windowing Fourier Transform Spectroscopy". In: *Journal of Photochemistry and Photobiology A: Chemistry* 168.1, pp. 117–132. ISSN: 1010-6030. DOI: 10/cw4rbn. URL: http://www.sciencedirect.com/science/article/pii/S1010603004001522 (visited on 03/20/2018) (cit. on p. 63).

FLETCHER, R. (**1987**). *Practical Methods of Optimization; (2Nd Ed.)* New York, NY, USA: Wiley-Interscience. ISBN: 978-0-471-91547-8 (cit. on p. 157).

FRIESS, U., H. KLEIN BALTINK, S. BEIRLE, K. CLÉMER, F. HENDRICK, B. HENZING, H. IRIE, G. de LEEUW, A. LI, M. M. MOERMAN, M. van ROOZENDAEL, R. SHAIGANFAR, T. WAGNER, Y. WANG, P. XIE, S. YILMAZ, and P. ZIEGER (**2016**). "Intercomparison of Aerosol Extinction Profiles Retrieved from MAX-DOAS Measurements". In: *Atmospheric Measurement Techniques* 9.7, pp. 3205–3222. ISSN: 1867-8548. DOI: 10.5194/amt-9-3205-2016. URL: https://www.atmos-meas-tech.net/9/3205/2016/ (visited on 09/12/2017) (cit. on pp. 106, 196, 197).

FRIESS, U., P. S. MONKS, J. J. REMEDIOS, A. ROZANOV, R. SINREICH, T. WAGNER, and U. PLATT (**2006**). "MAX-DOAS O_4 Measurements: A New Technique to Derive Information on Atmospheric Aerosols: 2. Modeling Studies". In: *Journal of Geophysical Research* 111.D14. ISSN: 0148-0227. DOI: 10.1029/2005JD006618. URL: http://doi.wiley.com/10.1029/2005JD006618 (visited on 09/12/2017) (cit. on pp. 2, 49, 106, 110, 116, 153).

GIELEN, C., M. VAN ROOZENDAEL, F. HENDRICK, G. PINARDI, T. VLEMMIX, V. DE BOCK, H. DE BACKER, C. FAYT, C. HERMANS, D. GILLOTAY, and P. WANG (**2014**). "A Simple and Versatile Cloud-Screening Method for MAX-DOAS Retrievals". In: *Atmospheric Measurement Techniques* 7.10, pp. 3509–3527. ISSN: 1867-8548. DOI: 10.5194/amt-7-3509-2014. URL: http://www.atmos-meas-tech.net/7/3509/2014/ (visited on 09/12/2017) (cit. on p. 67).

GRAINGER, J. F. and J. RING (**1962**). "Anomalous Fraunhofer Line Profiles". In: *Nature* 193.4817, p. 762. ISSN: 1476-4687. DOI: 10.1038/193762a0. URL: https://www.nature.com/articles/193762a0 (visited on 03/04/2018) (cit. on p. 30).

GRATSEA, Myrto, Mihalis VREKOUSSIS, Andreas RICHTER, Folkard WITTROCK, Anja SCHÖNHARDT, John BURROWS, Stelios KAZADZIS, Nikos MIHALOPOULOS, and Evangelos GERASOPOULOS (**2016**). "Slant Column

MAX-DOAS Measurements of Nitrogen Dioxide, Formaldehyde, Glyoxal and Oxygen Dimer in the Urban Environment of Athens". In: *Atmospheric Environment* 135, pp. 118–131. ISSN: 1352-2310. DOI: 10/gdh9th. URL: http://www.sciencedirect.com/science/article/pii/S1352231016302400 (visited on 05/24/2018) (cit. on pp. 3, 83).

GREENBLATT, Gary D., John J. ORLANDO, James B. BURKHOLDER, and A. R. RAVISHANKARA (**1990**). "Absorption Measurements of Oxygen between 330 and 1140 Nm". In: *Journal of Geophysical Research: Atmospheres* 95.D11, pp. 18577–18582. ISSN: 2156-2202. DOI: 10/bzk9dz. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JD095iD11p18577 (visited on 07/19/2018) (cit. on p. 27).

HAKEN, Hermann and Hans C. WOLF (**2006**). *Molekülphysik Und Quantenchemie: Einführung in Die Experimentellen Und Theoretischen Grundlagen*. 5th ed. Springer-Lehrbuch. Berlin Heidelberg: Springer-Verlag. ISBN: 978-3-540-30314-5 (cit. on p. 15).

HAKEN, Hermann and Hans Christoph WOLF (**2002**). *Atom- und Quantenphysik: Einführung in die experimentellen und theoretischen Grundlagen ; mit 29 Tabellen, 173 Aufgaben und vollständigen Lösungen*. 7., aktualisierte und erw. Aufl., 2. Nachdr. Springer-Lehrbuch. OCLC: 249389735. Berlin: Springer. 513 pp. ISBN: 978-3-540-67453-5 (cit. on p. 15).

HENDRICK, F., B. DILS, C. GIELEN, B. LANGEROCK, G. PINARDI, M. DE MAZIÈRE, M. VAN ROOZENDAEL, E. PETERS, A. RICHTER, A. PITERS, S. BEIRLE, T. WAGNER, T. DROSOGLOU, A. BAIS, S. WANG, C. CUEVAS, and SAIZ-LOPEZ A. (**2016a**). *QA4ECV D3.8*. URL: http://www.qa4ecv.eu/sites/default/files/QA4ECV_D3.8_v1.0_web.pdf (visited on 03/21/2018) (cit. on p. 62).

HENDRICK, F., J.-F. MÜLLER, K. CLÉMER, P. WANG, M. DE MAZIÈRE, C. FAYT, C. GIELEN, C. HERMANS, J. Z. MA, G. PINARDI, T. STAVRAKOU, T. VLEMMIX, and M. VAN ROOZENDAEL (**2014**). "Four Years of Ground-Based MAX-DOAS Observations of HONO and NO_2 in the Beijing Area". In: *Atmospheric Chemistry and Physics* 14.2, pp. 765–781. ISSN: 1680-7324. DOI: 10.5194/acp-14-765-2014. URL: https://www.atmos-chem-phys.net/14/765/2014/ (visited on 09/12/2017) (cit. on p. 153).

HENDRICK, F., M. VAN ROOZENDAEL, A. RICHTER, T. WAGNER, U. FRIESS, and K. KREHER (**2016b**). *Semi-Blind Intercomparison Protocol for CINDI-2 Second Cabauw Intercomparison of Nitrogen Dioxide Measuring Instruments*. URL: http://www.tropomi.eu/data-products/planning-information (visited on 03/20/2018) (cit. on p. 62).

HENDRICK, François, Brice BARRET, M. van ROOZENDAEL, Hartmut BOESCH, Andre BUTZ, M. DE MAZIÈRE, Florence GOUTAIL, Christian HERMANS, J.-C. LAMBERT, Klaus PFEILSTICKER, and J.-P. POMMEREAU (**2004**). "Retrieval of Nitrogen Dioxide Stratospheric Profiles from Ground-Based Zenith-Sky UV-Visible Observations: Validation of the Technique through Correlative Comparisons". In: *Atmospheric Chemistry and Physics* 4.8, pp. 2091–2106. DOI: 10.5194/acp-4-2091-2004. URL: http://www.atmos-chem-phys.net/4/2091/2004/acp-4-2091-2004.pdf (visited on 09/12/2017) (cit. on pp. 2, 106).

HENYEY, L. G. and J. L. GREENSTEIN (**1941**). "Diffuse Radiation in the Galaxy". In: *The Astrophysical Journal* 93, pp. 70–83. ISSN: 0004-637X. DOI: 10.1086/144246. URL: http://adsabs.harvard.edu/abs/1941ApJ...93...70H (visited on 09/12/2017) (cit. on pp. 29, 109).

HERMAN, Jay, Alexander CEDE, Elena SPINEI, George MOUNT, Maria TZORTZIOU, and Nader ABUHASSAN (**2009**). "NO_2 Column Amounts from Ground-Based Pandora and MFDOAS Spectrometers Using the Direct-Sun DOAS Technique: Intercomparisons and Application to OMI Validation". In: *Journal of Geophysical Research: Atmospheres* 114.D13, p. D13307. ISSN: 2156-2202. DOI: 10.1029/2009JD011848. URL: http://onlinelibrary.wiley.com/doi/10.1029/2009JD011848/abstract (visited on 02/12/2018) (cit. on p. 173).

HILBOLL, A., A. RICHTER, and J. P. BURROWS (**2013**). "Long-Term Changes of Tropospheric NO_2 over Megacities Derived from Multiple Satellite Instruments". In: *Atmospheric Chemistry and Physics* 13.8, pp. 4145–4169. ISSN: 1680-7316. DOI: 10/gb8rfx. URL: https://www.atmos-chem-phys.net/13/4145/2013/acp-13-4145-2013.html (visited on 11/02/2018) (cit. on p. 44).

HOLBEN, B. N., T. F. ECK, I. SLUTSKER, D. TANRÉ, J. P. BUIS, A. SETZER, E. VERMOTE, J. A. REAGAN, Y. J. KAUFMAN, T. NAKAJIMA, F. LAVENU, I. JANKOWIAK, and A. SMIRNOV (**1998**). "AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization". In: *Remote Sensing of Environment* 66.1, pp. 1–16. ISSN: 0034-4257. DOI: 10.1016/S0034-4257(98)00031-5. URL: http://www.sciencedirect.com/science/article/pii/S0034425798000315 (visited on 02/07/2018) (cit. on pp. 63, 169, 225).

HOLLOWAY, Ann M. and Richard P. WAYNE (**2015**). *Atmospheric Chemistry*. Royal Society of Chemistry. 286 pp. ISBN: 978-1-78262-593-3 (cit. on pp. 8, 13, 41).

HÖNNINGER, G., C. von FRIEDEBURG, and U. PLATT (**2004**). "Multi Axis Differential Optical Absorption Spectroscopy (MAX-DOAS)". In: *Atmospheric Chemistry and Physics* 4.1, pp. 231–254. DOI: 10.5194/acp-4-231-2004. URL: http://www.atmos-chem-phys.net/4/231/2004/acp-4-231-2004.pdf (visited on 09/12/2017) (cit. on p. 39).

HÖNNINGER, Gerd and U. PLATT (**2002**). "Observations of BrO and Its Vertical Distribution during Surface Ozone Depletion at Alert". In: *Atmospheric Environment*. Air/Snow/Ice Interactions in the Arctic: Results from ALERT 2000 and SUMMIT 2000 36.15, pp. 2481–2489. ISSN: 1352-2310. DOI: 10/bkjbwh. URL: http://www.sciencedirect.com/science/article/pii/S1352231002001048 (visited on 05/06/2018) (cit. on pp. 2, 105).

IPCC, 2013A (**2013**). *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Groups I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. In collab. with M. WILD, C. SCHÄR, N. LOEB, E. G. DUTTON, and G. KÖNIG-LANGLO. Cambridge, UK, New York, USA: Cambridge University Press (cit. on p. 12).

JACOB, Daniel (**1999**). *Introduction to Atmospheric Chemistry*. Princeton University Press. ISBN: 978-1-4008-4154-7 (cit. on pp. 41–43).

JACOBSON, Mark Z. (**2005**). *Fundamentals of Atmospheric Modeling*. Cambridge university press. ISBN: 978-1-139-16538-9 (cit. on p. 31).

JUNKERMANN, W. (**2009**). "On the Distribution of Formaldehyde in the Western Po-Valley, Italy, during FORMAT 2002/2003". In: *Atmospheric Chemistry and Physics* 9.23, pp. 9187–9196. ISSN: 1680-7324. DOI: 10/c59s2k. URL: https://www.atmos-chem-phys.net/9/9187/2009/ (visited on 06/11/2018) (cit. on p. 85).

KANAKIDOU, Maria, Robert A. DUCE, Joseph M. PROSPERO, Alex R. BAKER, Claudia BENITEZ-NELSON, Frank J. DENTENER, Keith A. HUNTER, Peter S. LISS, Natalie MAHOWALD, Gregory S. OKIN, Manmohan SARIN, Kostas TSIGARIDIS, Mitsuo UEMATSU, Lauren M. ZAMORA, and Tong ZHU (**2012**). "Atmospheric Fluxes of Organic N and P to the Global Ocean". In: *Global Biogeochemical Cycles* 26.3. ISSN: 1944-9224. DOI: 10/gdxw3g. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GB004277 (visited on 08/10/2018) (cit. on p. 100).

KATTAWAR, G. W., A. T. YOUNG, and T. J. HUMPHREYS (**1981**). "Inelastic Scattering in Planetary Atmospheres. I - The Ring Effect, without Aerosols". In: *The Astrophysical Journal* 243, pp. 1049–1057. ISSN: 0004-637X. DOI: 10.1086/158669. URL: http://adsabs.harvard.edu/abs/1981ApJ...243.1049K (visited on 03/04/2018) (cit. on p. 31).

KEBABIAN, Paul L., Ezra C. WOOD, Scott C. HERNDON, and Andrew FREEDMAN (**2008**). "A Practical Alternative to Chemiluminescence-Based Detection of Nitrogen Dioxide: Cavity Attenuated Phase Shift Spectroscopy". In: *Environmental Science & Technology* 42.16, pp. 6040–6045. ISSN: 0013-936X. DOI: 10/fs8pg7. URL: https://doi.org/10.1021/es703204j (visited on 05/06/2018) (cit. on p. 174).

KHAN, Waqas Ahmed, Muhammad Fahim KHOKHAR, Asadullah SHOAIB, and Rab NAWAZ (**2018**). "Monitoring and Analysis of Formaldehyde Columns over Rawalpindi-Islamabad, Pakistan Using MAX-DOAS and Satellite Observation". In: *Atmospheric Pollution Research*. ISSN: 1309-1042. DOI: 10/gdkxxq. URL: http://www.sciencedirect.com/science/article/pii/S1309104217303732 (visited on 06/04/2018) (cit. on pp. 82, 206).

KNEPP, T., M. PIPPIN, J. CRAWFORD, G. CHEN, J. SZYKMAN, R. LONG, L. COWEN, A. CEDE, N. ABUHASSAN, J. HERMAN, R. DELGADO, J. COMPTON, T. BERKOFF, J. FISHMAN, D. MARTINS, R. STAUFFER, A. M. THOMPSON, A. WEINHEIMER, D. KNAPP, D. MONTZKA, D. LENSCHOW, and D. NEIL (**2015**). "Estimating Surface NO_2 and SO_2 Mixing Ratios from Fast-Response Total Column Observations and Potential Application to Geostationary Missions". In: *Journal of Atmospheric Chemistry* 72.3-4, pp. 261–286. ISSN: 0167-7764, 1573-0662. DOI: 10.1007/s10874-013-9257-6. URL: https://link.springer.com/article/10.1007/s10874-013-9257-6 (visited on 02/12/2018) (cit. on p. 173).

KOELEMEIJER, R. B. A., J. F. de HAAN, and P. STAMMES (**2003**). "A Database of Spectral Surface Reflectivity in the Range 335–772 Nm Derived from 5.5 Years of GOME Observations". In: *Journal of Geophysical Research: Atmospheres* 108.D2. ISSN: 0148-0227. DOI: 10/chkvvx. URL: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2002JD002429 (visited on 03/22/2018) (cit. on p. 63).

KRAMER, Louisa J., Roland J. LEIGH, John J. REMEDIOS, and Paul S. MONKS (**2008**). "Comparison of OMI and Ground-Based in Situ and MAX-DOAS Measurements of Tropospheric Nitrogen Dioxide in an Urban Area". In: *Journal of Geophysical Research: Atmospheres* 113.D16, D16S39. ISSN: 2156-2202. DOI: 10.1029/2007JD009168. URL: http://onlinelibrary.wiley.com/doi/10.1029/2007JD009168/abstract (visited on 01/08/2018) (cit. on pp. 78, 201).

KURUCZ, R. L. (**1984**). "Solar Flux Atlas from 296 to 1300 Nm". In: *National Solar Observatory Atlas* 1. URL: https://ci.nii.ac.jp/naid/10003938930/ (visited on 03/17/2018) (cit. on p. 57).

LATIF, Mojib (**2009**). *Klimawandel Und Klimadynamik*. Stuttgart: UTB Ulmer. 220 pp. ISBN: 978-3-8252-3178-1. URL: http://eprints.uni-kiel.de/6962/ (visited on 03/01/2018) (cit. on pp. 11, 13).

LEE, Hanlim, H. IRIE, Young J. KIM, Youngmin NOH, Chulkyu LEE, Yeosook KIM, and Kee J. CHUN (**2009**). "Retrieval of Aerosol Extinction in the Lower Troposphere Based on UV MAX-DOAS Measurements". In: *Aerosol Science and Technology* 43.5, pp. 502–509. ISSN: 0278-6826. DOI: 10.1080/02786820902769691. URL: http://dx.doi.org/10.1080/02786820902769691 (visited on 09/13/2017) (cit. on p. 106).

LEIGHTON, Philip (**1961**). *Photochemistry of Air Pollution*. Elsevier. 313 pp. ISBN: 978-0-323-15645-5 (cit. on p. 43).

LESER, H., G. HÖNNINGER, and U. PLATT (**2003**). "MAX-DOAS Measurements of BrO and NO_2 in the Marine Boundary Layer". In: *Geophysical Research Letters* 30.10. ISSN: 0094-8276. DOI: 10/fjm8bv. URL: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2002GL015811 (visited on 05/06/2018) (cit. on pp. 2, 105).

LI, X., T. BRAUERS, M. SHAO, R. M. GARLAND, T. WAGNER, T. DEUTSCHMANN, and A. WAHNER (**2010**). "MAX-DOAS Measurements in Southern China: Retrieval of Aerosol Extinctions and Validation Using Ground-Based in-Situ Data". In: *Atmospheric Chemistry and Physics* 10.5, pp. 2079–2089. ISSN: 1680-7324. DOI: 10.5194/acp-10-2079-2010. URL: https://www.atmos-chem-phys.net/10/2079/2010/ (visited on 09/13/2017) (cit. on pp. 106, 197).

LIOU, Kuo-Nan (**2002**). *An Introduction to Atmospheric Radiation*. Vol. 84. Elsevier Science. ISBN: 978-0-08-049167-7 (cit. on pp. 19, 31).

MA, J. Z., S. BEIRLE, J. L. JIN, R. SHAIGANFAR, P. YAN, and T. WAGNER (**2013**). "Tropospheric NO_2 Vertical Column Densities over Beijing: Results of the First Three Years of Ground-Based MAX-DOAS Measurements (2008–2011) and Satellite Validation". In: *Atmospheric Chemistry and Physics* 13.3, pp. 1547–1567. ISSN: 1680-7316. DOI: 10/f4qdfx. URL: https://www.atmos-chem-phys.net/13/1547/2013/acp-13-1547-2013.html (visited on 04/17/2018) (cit. on pp. 78, 81, 201).

MAD-CAT (**2013**). *Satellite Group Mainz*. URL: http://joseba.mpch-mainz.mpg.de/mad_cat.htm (visited on 05/23/2018) (cit. on p. 74).

MAMALI, D., J. MIKKILÄ, B. HENZING, R. SPOOR, M. EHN, T. PETÄJÄ, H. RUSSCHENBERG, and G. BISKOS (**2018**). "Long-Term Observations of the Background Aerosol at Cabauw, The Netherlands". In: *Science of The Total Environment* 625, pp. 752–761. ISSN: 0048-9697. DOI: 10.1016/j.scitotenv.2017.12.136. URL: https:

//www.sciencedirect.com/science/article/pii/S0048969717335623 (visited on 01/09/2018) (cit. on p. 198).

MARQUARD, L. C., T. WAGNER, and U. PLATT (**2000**). "Improved Air Mass Factor Concepts for Scattered Radiation Differential Optical Absorption Spectroscopy of Atmospheric Species". In: *Journal of Geophysical Research: Atmospheres* 105.D1, pp. 1315–1327. ISSN: 2156-2202. DOI: 10.1029/1999JD900340. URL: http://onlinelibrary.wiley.com/doi/10.1029/1999JD900340/abstract (visited on 03/06/2018) (cit. on p. 34).

MELLER, R and G. K. MOORTGAT (**2000**). "Temperature Dependence of the Absorption Cross Sections of Formaldehyde between 223 and 323 K in the Wavelength Range 225–375 Nm". In: *Journal of Geophysical Research: Atmospheres* 105.D6, pp. 7089–7101. ISSN: 0148-0227. DOI: 10/chd6dh. URL: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/1999JD901074 (visited on 03/20/2018) (cit. on pp. 46, 63).

MENDOLIA, D., R. J. C. D'SOUZA, G. J. EVANS, and J. BROOK (**2013**). "Comparison of Tropospheric NO2 Vertical Columns in an Urban Environment Using Satellite, Multi-Axis Differential Optical Absorption Spectroscopy, and in Situ Measurements". In: *Atmospheric Measurement Techniques* 6.10, pp. 2907–2924. ISSN: 1867-8548. DOI: 10/gb8shv. URL: https://www.atmos-meas-tech.net/6/2907/2013/ (visited on 06/04/2018) (cit. on pp. 78, 201, 202).

MOLINA, Mario J. and F. S. ROWLAND (**1974**). "Stratospheric Sink for Chlorofluoromethanes: Chlorine Atom-Catalysed Destruction of Ozone". In: *Nature* 249.5460, pp. 810–812. ISSN: 1476-4687. DOI: 10/frg72c. URL: https://www.nature.com/articles/249810a0 (visited on 09/14/2018) (cit. on p. 1).

NASA (**1976**). *U.S. Standard Atmosphere, 1976*. Washington, D.C.: Office, U.S. Government Printing. URL: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19770009539.pdf (visited on 01/09/2018) (cit. on pp. 88, 110, 114, 184).

– (**2018**). *AVDC | DATA - AURA - NASA EOS -Aura Validation Data Center - Correlative Data, Field of View Predictions, Data Subsets*. URL: https://avdc.gsfc.nasa.gov/index.php?site=2045907950 (visited on 02/12/2018) (cit. on pp. 64, 173).

NOXON, J. F. (**1975**). "Nitrogen Dioxide in the Stratosphere and Troposphere Measured by Ground-Based Absorption Spectroscopy". In: *Science* 189.4202, pp. 547–549. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.189.4202.547. pmid: 17798301. URL: http://science.sciencemag.org/content/189/4202/547 (visited on 09/12/2017) (cit. on pp. 2, 34).

ORTEGA, I., T. KOENIG, R. SINREICH, D. THOMSON, and R. VOLKAMER (**2015**). "The CU 2-D-MAX-DOAS Instrument – Part 1: Retrieval of 3-D Distributions of NO_2 and Azimuth-Dependent OVOC Ratios". In: *Atmospheric Measurement Techniques* 8.6, pp. 2371–2395. ISSN: 1867-8548. DOI: 10.5194/amt-8-2371-2015. URL: http://www.atmos-meas-tech.net/8/2371/2015/ (visited on 09/12/2017) (cit. on pp. 86, 87, 91).

PERNER, D., D. H. EHHALT, H. W. PÄTZ, U. PLATT, E. P. RÖTH, and A. VOLZ (**1976**). "OH - Radicals in the Lower Troposphere". In: *Geophysical Research Letters* 3.8, pp. 466–468. ISSN: 1944-8007. DOI: 10/fhtggw. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GL003i008p00466 (visited on 09/14/2018) (cit. on p. 2).

PETERS, E., G. PINARDI, A. SEYLER, A. RICHTER, F. WITTROCK, T. BÖSCH, M. VAN ROOZENDAEL, F. HENDRICK, T. DROSOGLOU, A. F. BAIS, Y. KANAYA, X. ZHAO, K. STRONG, J. LAMPEL, R. VOLKAMER, T. KOENIG, I. ORTEGA, O. PUENTEDURA, M. NAVARRO-COMAS, L. GÓMEZ, M. YELA GONZÁLEZ, A. PITERS, J. REMMERS, Y. WANG, T. WAGNER, S. WANG, A. SAIZ-LOPEZ, D. GARCÍA-NIETO, C. A. CUEVAS, N. BENAVENT, R. QUEREL, P. JOHNSTON, O. POSTYLYAKOV, A. BOROVSKI, A. ELOKHOV, I. BRUCHKOUSKI, H. LIU, C. LIU, Q. HONG, C. RIVERA, M. GRUTTER, W. STREMME, M. F. KHOKHAR, J. KHAYYAM, and J. P. BURROWS (**2017**). "Investigating Differences in DOAS Retrieval Codes Using MAD-CAT Campaign Data". In: *Atmospheric Measurement Techniques* 10.3, pp. 955–978. ISSN: 1867-8548. DOI: 10/gbpwmh. URL: https://www.atmos-meas-tech.net/10/955/2017/ (visited on 05/23/2018) (cit. on p. 74).

PETERS, E., F. WITTROCK, K. GROSSMANN, U. FRIESS, A. RICHTER, and J. P. BURROWS (**2012**). "Formaldehyde and Nitrogen Dioxide over the Remote Western Pacific Ocean: SCIAMACHY and GOME-2 Validation Using Ship-Based MAX-DOAS Observations". In: *Atmospheric Chemistry and Physics* 12.22, pp. 11179–11197. ISSN: 1680-7324. DOI: 10/f4hqqk. URL: https://www.atmos-chem-phys.net/12/11179/2012/ (visited on 03/23/2018) (cit. on pp. 66, 273).

PETERS, Enno (**2013**). "Improved MAX-DOAS Measurements and Retrievals Focused on the Marine Boundary Layer". Phd thesis. University of Bremen. URL: http://nbn-resolving.de/urn:nbn:de:gbv:46-00103447-17 (cit. on p. 54).

PETTY, G.W. (**2006**). *A First Course in Atmospheric Radiation*. Madison, Wisconsin, USA: Sundog publishing. ISBN: 978-0-9729033-1-8. URL: https://sundogpublishingstore.myshopify.com/products/a-first-course-in-atmospheric-radiation-g-w-petty (visited on 03/01/2018) (cit. on pp. 19, 28, 31).

PINARDI, G., M. VAN ROOZENDAEL, N. ABUHASSAN, C. ADAMS, A. CEDE, K. CLÉMER, C. FAYT, U. FRIESS, M. GIL, J. HERMAN, C. HERMANS, F. HENDRICK, H. IRIE, A. MERLAUD, M. NAVARRO COMAS, E. PETERS, A. J. M. PITERS, O. PUENTEDURA, A. RICHTER, A. SCHÖNHARDT, R. SHAIGANFAR, E. SPINEI, K. STRONG, H. TAKASHIMA, M. VREKOUSSIS, T. WAGNER, F. WITTROCK, and S. YILMAZ (**2013**). "MAX-DOAS Formaldehyde Slant Column Measurements during CINDI: Intercomparison and Analysis Improvement". In: *Atmospheric Measurement Techniques* 6.1, pp. 167–185. ISSN: 1867-8548. DOI: 10.5194/amt-6-167-2013. URL: https://www.atmos-meas-tech.net/6/167/2013/ (visited on 09/12/2017) (cit. on pp. 3, 83, 167, 208).

PITERS, A. J. M., K. F. BOERSMA, M. KROON, J. C. HAINS, M. VAN ROOZENDAEL, F. WITTROCK, N. ABUHASSAN, C. ADAMS, M. AKRAMI, M. A. F. ALLAART, A. APITULEY, S. BEIRLE, J. B. BERGWERFF, A. J. C. BERKHOUT, D. BRUNNER, A. CEDE, J. CHONG, K. CLÉMER, C. FAYT, U. FRIESS, L. F. L. GAST, M. GIL-OJEDA, F. GOUTAIL, R. GRAVES, A. GRIESFELLER, K. GROSSMANN, G. HEMERIJCKX, F. HENDRICK, B. HENZING, J. HERMAN, C. HERMANS, M. HOEXUM, G. R. van der HOFF, H. IRIE, P. V. JOHNSTON, Y. KANAYA, Y. J. KIM, H. KLEIN BALTINK, K. KREHER, G. de LEEUW, R. LEIGH, A. MERLAUD, M. M. MOERMAN, P. S. MONKS, G. H. MOUNT, M. NAVARRO-COMAS, H. OETJEN, A. PAZMINO, M. PEREZ-CAMACHO, E. PETERS, A. du PIESANIE, G. PINARDI, O. PUENTEDURA, A. RICHTER, H. K. ROSCOE, A. SCHÖNHARDT, B. SCHWARZENBACH, R. SHAIGANFAR, W. SLUIS, E. SPINEI, A. P. STOLK, K. STRONG, D. P. J. SWART, H. TAKASHIMA, T. VLEMMIX, M. VREKOUSSIS, T. WAGNER, C. WHYTE, K. M. WILSON, M. YELA, S. YILMAZ, P. ZIEGER, and Y. ZHOU (**2012**). "The Cabauw Intercomparison Campaign for Nitrogen Dioxide Measuring Instruments (CINDI): Design, Execution, and Early Results". In: *Atmospheric Measurement Techniques* 5.2, pp. 457–485. ISSN: 1867-8548. DOI: 10.5194/amt-5-457-2012. URL: https://www.atmos-meas-tech.net/5/457/2012/ (visited on 02/07/2018) (cit. on p. 167).

PLATT, U. and D. PERNER (**1980**). "Direct Measurements of Atmospheric CH_2O, HNO_2, O_2 NO_2, and SO_2 by Differential Optical Absorption in the near UV". In: *Journal of Geophysical Research: Oceans* 85.C12, pp. 7453–7458. ISSN: 2156-2202. DOI: 10/d6v39p. URL: http://onlinelibrary.wiley.com/doi/10.1029/JC085iC12p07453/abstract (visited on 03/07/2018) (cit. on pp. 2, 34).

PLATT, U. and J. STUTZ (**2008**). *Differential Optical Absorption Spectroscopy: Principles and Applications*. Physics of Earth and space environments. Berlin: Springer. 597 pp. ISBN: 978-3-540-75776-4 (cit. on pp. 19, 30, 34, 66).

PÖHLER, Denis, Erik LUTZ, Martin HORBANSKI, Johannes LAMPEL, and Ulrich PLATT (**2017**). "High Precision NO_2 and NO Measurements with the ICAD Instrument during S-b-s Campaign Hohenpeißenberg 2016". In: EGU General Assembly Conference Abstracts. Vol. 19, p. 14336. URL: http://adsabs.harvard.edu/abs/2017EGUGA..1914336P (visited on 02/12/2018) (cit. on p. 174).

PÖHLER, Denis, Leif VOGEL, Udo FRIESS, and Ulrich PLATT (**2010**). "Observation of Halogen Species in the Amundsen Gulf, Arctic, by Active Long-Path Differential Optical Absorption Spectroscopy". In: *Proceedings of the National Academy of Sciences* 107.15, pp. 6582–6587. ISSN: 0027-8424, 1091-6490. DOI: 10/cqnpj4. pmid: 20160121. URL: http://www.pnas.org/content/107/15/6582 (visited on 05/06/2018) (cit. on p. 174).

POMMEREAU, Jean Pierre and Florence GOUTAIL (**1988**). "O3 and NO2 Ground-Based Measurements by Visible Spectrometry during Arctic Winter and Spring 1988". In: *Geophysical Research Letters*, pp. 891–894. ISSN: 1944-8007. DOI: 10.1029/GL015i008p00891@10.1002/(ISSN)1944-8007.POLOZ1. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/GL015i008p00891%4010.1002/%28ISSN%291944-8007.POLOZ1 (visited on 12/14/2018) (cit. on p. 2).

PRESS, William H (**2007**). *Numerical Recipes: The Art of Scientific Computing*. OCLC: 212427139. Cambridge, UK; New York: Cambridge University Press. ISBN: 978-0-511-33555-6 (cit. on p. 58).

PRICE, S.D. and R.G. WALKER (**1976**). *The AFGL Four Color Infrared Sky Survey: Catalog of Observations at 4. 2, 11.0, 19.8 and 27.4 Micrometers*. URL: http://www.dtic.mil/docs/citations/ADA034448 (visited on 03/22/2018) (cit. on p. 63).

PUĶĪTE, J., S. KÜHL, T. DEUTSCHMANN, U. PLATT, and T. WAGNER (**2010**). "Extending Differential Optical Absorption Spectroscopy for Limb Measurements in the UV". In: *Atmospheric Measurement Techniques* 3.3, pp. 631–653. ISSN: 1867-8548. DOI: 10.5194/amt-3-631-2010. URL: http://www.atmos-meas-tech.net/3/631/2010/ (visited on 09/12/2017) (cit. on pp. 34, 57).

RAITH, Wilhelm and Siegfried J. BAUER, eds. (**2001**). *Erde und Planeten*. 2., aktualisierte Aufl. Lehrbuch der Experimentalphysik 7. Berlin: de Gruyter. 727 pp. ISBN: 978-3-11-016837-2 (cit. on pp. 9, 13).

RAYLEIGH, Robert John Strutt (**1899**). "XXXIV. On the Transmission of Light through an Atmosphere Containing Small Particles in Suspension, and on the Origin of the Blue of the Sky". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 47.287, pp. 375–384. ISSN: 1941-5982. DOI: 10.1080/14786449908621276. URL: https://doi.org/10.1080/14786449908621276 (visited on 03/01/2018) (cit. on pp. 25, 27).

RICHTER, A., M. EISINGER, A. LADSTÄTTER-WEISSENMAYER, and J.P. BURROWS (**1999**). "DOAS Zenith Sky Observations: 2. Seasonal Variation of BrO Over Bremen (53°N) 1994-1995". In: *Journal of Atmospheric Chemistry* 32.1, pp. 83–99. ISSN: 1573-0662. DOI: 10/dg379k. URL: https://doi.org/10.1023/A:1006077725894 (visited on 09/14/2018) (cit. on p. 2).

RICHTER, Andreas (**1997**). "Absorptionsspektroskopische Messungen Stratosphaerischer Spurengase Ueber Bremen, 53°N". Phd thesis. Bremen: University of Bremen. URL: https://mm.suub.uni-bremen.de/peid=B23948594&LAN=DE&CID=6225584&index=L&Hitnr=3&dtyp=h&rtyp=a&Exemplar=1%0A%09%09%09%09%09%09%09 (cit. on pp. 34, 56, 57, 60).

RODGERS, Clive D. (**2004**). *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Reprinted. Series on atmospheric oceanic and planetary physics 2. OCLC: 254137862. Singapore: World Scientific. 238 pp. ISBN: 978-981-02-2740-1 (cit. on pp. 106, 111–113, 115, 127, 128, 133, 134, 157).

RODGERS, Clive D. and Brian J. CONNOR (**2003**). "Intercomparison of Remote Sounding Instruments". In: *Journal of Geophysical Research: Atmospheres* 108.D3, p. 4116. ISSN: 2156-2202. DOI: 10.1029/2002JD002299. URL: http://onlinelibrary.wiley.com/doi/10.1029/2002JD002299/abstract (visited on 01/11/2018) (cit. on p. 172).

ROSCOE, H. K., P. V. JOHNSTON, M. VAN ROOZENDAEL, A. RICHTER, A. SARKISSIAN, J. ROSCOE, K. E. PRESTON, J-C. LAMBERT, C. HERMANS, W. DECUYPER, S. DZIENUS, T. WINTERRATH, J. BURROWS, F. GOUTAIL, J-P. POMMEREAU, E. D'ALMEIDA, J. HOTTIER, C. COUREUL, R. DIDIER, I. PUNDT, L. M. BARTLETT, C. T. MCELROY, J. E. KERR, A. ELOKHOV, G. GIOVANELLI, F. RAVEGNANI, M. PREMUDA, I. KOSTADINOV, F. ERLE, T. WAGNER, K. PFEILSTICKER, M. KENNTNER, L. C. MARQUARD, M. GIL, O. PUENTEDURA, M. YELA, D. W. ARLANDER, B. A. KASTAD HOISKAR, C. W. TELLEFSEN, K. KARLSEN TORNKVIST, B. HEESE, R. L. JONES, S. R. ALIWELL, and R. A. FRESHWATER (**1999**). "Slant Column Measurements of O3 and NO2 During the NDSC Intercomparison of Zenith-Sky UV-Visible Spectrometers in June 1996". In: *Journal of Atmospheric Chemistry* 32.2, pp. 281–314. ISSN: 1573-0662. DOI: 10/b54gdz. URL: https://doi.org/10.1023/A:1006111216966 (visited on 09/14/2018) (cit. on p. 2).

ROTHMAN, L. S., I. E. GORDON, R. J. BARBER, H. DOTHE, R. R. GAMACHE, A. GOLDMAN, V. I. PEREVALOV, S. A. TASHKUN, and J. TENNYSON (**2010**). "HITEMP, the High-Temperature Molecular Spectroscopic Data-

base". In: *Journal of Quantitative Spectroscopy and Radiative Transfer*. XVIth Symposium on High Resolution Molecular Spectroscopy (HighRus-2009) 111.15, pp. 2139–2150. ISSN: 0022-4073. DOI: 10/c93s6j. URL: http://www.sciencedirect.com/science/article/pii/S002240731000169X (visited on 03/20/2018) (cit. on p. 63).

ROZANOV, V. V. and A. V. ROZANOV (**2010**). "Differential Optical Absorption Spectroscopy (DOAS) and Air Mass Factor Concept for a Multiply Scattering Vertically Inhomogeneous Medium: Theoretical Consideration". In: *Atmospheric Measurement Techniques* 3.3, pp. 751–780. ISSN: 1867-8548. DOI: 10.5194/amt-3-751-2010. URL: http://www.atmos-meas-tech.net/3/751/2010/ (visited on 09/12/2017) (cit. on pp. 34, 110).

ROZANOV, V. V., A. V. ROZANOV, A. A. KOKHANOVSKY, and J. P. BURROWS (**2014**). "Radiative Transfer through Terrestrial Atmosphere and Ocean: Software Package SCIATRAN". In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 133, pp. 13–71. ISSN: 0022-4073. DOI: 10.1016/j.jqsrt.2013.07.004. URL: http://www.sciencedirect.com/science/article/pii/S0022407313002872 (visited on 03/06/2018) (cit. on pp. 33, 34, 107).

SANDERS, R. W. (**1996**). "Improved Analysis of Atmospheric Absorption Spectra by Including the Temperature Dependence of NO2". In: *Journal of Geophysical Research: Atmospheres* 101.D15, pp. 20945–20952. ISSN: 2156-2202. DOI: 10/b9s5v3. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD01699 (visited on 07/22/2018) (cit. on p. 57).

SAYER, A. M., N. C. HSU, C. BETTENHAUSEN, and M.-J. JEONG (**2013**). "Validation and Uncertainty Estimates for MODIS Collection 6 "Deep Blue" Aerosol Data". In: *Journal of Geophysical Research: Atmospheres* 118.14, pp. 7864–7872. ISSN: 2169-8996. DOI: 10.1002/jgrd.50600. URL: http://onlinelibrary.wiley.com/doi/10.1002/jgrd.50600/abstract (visited on 02/10/2018) (cit. on p. 169).

SEINFELD, John H. and Spyros N. PANDIS (**2006**). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. 2nd ed. OCLC: ocm62493628. Hoboken, N.J: J. Wiley. 1203 pp. ISBN: 978-0-471-72017-1 (cit. on pp. 8, 41, 45).

SERDYUCHENKO, A., V. GORSHELEV, M. WEBER, W. CHEHADE, and J. P. BURROWS (**2014**). "High Spectral Resolution Ozone Absorption Cross-Sections – Part 2: Temperature Dependence". In: *Atmospheric Measurement Techniques* 7.2, pp. 625–636. ISSN: 1867-8548. DOI: 10/gb8szw. URL: https://www.atmos-meas-tech.net/7/625/2014/ (visited on 03/20/2018) (cit. on p. 63).

SHAIGANFAR, R., S. BEIRLE, M. SHARMA, A. CHAUHAN, R. P. SINGH, and T. WAGNER (**2011**). "Estimation of NO_x Emissions from Delhi Using Car MAX-DOAS Observations and Comparison with OMI Satellite Data". In: *Atmospheric Chemistry and Physics* 11.21, pp. 10871–10887. ISSN: 1680-7324. DOI: 10.5194/acp-11-10871-2011. URL: http://www.atmos-chem-phys.net/11/10871/2011/ (visited on 09/12/2017) (cit. on p. 64).

SHANNON, C. E. (**2001**). "A Mathematical Theory of Communication". In: *SIGMOBILE Mobile Computing and Communications Review* 5.1, pp. 3–55. ISSN: 1559-1662. DOI: 10/d5szst. URL: http://doi.acm.org/10.1145/584091.584093 (visited on 07/06/2018) (cit. on p. 116).

SINREICH, R., U. FRIESS, T. WAGNER, and U. PLATT (**2005**). "Multi Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) of Gas and Aerosol Distributions". In: *Faraday Discussions* 130, p. 153. ISSN: 1359-6640, 1364-5498. DOI: 10.1039/b419274p. URL: http://xlink.rsc.org/?DOI=b419274p (visited on 09/12/2017) (cit. on p. 106).

SINREICH, R., A. MERTEN, L. MOLINA, and R. VOLKAMER (**2013**). "Parameterizing Radiative Transfer to Convert MAX-DOAS dSCDs into near-Surface Box-Averaged Mixing Ratios". In: *Atmospheric Measurement Techniques* 6.6, pp. 1521–1532. ISSN: 1867-8548. DOI: 10.5194/amt-6-1521-2013. URL: http://www.atmos-meas-tech.net/6/1521/2013/ (visited on 09/12/2017) (cit. on pp. 85–87, 90, 106).

SOLOMON, Susan (**1988**). "The Mystery of the Antarctic Ozone "Hole"". In: *Reviews of Geophysics* 26.1, pp. 131–148. ISSN: 1944-9208. DOI: 10/dq39gr. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG026i001p00131 (visited on 09/14/2018) (cit. on p. 1).

SOLOMON, Susan, Arthur L. SCHMELTEKOPF, and Ryan W. SANDERS (**1987**). "On the Interpretation of Zenith Sky Absorption Measurements". In: *Journal of Geophysical Research: Atmospheres* 92.D7, pp. 8311–8319. ISSN: 2156-2202. DOI: 10.1029/JD092iD07p08311. URL: http://onlinelibrary.wiley.com/doi/10.1029/JD092iD07p08311/abstract (visited on 09/12/2017) (cit. on pp. 34, 39).

SPURR, Robert (**2008**). "LIDORT and VLIDORT: Linearized Pseudo-Spherical Scalar and Vector Discrete Ordinate Radiative Transfer Models for Use in Remote Sensing Retrieval Problems". In: *Light Scattering Reviews 3*. Springer Praxis Books. Springer, Berlin, Heidelberg, pp. 229–275. ISBN: 978-3-540-48546-9. DOI: 10.1007/978-3-540-48546-9_7. URL: https://link.springer.com/chapter/10.1007/978-3-540-48546-9_7 (visited on 03/22/2018) (cit. on p. 63).

STAVRAKOU, T., J.-F. MÜLLER, I. De SMEDT, M. Van ROOZENDAEL, M. KANAKIDOU, M. VREKOUSSIS, F. WITTROCK, A. RICHTER, and J. P. BURROWS (**2009**). "The Continental Source of Glyoxal Estimated by the Synergistic Use of Spaceborne Measurements and Inverse Modelling". In: *Atmospheric Chemistry and Physics* 9.21, pp. 8431–8446. ISSN: 1680-7316. DOI: 10/ftj8bm. URL: https://www.atmos-chem-phys.net/9/8431/2009/acp-9-8431-2009.html (visited on 09/19/2018) (cit. on p. 45).

STOCKER, Thomas, ed. (**2014**). *Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. New York: Cambridge University Press. 1535 pp. ISBN: 978-1-107-66182-0 (cit. on p. 14).

THALMAN, Ryan and Rainer VOLKAMER (**2013**). "Temperature Dependent Absorption Cross-Sections of O_2 –O_2 Collision Pairs between 340 and 630 Nm and at Atmospherically Relevant Pressure". In: *Physical Chemistry Chemical Physics* 15.37, pp. 15371–15381. DOI: 10/gc5s9g. URL: http://pubs.rsc.org/en/Content/ArticleLanding/2013/CP/C3CP50968K (visited on 03/20/2018) (cit. on p. 63).

TIAN, Xin, Pinhua XIE, Jin XU, Ang LI, Yang WANG, Min QIN, and Zhaokun HU (**2018**). "Long-Term Observations of Tropospheric NO2, SO2 and HCHO by MAX-DOAS in Yangtze River Delta Area, China". In: *Journal of Environmental Sciences*. ISSN: 1001-0742. DOI: 10/gdkxw9. URL: http://www.sciencedirect.com/science/article/pii/S1001074217329613 (visited on 06/04/2018) (cit. on pp. 78, 81–83, 206).

TOLEDANO, C., E.V. CACHORRO, A. BERJON, A.M. de FRUTOS, M. SORRIBAS, A.B. de la MORENA, and P. GOLOUB (**2007**). "Aerosol Optical Depth and Ångström Exponent Climatology at El Arenosillo AERONET Site (Huelva, Spain)". In: *Quarterly Journal of the Royal Meteorological Society* 133.624, pp. 795–807. ISSN: 0035-9009. DOI: 10/b57xrq. URL: https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/qj.54 (visited on 04/25/2018) (cit. on p. 49).

TOUBLANC, Dominique (**1996**). "Henyey–Greenstein and Mie Phase Functions in Monte Carlo Radiative Transfer Computations". In: *Applied Optics* 35.18, pp. 3270–3274. ISSN: 2155-3165. DOI: 10.1364/AO.35.003270. URL: https://www.osapublishing.org/abstract.cfm?uri=ao-35-18-3270 (visited on 09/12/2017) (cit. on p. 30).

Van ROOZENDAEL, M., C. FAYT, P. POST, C. HERMANS, and J. C. LAMBERT (**2004**). "Retrieval of Tropospheric BrO and NO2 from UV-Visible Observations". In: *Sounding the Troposphere from Space*. 10.1007/978-3-642-18875-6_19. Springer, Berlin, Heidelberg, pp. 155–165. ISBN: 978-3-642-18875-6. URL: https://link.springer.com/chapter/10.1007/978-3-642-18875-6_19 (visited on 05/06/2018) (cit. on pp. 2, 105).

VANDAELE, A. C., C. HERMANS, P. C. SIMON, M. van ROOZENDAEL, J. M. GUILMOT, M. CARLEER, and R. COLIN (**1996**). "Fourier Transform Measurement of NO2 Absorption Cross-Section in the Visible Range at Room Temperature". In: *Journal of Atmospheric Chemistry* 25.3, pp. 289–305. ISSN: 0167-7764, 1573-0662. DOI: 10/bcjjc3. URL: https://link.springer.com/article/10.1007/BF00053797 (visited on 03/20/2018) (cit. on p. 63).

VANDAELE, A. C., P. C. SIMON, J. M. GUILMOT, M. CARLEER, and R. COLIN (**1994**). "SO_2 Absorption Cross Section Measurement in the UV Using a Fourier Transform Spectrometer". In: *Journal of Geophysical Research: Atmospheres* 99.D12, pp. 25599–25605. ISSN: 2156-2202. DOI: 10.1029/94JD02187. URL: http:

//onlinelibrary.wiley.com/doi/10.1029/94JD02187/abstract (visited on 03/01/2018) (cit. on p. 24).

VEEFKIND, J. P., I. ABEN, K. MCMULLAN, H. FÖRSTER, J. de VRIES, G. OTTER, J. CLAAS, H. J. ESKES, J. F. de HAAN, Q. KLEIPOOL, M. van WEELE, O. HASEKAMP, R. HOOGEVEEN, J. LANDGRAF, R. SNEL, P. TOL, P. INGMANN, R. VOORS, B. KRUIZINGA, R. VINK, H. VISSER, and P. F. LEVELT (**2012**). "TROPOMI on the ESA Sentinel-5 Precursor: A GMES Mission for Global Observations of the Atmospheric Composition for Climate, Air Quality and Ozone Layer Applications". In: *Remote Sensing of Environment*. The Sentinel Missions - New Opportunities for Science 120, pp. 70–83. ISSN: 0034-4257. DOI: 10/f3xw4m. URL: http://www.sciencedirect.com/science/article/pii/S0034425712000661 (visited on 10/25/2018) (cit. on p. 2).

VLEMMIX, T., F. HENDRICK, G. PINARDI, I. DE SMEDT, C. FAYT, C. HERMANS, A. PITERS, P. WANG, P. LEVELT, and M. VAN ROOZENDAEL (**2015**). "MAX-DOAS Observations of Aerosols, Formaldehyde and Nitrogen Dioxide in the Beijing Area: Comparison of Two Profile Retrieval Approaches". In: *Atmospheric Measurement Techniques* 8.2, pp. 941–963. ISSN: 1867-8548. DOI: 10.5194/amt-8-941-2015. URL: https://www.atmos-meas-tech.net/8/941/2015/ (visited on 09/13/2017) (cit. on pp. 3, 106, 116, 202, 206).

VLEMMIX, T., A. J. M. PITERS, A. J. C. BERKHOUT, L. F. L. GAST, P. WANG, and P. F. LEVELT (**2011**). "Ability of the MAX-DOAS Method to Derive Profile Information for NO_2: Can the Boundary Layer and Free Troposphere Be Separated?" In: *Atmospheric Measurement Techniques* 4.12, pp. 2659–2684. ISSN: 1867-8548. DOI: 10.5194/amt-4-2659-2011. URL: https://www.atmos-meas-tech.net/4/2659/2011/ (visited on 09/13/2017) (cit. on p. 106).

VOLKAMER, Rainer, Peter SPIETZ, John BURROWS, and Ulrich PLATT (**2005**). "High-Resolution Absorption Cross-Section of Glyoxal in the UV–Vis and IR Spectral Ranges". In: *Journal of Photochemistry and Photobiology A: Chemistry* 172.1, pp. 35–46. ISSN: 1010-6030. DOI: 10/cwgfv5. URL: http://www.sciencedirect.com/science/article/pii/S1010603004005143 (visited on 06/05/2018) (cit. on p. 275).

WAGNER, T., A. APITULEY, S. BEIRLE, S. DÖRNER, U. FRIESS, J. REMMERS, and R. SHAIGANFAR (**2014**). "Cloud Detection and Classification Based on MAX-DOAS Observations". In: *Atmospheric Measurement Techniques* 7.5, pp. 1289–1320. ISSN: 1867-8548. DOI: 10.5194/amt-7-1289-2014. URL: http://www.atmos-meas-tech.net/7/1289/2014/ (visited on 09/12/2017) (cit. on pp. 67, 68).

WAGNER, T., S. BEIRLE, T. BRAUERS, T. DEUTSCHMANN, U. FRIESS, C. HAK, J. D. HALLA, K. P. HEUE, W. JUNKERMANN, X. LI, U. PLATT, and I. PUNDT-GRUBER (**2011**). "Inversion of Tropospheric Profiles of Aerosol Extinction and HCHO and NO_2 Mixing Ratios from MAX-DOAS Observations in Milano during the Summer of 2003 and Comparison with Independent Data Sets". In: *Atmospheric Measurement Techniques* 4.12, pp. 2685–2715. ISSN: 1867-8548. DOI: 10.5194/amt-4-2685-2011. URL: https://www.atmos-meas-tech.net/4/2685/2011/ (visited on 09/13/2017) (cit. on pp. 3, 106).

WAGNER, T., J. P. BURROWS, T. DEUTSCHMANN, B. DIX, C. von FRIEDEBURG, U. FRIESS, F. HENDRICK, K.-P. HEUE, H. IRIE, H. IWABUCHI, Y. KANAYA, J. KELLER, C. A. MCLINDEN, H. OETJEN, E. PALAZZI, A. PETRITOLI, U. PLATT, O. POSTYLYAKOV, J. PUKITE, A. RICHTER, M. van ROOZENDAEL, A. ROZANOV, V. ROZANOV, R. SINREICH, S. SANGHAVI, and F. WITTROCK (**2007**). "Comparison of Box-Air-Mass-Factors and Radiances for Multiple-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) Geometries Calculated from Different UV/Visible Radiative Transfer Models". In: *Atmospheric Chemistry and Physics* 7.7, pp. 1809–1833. ISSN: 1680-7324. DOI: 10/cdf2hm. URL: https://www.atmos-chem-phys.net/7/1809/2007/ (visited on 04/24/2018) (cit. on p. 40).

WAGNER, T., B. DIX, C. v. FRIEDEBURG, U. FRIESS, S. SANGHAVI, R. SINREICH, and U. PLATT (**2004**). "MAX-DOAS O_4 Measurements: A New Technique to Derive Information on Atmospheric Aerosols-Principles and Information Content: MAX-DOAS O $_4$ MEASUREMENTS OF AEROSOLS". In: *Journal of Geophysical Research: Atmospheres* 109.D22, n/a–n/a. ISSN: 01480227. DOI: 10.1029/2004JD004904. URL: http://doi.wiley.com/10.1029/2004JD004904 (visited on 09/12/2017) (cit. on pp. 2, 49, 105, 106, 110).

WAGNER, T., F. ERLE, L. C. MARQUARD, C. OTTEN, K. PFEILSTICKER, and T. SENNE (**1998**). "Cloudy Sky Optical Paths as Derived from Differential Optical Absorption Spectroscopy Observations". In: *Journal of*

Geophysical Research: Atmospheres 103.D19, pp. 25307–25321. ISSN: 2156-2202. DOI: 10/cdgjw3. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JD01021 (visited on 09/14/2018) (cit. on p. 2).

WAGNER, Thomas, Steffen BEIRLE, Nuria BENAVENT, Tim BÖSCH, Kai Lok CHAN, Sebastian DONNER, Steffen DÖRNER, Caroline FAYT, Udo FRIESS, David GARCÍA-NIETO, Clio GIELEN, David GONZÁLEZ-BARTOLOME, Laura GOMEZ, François HENDRICK, Bas HENZING, Jun Li JIN, Johannes LAMPEL, Jianzhong MA, Kornelia MIES, Mónica NAVARRO, Enno PETERS, Gaia PINARDI, Olga PUENTEDURA, Janis PUĶĪTE, Julia REMMERS, Andreas RICHTER, Alfonso SAIZ-LOPEZ, Reza SHAIGANFAR, Holger SIHLER, Michel Van ROOZENDAEL, Yang WANG, and Margarita YELA (**2018**). "Is a Scaling Factor Required to Obtain Closure between Measured and Modelled Atmospheric O_4 Absorptions? – A Case Study for Two Days during the MADCAT Campaign". In: *Atmospheric Measurement Techniques Discussions*, pp. 1–85. ISSN: 1867-1381. DOI: 10/gfb74z. URL: https://www.atmos-meas-tech-discuss.net/amt-2018-238/ (visited on 10/11/2018) (cit. on pp. 184, 188, 213).

WANG, Y., J. LAMPEL, P. XIE, S. BEIRLE, A. LI, D. WU, and T. WAGNER (**2017**). "Ground-Based MAX-DOAS Observations of Tropospheric Aerosols, NO_2, SO_2 and HCHO in Wuxi, China, from 2011 to 2014". In: *Atmospheric Measurement Techniques* 17.3, pp. 2189–2215. ISSN: 1680-7324. DOI: 10.5194/acp-17-2189-2017. URL: https://www.atmos-chem-phys.net/17/2189/2017/ (visited on 09/13/2017) (cit. on pp. 2, 82, 83, 106, 116, 183, 198, 205, 206, 208).

WANG, Y., A. LI, P. H. XIE, T. WAGNER, H. CHEN, W. Q. LIU, and J. G. LIU (**2014**). "A Rapid Method to Derive Horizontal Distributions of Trace Gases and Aerosols near the Surface Using Multi-Axis Differential Optical Absorption Spectroscopy". In: *Atmospheric Measurement Techniques* 7.6, pp. 1663–1680. ISSN: 1867-8548. DOI: 10.5194/amt-7-1663-2014. URL: http://www.atmos-meas-tech.net/7/1663/2014/ (visited on 09/12/2017) (cit. on pp. 87, 88, 90).

WITTROCK, F., H. OETJEN, A. RICHTER, S. FIETKAU, T. MEDEKE, A. ROZANOV, and J. P. BURROWS (**2004**). "MAX-DOAS Measurements of Atmospheric Trace Gases in Ny-Ålesund - Radiative Transfer Studies and Their Application". In: *Atmospheric Chemistry and Physics* 4.4, pp. 955–966. ISSN: 1680-7324. DOI: 10.5194/acp-4-955-2004. URL: https://www.atmos-chem-phys.net/4/955/2004/ (visited on 09/12/2017) (cit. on pp. 2, 105, 106).

WITTROCK, Folkard (**2006**). "The Retrieval of Oxygenated Volatile Organic Compounds by Remote Sensing Techniques". Phd thesis. Bremen: University of Bremen. URL: http://nbn-resolving.de/urn:nbn:de:gbv:46-diss000104818 (cit. on pp. 2, 106, 107, 211).

ZIEGER, P., E. WEINGARTNER, J. HENZING, M. MOERMAN, G. de LEEUW, J. MIKKILÄ, M. EHN, T. PETÄJÄ, K. CLÉMER, M. van ROOZENDAEL, S. YILMAZ, U. FRIESS, H. IRIE, T. WAGNER, R. SHAIGANFAR, S. BEIRLE, A. APITULEY, K. WILSON, and U. BALTENSPERGER (**2011**). "Comparison of Ambient Aerosol Extinction Coefficients Obtained from In-Situ, MAX-DOAS and LIDAR Measurements at Cabauw". In: *Atmospheric Chemistry and Physics* 11.6, pp. 2603–2624. ISSN: 1680-7324. DOI: 10.5194/acp-11-2603-2011. URL: https://www.atmos-chem-phys.net/11/2603/2011/ (visited on 09/13/2017) (cit. on pp. 106, 167).