Investigation of metal atom and ion number densities in the mesosphere and lower thermosphere by satellite remote sensing with SCIAMACHY/Envisat

Doktorarbeit

Universität Bremen
Fachbereich I
Institut für Physik
Institut für Umweltphysik

eingereicht von
Martin Langowski

1. Gutachter:
John P. Burrows

2. Gutachter:
Arthur C. Aikin

Datum des Kolloquiums
20.01.2016

Abstract

Extraterrestrial meteoroids continuously enter the Earth's atmosphere. The bulk of the meteoric mass comes from about 220 µm sized particles coming mostly from asteroids in the asteroid belt between Mars and Jupiter. The particle size ranges from atomic scale particles to macroscopic bodies with a mass of several tons. The input rate is very uncertain, ranging from 2 to 300 tons per day. The meteoroids collide with the Earth's atmosphere at very high velocities (on the order of 10 to several 10 km per second), which leads to frictional heating of the meteoroids. This leads to full or partial ablation of the meteoroids in the upper atmosphere. Since the meteoroids consist partly of metals, this ablation leads to the formation of metal atom layers in the mesosphere and lower thermosphere (MLT). The metals are partly ionized in the MLT and metal ion layers are also formed.

The MLT cannot easily be probed by in situ measurements. The pressure is too low for balloons to fly in this region, and the drag on satellites is too high there for stable satellite orbits. Therefore, only rocket-borne in situ measurements can be performed in the MLT. Because of this, remote sensing techniques are preferred to investigate the MLT. The atmospheric metal atoms and ions can be observed by remote sensing, because they are strong emitters of resonance fluorescence. Therefore, they are valuable tracer species for wave propagation and other transport processes in the MLT. In addition to their role as tracers, meteoric metals are also important for the chemistry in the middle atmosphere. Information on their absolute number densities is therefore of interest. The metals are involved in the formation of cluster ions and meteoric smoke particles (MSP), which are important for the heterogeneous condensation and nucleation of aerosols and clouds in the middle atmosphere. To estimate the impact of meteoric metals on the chemistry of the middle atmosphere, global observations are necessary. These need to be combined with models, which also include the yet undetectable particle species like, e.g., the metal molecules.

Only in the last decades have global long-term satellite observations of the Earth's atmosphere with good local, temporal, and vertical resolution and coverage been available. One of these data sets comes from the observations of SCIAMACHY/Envisat, which provides daily limb and nadir observations from 2002 to 2012. Beginning in mid-2008, a special limb observation mode for the investigation of the MLT was performed. This data set was used to retrieve the number densities of Mg, Mg^+, and Na from 2008-2012. Therefore, an existing retrieval algorithm for Mg and Mg^+ was significantly improved and also extended for application in the visible spectral region for Na. The improved retrieval algorithm and the results for Mg, Mg^+ and Na are presented and discussed in this thesis, and the metal's interaction with polar mesospheric clouds is investigated. Furthermore, a survey of spectral signatures of other target species for the retrieval algorithm has been carried out. The measurements have been compared with model results, which include the total daily meteoric mass influx as well as other measurement results. The seasonal, altitudinal and latitudinal variation of metal species is in good agreement with the most recent model results.

Contents

Part I
Introduction and background information

1 Outline

The metal atom and ion layers in the mesosphere and lower thermosphere (MLT) are investigated in this thesis. The number densities of different metal atom and ion species are retrieved from the SCIAMACHY/Envisat limb observations and the results are analysed. The unique SCIAMACHY limb MLT data set, providing resonance fluorescence radiation, is used for the first time to retrieve number densities of Mg, Mg^+ and Na. The data set has been corrected for several previously unaccounted sources of error making it the most accurate global measurement of Mg and Mg^+ distributions. In addition, for the first time the measured global Mg, Mg^+ densities are compared to a two-dimensional meteorological model, which includes metal, ion, and neutral chemistry. The retrieved Na data set is beside the data sets from OSIRIS/Odin and GOMOS/Envisat one of only few existing global data sets, which differ in local time and/or used measurement geometry.

The Thesis is structured as follows: as an introduction, Sect. 2 discusses the general vertical structure of the Earth's atmosphere with a focus on the upper atmosphere where the metal layers reside. The metal layers are introduced in Sect. 3. A short overview on SCIAMACHY is given in Sect. 4, and in Sect. 5 a survey of suitable metal emission lines, as well as other emission features in the SCIAMACHY limb MLT observations is presented. The focus of this thesis is the improvement of the existing number density retrieval algorithm for magnesium atoms (Mg) and ions (Mg^+). Several new features have been implemented and the algorithm is extended to sodium atoms (Na). A detailed description of the number density retrieval algorithm is provided in Sect. 6. The results for the retrieved metal number densities are presented and discussed in Sect. 7 (Mg/Mg^+) and Sect. 8 (Na). Noctilucent clouds (NLCs) can be found in the same altitude region as the mesospheric metal layers, and in Sect. 9 the mesospheric metal data set is investigated in terms of signs of interactions with the NLCs. In Sect. 10 the global mass input of meteoric material into the Earth's atmosphere is estimated. Finally, Sect. 11 provides a summary of this PhD thesis and an outlook for further investigations.

2 The Earth's atmosphere

This chapter gives a short introduction to the vertical structure of the atmosphere. Most of this material is well known and can be found in various textbooks (see, e.g., Rödel and Wagner, 2011; Brasseur and Solomon, 2005, and others). The focus of this overview will be on the mesosphere and lower thermosphere, since the metal atom and ion layers investigated in this thesis are located in this region.

2.1 General vertical structure

Using maxima and minima in the vertical profile of atmospheric temperature T leads to a convenient subdivision of the atmosphere into regions with similar properties and driving physical effects. This kind of subdivision is a good choice, because the pressure p and the density n in the ideal gas law, $p = nk_BT$ (with the Boltzmann-constant k_B), are nearly exponentially decreasing with altitude following the barometric formula $p = p_0 e^{-\frac{Mg}{RT}(h-h_0)}$, with p_0 being the pressure at altitude $h = h_0$, M being the molar mass of air, g the local gravitational acceleration and R the universal gas constant. Therefore, only the temperature shows well pronounced minima and maxima. If there were only adiabatic vertical transport processes in the atmosphere, the temperature would slowly decrease with altitude, because thermal energy of an air parcel, which moves upwards, is converted into potential energy. Because p and T are coupled by $T^\gamma p^{1-\gamma} = const$ for adiabatic processes, a quantity for T at $p = p_0 = p(h_0)$ can be defined that is constant for adiabatic processes, which is called the potential temperature Θ, $\Theta = T\left(\frac{p_0}{p}\right)^{\frac{\gamma-1}{\gamma}}$. The adiabatic index γ is the ratio of the heat capacity at constant pressure C_p and heat capacity at constant volume C_v, $\gamma = \frac{C_p}{C_v}$, which is $\frac{f+2}{f}$ for ideal gases, with f being the degree of freedom. For typical atmospheric conditions only translational and rotational degrees are considered, while vibrational degrees are only activated at high energies, which leads to $f = 3$ for mono-atomic gases, e.g., Ar, $f = 5$ for linearly formed molecules, e.g., N_2 and O_2 and $f = 6$ for most other molecules, e.g., H_2O. It should be noted that $H = \frac{RT}{Mg}$ is the scale height, the altitude range at which the pressure falls by a factor e, which is around 7.8 km close to Earth. Furthermore, the barometric pressure law is actually a bit more complicated, due to the fact that the temperature T and local gravitational acceleration g are not constant.

The potential temperature is convenient to visualize how different the vertical temperature profile is compared to the adiabatic temperature profile, as humans are usually better trained in seeing deviations of straight lines then deviations of exponential functions. The real temperature does not follow an adiabatic decompression with altitude, which has significant influence on the vertical transport within the different atmospheric layers and, therefore, explains why it is so convenient to separate the atmosphere by temperature extrema. The reason for the non-adiabatic behavior mainly comes from the input of solar energy into different regions of the Earth's atmosphere through absorption of solar radiation by certain absorbers at certain wavelength regions, or loss/gain of energy through emission of the different gases forming the atmosphere (radiative cooling/heating). Furthermore, there are

other energy sources and sinks, e.g., phase transitions of water and atmospheric wave creation and dissipation processes.

The different sources and sinks of energy lead to either steeper or flatter vertical temperature gradients than the adiabatic case. The adiabatic temperature profile is called the dry adiabatic lapse rate, if no phase transitions of water are considered, and the saturated adiabatic lapse rate, if phase transitions of water are considered. If the true vertical temperature gradient is steeper than the saturated adiabatic lapse rate, the atmosphere is layered instable. Instable means that if an air parcel rises upward to a higher level, it is forced to rise even further upward to a higher level. This instability is contradictory to an equilibrium condition in the vertical structure and, therefore, must lead to a countering transport mechanism, which preserves the equilibrium. These mostly turbulent transport mechanisms are much faster than, e.g., the slower transport by molecular diffusion. For the case that the true lapse rate is less steep than the saturated adiabatic lapse rate, the atmosphere is layered stable, and if an air parcel moves upwards, it is forced to move down again and vice versa. As instabilities lead to much faster vertical transport than stable conditions, stable regions in the atmosphere function as transport barriers and, therefore, the physical mechanisms in each of the layers can be treated in a closed system, to a certain degree. The vertical structure of the atmosphere therefore is strongly depending on the difference between the true vertical gradient of the temperature to the adiabatic vertical gradient. However, using the maxima and minima of the vertical temperature profile gives a good approximation of these different regimes. Regions with increasing temperature with altitudes are stable, while regions with temperatures that decrease with altitude can still be stable, but especially in the troposphere can also be instable.
Figure 2.1 shows a typical vertical temperature profile of the atmosphere, and Fig. 2.2 shows the vertical and latitudinal temperature distribution in July.

2.2 The troposphere

The temperature decreases with altitude from the Earth's surface to a first minimum. This minimum is called the tropopause and is the boundary region between the troposphere and the stratosphere. Interactions with the surface are non negligible in the vicinity of the surface. This region, which can have a vertical extension of up to 2 km, is called the Planetary Boundary Layer and is part of the troposphere. The troposphere, derived from the Greek word tropos for "mixing", contains around 80% of the mass of the atmosphere and around 99% of the water vapor and aerosols. The tropopause altitude varies with time and latitude. This variation is caused by different solar influence and by the local terrain. The tropopause altitude varies between 8-15 km. It is sufficient to only consider this lowest layer, which is well separated from higher layers, for most weather and climate phenomena. However, the chemistry and physics of the upper layers, which may only have a small influence on weather and climate, are important too, as, e.g., the upper atmosphere protects the living species on the surface from several extraterrestrial threats. The most famous example for this is located in the next higher layer, the stratosphere.

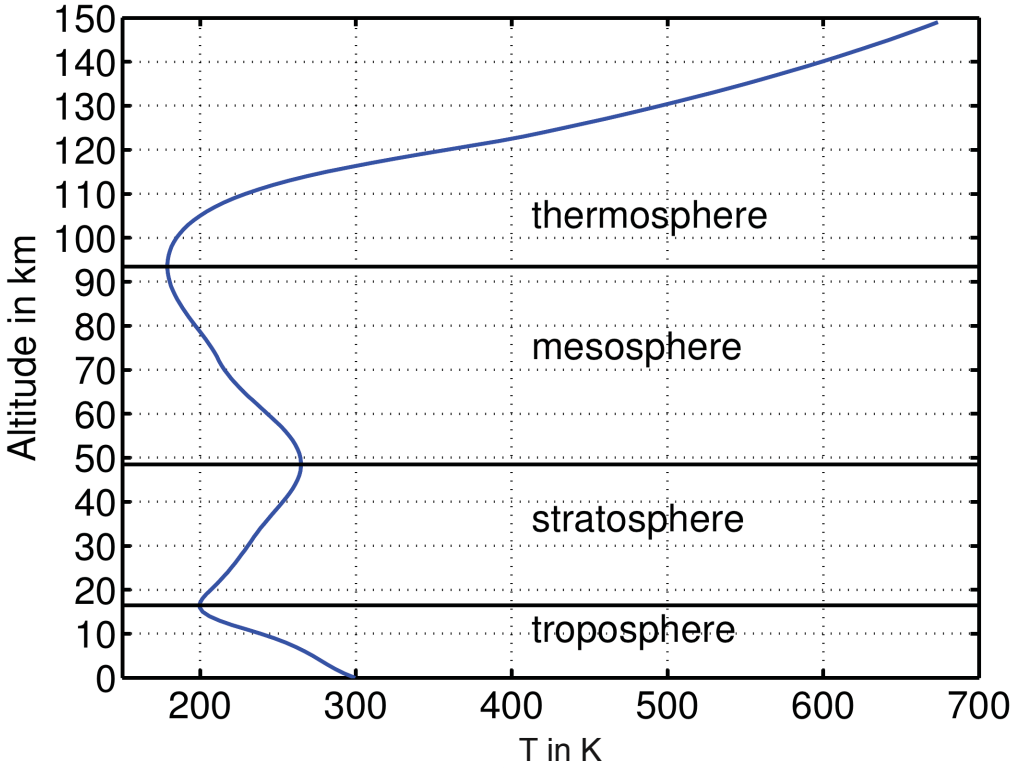

Fig. 2.1: Vertical temperature profile of the atmosphere from $0 - 150$ km altitude at equatorial latitudes. The data used is from the MSISE-00-model (Picone et al., 2002).

2.3 The stratosphere and stratospheric ozone

The stratosphere ranges from the tropopause to the stratopause at around 50 km altitude. The name of the stratosphere comes from the stably stratified layering of the atmosphere there with a temperature profile with warmer layers being always above colder layers. This temperature increase with altitude comes from the absorption of ultraviolet (UV) radiation by ozone in the stratosphere. This absorption of UV radiation occurs mostly in the O_3 Hartley-Bands from around 200 to 300 nm wavelength, see Fig. 6.1.

Figure 2.3 shows a vertical profile of ozone density, as well as a vertical profile of the volume mixing ratio (vmr), which is the ratio of ozone number density to the total atmospheric number density at each altitude. The density of ozone has its maximum at around 20 km, the vmr of ozone peaks at higher altitudes (30-35 km), and the temperature maximum, the stratopause, can be typically found much higher at around 50 km. The four main reactions of atmospheric oxygen photo-chemistry were proposed by Chapman (1930):

$$O_2 + h\nu \rightarrow 2O \tag{2.1}$$

$$O + O_2 + M \rightarrow O_3 + M \tag{2.2}$$

$$O_3 + h\nu \rightarrow O_2 + O \tag{2.3}$$

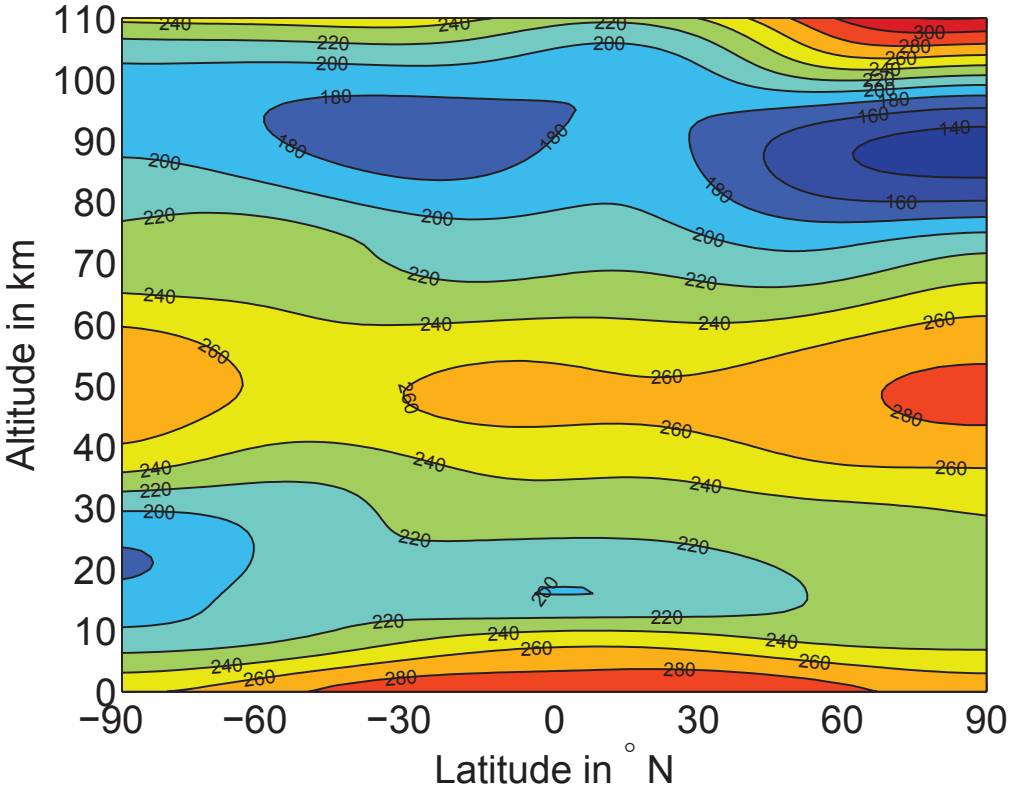

Fig. 2.2: Temperature in K at different latitudes and altitudes in July. The data used is from the MSISE-00-model for July 20, 2013 (Picone et al., 2002).

$$O_3 + O \rightarrow 2O_2. \tag{2.4}$$

Reactions (2.2) and (2.3) are much faster than reaction (2.1) and (2.4). O_2 is photo-dissociated at wavelengths below 240 nm and ozone mainly is photo-dissociated in the Hartley-band at wavelengths below 320 nm. The photo-dissociation of O_2 mainly occurs in the thermosphere and there is more O available at higher altitudes. On the other hand side the density of O_2 and M for reaction (2.2) are higher at lower altitudes. These two facts explain, why the maximum of ozone vmr is higher than the density maximum. The absorption of photons by ozone (reaction (2.3)) is so strong, that it significantly reduces the incoming radiation, which can be absorbed by ozone, at the upper edge of the ozone profile, and thus the temperature maximum is even higher than the ozone vmr maximum. The order of these 3 characteristic maxima can be well explained by Chapman's reactions. However, these reactions alone lead to a stratospheric profile with about a factor 2 too high concentrations than observed. Furthermore, the density's peak altitude is too high.

Additional catalytic cycles of HO_x(OH and HO_2) (see, e.g., Bates and Nicolet, 1950) and NO_x (NO and NO_2) (see, e.g., Crutzen, 1970), that effectively convert ozone into molecular oxygen, lead to very exact models of the stratospheric ozone. The discovery of the NO_x cycle led to the proposal that in addition to natural sources exhaust gases of airplanes in the stratosphere could add to a depletion of the ozone layer, which was very important in that time, as the technology for mass production

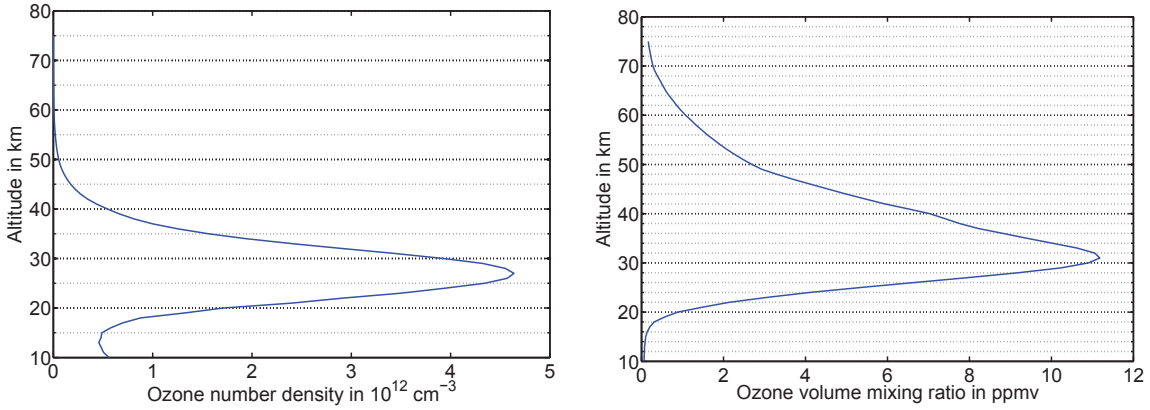

Fig. 2.3: Equatorial ozone number density and volume mixing ratio profiles from SCIAMACHY.

of airplanes flying in the stratosphere had been made accessible. In addition to this reactions, it is known that halogen molecules like Cl_2 or Br_2 are photo-dissociated in the stratosphere to its atomic radicals which are catalysts of the reaction of O_3 to O_2.

Due to the work by Molina and S. (1974) the ozone depleting role of chlorofluorocarbons (CFCs), which have been widely used as refrigerants, propellants and solvents, was found and a slow but steady decrease of the ozone concentration in the stratospheric ozone layer was expected and investigated. However, in the 1980s unexpectedly strong depletions in the ozone concentration at the end of the Antarctic winter were observed. The reason for this additional ozone loss was quickly found in the heterogenic chemistry at the surface of polar stratospheric clouds (PSCs). PSCs are formed in polar regions in the winter. The chemistry on the surface of these clouds leads to very effective reactions of chemical sinks for halogens, which are actually inert, into halogens in their molecular form. At the end of the polar night UV radiation splits these halogen molecules into radicals, which convert ozone to molecular oxygen in catalytic cycles. As a consequence of these rapid ozone losses, the industrial use of CFCs, was finally limited in the Montreal Protocol 1987.

The formation of stratospheric clouds is more pronounced in the southern hemisphere, because the polar vortex, a persistent, large thermal wind, circling the geographical poles, which strengthens in the winter due to higher temperature differences between the pole and lower latitudes, is stronger in the Southern Hemisphere. The Southern Hemisphere is less perturbed by atmospheric wave activity than the Northern Hemisphere, where there is more land mass and therefore orographic obstacles that causes wave activity. Dissipation of waves in the middle atmosphere can lead to so-called sudden stratospheric warmings, which dramatically change the temperature by several tens of degrees, as well as the wind direction and the momentum flux (see, e.g., de Wit et al., 2014), and the stratopause is lifted tens of kilometers upwards into the mesosphere. These sudden stratospheric warmings, which occur nearly every two years at the end of the Arctic winter, weaken the polar vortex. A more detailed overview of the stratospheric ozone chemistry can be found, e.g., in

Solomon (1999). The influence of the stratospheric ozone layer is very important for the radiative transfer model used for retrieving metal densities in the mesosphere and lower thermosphere. In Sect. 6, it is described, that the absorption of UV radiation by ozone leads to important simplifications of the radiative transfer model used for this work.

2.4 The mesosphere

The mesosphere reaches from the stratopause at around 50 km to the mesopause at around 90-100 km. "Meso" is the Greek word for middle, because this sphere is the middle layer of the five layers defined by temperature extrema. The mesosphere is not as strongly diabatic heated by absorption of solar radiation as the thermosphere, because solar radiation, which is absorbed in the above laying thermosphere by O_2, cannot reach the mesosphere. Nevertheless, diabatic solar heating is one of the most important heating mechanisms in the mesosphere due to absorption of UV radiation by O_3. In addition, ozone plays an important role as one of the most chemically reactive species in this region, and there is a second maximum in ozone vmr in the mesopause region (see, e.g., Smith et al., 2013, Fig. 2). Molecules, mainly CO_2 and to a smaller extent O_3 and H_2O, cool the mesosphere by radiative cooling in the IR-wavelength region. This occurs because the thermal energy, which is partly deposited in the rotation of the molecules, is lost by radiation through transitions between different vibrational-rotational levels. With decreasing density the mean free path for collisions and electron ion recombination gets longer, which increases the lifetime and therefore the density of unstable/reactive atoms, molecules and ions like O, H, O^+, NO^+, O_2^+, cluster molecules and free electrons (see, e.g., Plane, 2003, Fig. 3 and 4). Beside radiative processes, there is also heating from exothermic chemical reactions, breaking gravity waves in the thinning atmosphere, and frictional heating of incoming meteoroids. Although the temperature T decreases with altitude in the mesosphere, the potential temperature Θ rises with altitude and the mesosphere is therefore stably layered.

The mesopause region is the coldest region in the Earth's atmosphere with typical temperatures of 140-200 K. Amazingly, the summer polar mesopause is even colder than the winter mesopause. The reason for this is associated by upwelling – and adiabatic cooling – above the summer pole, and downwelling – and adiabatic heating – above the winter pole, which is driven by different gravity wave breaking in both hemispheres. As the real temperature gradient is less steep than the adiabatic one, the upwelling brings air particles with lower potential temperature Θ to higher altitudes. This adiabatic cooling cools the summer mesopause, while there is adiabatic heating at the winter pole, due to the downward transport. The mesopause also shows a seasonal variability in altitude and is around 10 km lower at the summer pole.

The mesosphere and lower thermosphere region from 50 to 150 km altitude is currently one of the most expensive accessible region in the Earth's atmosphere for in-situ measurements. The air is too thin for aircraft to fly in this region, but too dense for long-term satellite operation. In-situ measurements, therefore, are only available from rocket measurements, which are expensive and can only be performed

at limited locations around the world. As in-situ measurements are difficult and expensive, remote sensing methods are important in this region. Meteoric metals show significant emission features in this region, and since this thesis is focused on the investigation of meteoric metal layers, these will be discussed in more detail in Sect. 3.

2.5 The thermosphere and the exosphere

The thermosphere (thermos meaning heat), reaching from the mesopause to $\approx$ 600 km altitude, is the hottest part of the Earth's atmosphere, with temperatures of up to over 1000 K. The reason for this is the absorption of UV radiation (in the Herzberg continuum between 200 and 240 nm as well in the Schumann-Runge continuum between 100 and 176 nm) by O_2 and the photo-dissociation and photo-ionisation of O_2 for these and shorter wavelength regions. The photo-dissociation of O_2 leads to a higher density of O than the one of O_2 above 110 km altitude. As the air density is low, the mean free path of collisions becomes long enough, that the conditions for a local thermodynamic equilibrium are less and less present for rising altitudes. Therefore, the temperature is not well defined in this altitude.

One definition of the border between Earth and space is an altitude of 100 km, and this altitude is in the lower thermosphere. Around 100 km is also the altitude of the turbopause, where turbulent mixing and molecular diffusion are equally strong. The atmosphere below the turbopause is also called the homosphere, because through the effective turbulent mixing the vmr of the main constituents are nearly constant at each altitude. Therefore, changes in the vmr indicate production and destruction reactions of the constituents at certain altitudes, which is, e.g., the case for the significant O_2 loss and O gain due to photo-dissociation around the turbopause. In the atmosphere above the turbopause, the constituents are more and more separated by their masses with lighter species having higher vmr at higher altitudes. Figure 2.4 shows the vmr profiles for the main atmospheric constituents. The turbopause is most easily identified for the inert noble gas species, of which argon shows a significant decrease in vmr above 100 km, while helium, which is lighter than argon, shows an increase of its vmr with altitude above 100 km. Because of the photo-dissociation of O_2 and the lower weight of the O atom compared to the N_2 and O_2 molecules, atomic oxygen O becomes the dominant species above 200 km. Atomic oxygen is succeeded by the lighter species helium He and hydrogen H as the dominant species of the thermosphere for higher altitudes.

The thermosphere ends at the thermopause at around 600-800 km altitude, which is the region where the atmosphere is thin enough and the mean free paths are long enough that the particles nearly do not collide at all and are in Keplerian orbits around the Earth. A temperature can still be defined through the mean kinetic energy of all particles, which remains constant for all altitudes above the thermopause. However, because there are few collisions, the velocity distribution of the constituents is not Maxwellian. The exosphere (exo is Greek for external) includes all particles that are still gravitationally bound to Earth. As there are almost no collisions at all at this altitude, the particles there can be treated as single bodies with elliptic orbits around the Earth.

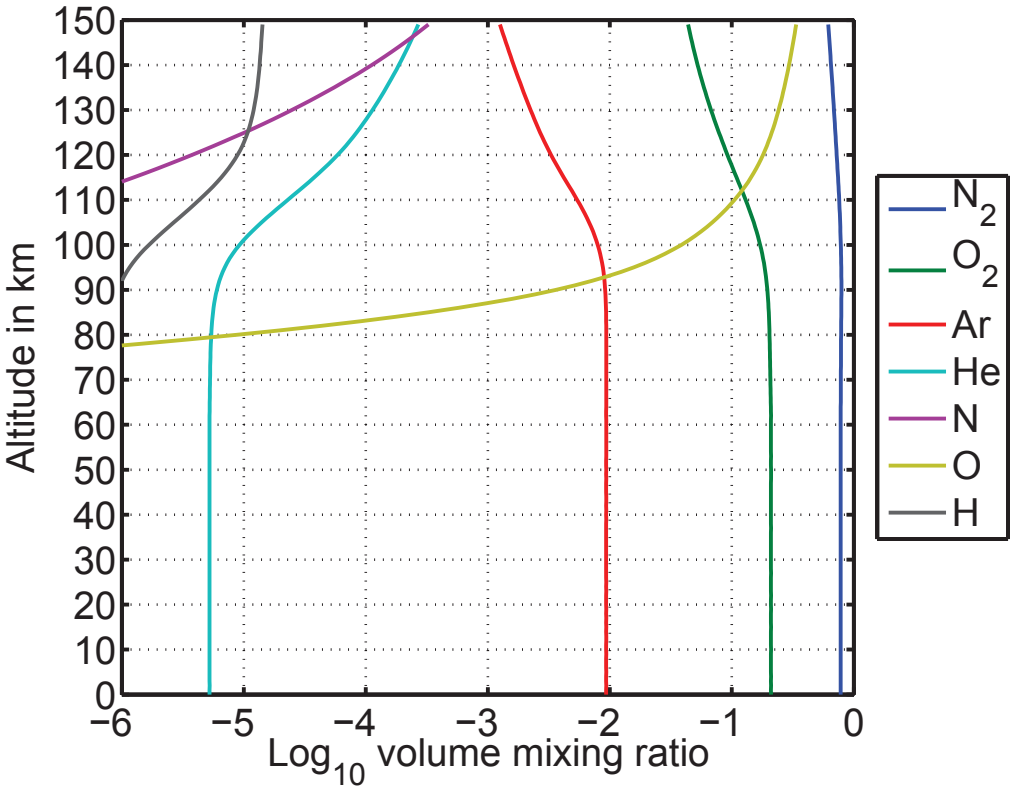

Fig. 2.4: Vertical volume mixing ratio profiles of different atmospheric constituents at the equator. The used data is from the MSISE-00-model for July 20, 2013 (Picone et al., 2002). Molecular diffusion is more important than other mixing processes in most of the thermosphere, so that constituents are distributed by mass.

2.6 The ionosphere

A part of the Atoms and molecules above 70 km are ionised by absorption of highly energetic UV and X-ray radiation. Furthermore, the ion electron recombination rate decreases with altitude, which increases the lifetime of the formed ions. Thus, the upper atmosphere is partly ionized and forms a very thin plasma of ions and electrons.

The far field forces of charged particles are much stronger than the weak interaction between the neutral species, which leads to strong collective effects and electric fields, currents, and interaction with the terrestrial magnetic field, that influence particle transport mechanisms in the upper atmosphere, and also imply radiative transfer barriers, which are especially important for radio communication. The first observations of the ionosphere date back to the development of radio communication at the beginning of the 20th century. It was found, that with a sufficient powerful emitter radio signals could be sent around the whole Earth. However, this was only true for long wavelength radiation. This is, because electromagnetic waves can only penetrate plasmas if their frequency is higher than the plasma frequency of the plasma $\omega_p = \sqrt{\frac{n_e e^2}{\epsilon m_e}}$ (with electron density n_e, electron charge e, permittiv-

ity ϵ and electron mass m_e; a very good introduction to the most basic types of waves in plasmas can be found, e.g., in Chen (1984)). Because of this, local radio stations typically use wavelengths between 1 m and 10 m (VHF), while wavelengths from 100 m to 10000 m (MF and LF) are used for global broadcasting. With the beginning of the space age, this aspect became even more important for satellite communication and satellite control.

For the classification of the different ionospheric layers the extrema of (free) electron density n_e are used, as this quantity nearly directly defines the plasma frequency. As the ionization is strongly dependent on the solar irradiation, the layers are very different for day and night conditions and vary with the solar cycle.

The first layer that was found was called the E(lectric)-layer. The layer below, which was later found, is called the D-layer and the layer above the F-layer. The D-layer between 60 km and 90 km is the lowermost layer and is formed by ionization of NO through Lyman-α radiation at 121.6 nm. In addition, during times of high solar activity, hard X-rays ($\lambda < 1$ nm) can reach this altitude region and directly ionize N_2 and O_2. The D-layer is nearly destroyed several minutes after sunset, because recombination is very fast at this altitude. In the night, negative ions are formed in the D-Region.

The E-layer between 90 km and 150 km is formed by ionization of O, N_2, and O_2 by soft X-ray-radiation (1-10 nm) and Lyman-(β-∞) radiation between 91-102 nm. It has roughly a free electron concentration of $10^5\,\mathrm{cm}^{-3}$ and only 0.1% of the atoms are ionized. After sunset, the E-layer exists longer than the D-layer, but it is also depleted after an hour. Sporadic E_s-layers can form at the same altitude as the E-layers. These layers only occur for several hours and have a highly increased electron density. The formation of the E_s-layers is due in part to the presence of long lived metallic ions focused by wind shears. Frequencies in the VHF regions are also reflected at the E_s-layer and MF frequency signals are more strongly absorbed in the E_s-layer.

The ions in the F-layer, which extends from 150-500 km, are formed by extreme UV (EUV) radiation ($\lambda = 14 - 18$ nm), which ionizes atomic oxygen O and atomic nitrogen N. The F-layer can be separated into the F_1- and the F_2-layer. The highest ion production occurs in the F_1 layer, while the highest ion concentration can be found in the higher F_2 layer, since the recombination is weaker in this layer. The F_1 disappears during the night because of recombination, while the F_2-layer also exists at night.

2.7 The magnetosphere

The Earth has a magnetic field which is formed in the Earth's core. The region in space where charged particles are controlled by this field is called the magnetosphere. The magnetic field protects the Earth from fast and highly energetic particles, which originate mainly from the solar wind, which is a plasma stream originating at the sun. The strength of the solar wind is a function of solar activity. The influence of the solar wind is so strong, that it strongly distorts the magnetic field. At solar maximum, the Sun produces roughly three coronal mass ejections (CMEs) every day, while at solar minimum, there is only one CME every five days. When a CME

points towards Earth, it produces a shockwave that causes geomagnetic storms in the magnetosphere and heavily distorts the magnetic field temporarily. The magnetic field lines are closed at the magnetic poles of Earth, and highly energetic charged particles can penetrate deeper into the atmosphere at the geomagnetic poles. This causes aurorae around the magnetic poles, which are increased during solar energetic particle (SEP) events. The currents during SEP may become so high that they can cause power outages on Earth. During a very strong SEP in 1989, e.g., the power grid was interrupted in Quebéc, Canada, for more than 9 hours. Geomagnetic storms may also have severe consequences for satellites.

3 Extraterrestrial material in the Earth's atmosphere

3.1 Meteoroids

Meteoroids crossing the Earth's orbit enter the Earth's atmosphere every day and add several tons of extraterrestrial material into the Earth's atmosphere, which eventually ends at the Earth's surface. An extensive introduction into this topic can, e.g., be found in Ceplecha et al. (1998) or Murad and Williams (2002).

Two main sources have been identified as the origin of the meteoroids. One source are the dust trails of comets orbiting the Sun close to the Earth's orbit, which come close to Earth in time spans of roughly 100 years. These dust trails pass by the Earth at certain time periods of the year, which leads to meteor showers in the Earth's atmosphere. There are at least 20 meteor shower periods known, and the strongest showers are listed in table 3.1. The most famous of these showers are

Name	Dates	Max (2012)	Max ZHR	Associated celestial body
Quadrantides	28 Dec - 12 Jan	4 Jan	120	minor planet 2003 EH_1/ comet C/1490 Y1
Geminids	7 Dec - 17 Dec	13 Dec	120	3200 Phaeton
Perseids	17 Jul - 24 Aug	12 Aug	100	Comet Swift-Tuttle
Eta Aquariids	19 Apr - 28 May	5 May	65	Halley's Comet
Arietids	22 May - 2 July	7 Jun	54	unknown
Leonids	6 Nov - 30 Nov	17 Nov	15	Comet Temple-Tuttle

Table 3.1: List of meteor showers with a zenithal hourly rate > 50 and Leonids (Taken from Wikipedia on 27.2.2014).

probably the Perseids and the Leonids. The Perseids have a high zenithal hourly rate (ZHR) and can be observed in mid summer, where, at least in central Europe, mild weather conditions favor night observation by a larger audience than the other strong showers, which occur in winter. The Leonids, which in most years have a lower ZHR than the Perseids, show a drastically increased ZHR roughly every 33 years, when the Earth crosses the trail of the comet Tempel-Tuttle. This was most recently the case in the years 1999 and 1966. In 1833, ZHRs of more than 100,000 were reported (roughly 300 per second).

The second source of the meteoroids are trails of long-decayed comets and the asteroids from the asteroid belt between the orbits of Mars and Jupiter. Although not as spectacular as the meteor showers, this rather constant source of meteoric material input dominates the total influx into the Earth atmosphere and, e.g., Nesvorný et al. (2010) claims that roughly 85% of the meteoric input on Earth has its origin in the Jupiter family comets. The dominance of the constant background signal over the meteor shower variations was also shown in Scharringhausen (2007), Correira et al. (2010), and other observations. However, sometimes an increased metal density in the upper atmosphere can be observed (e.g., Huang, 2011). Although the material input does not appear to change significantly during meteor showers,

the fraction of larger meteoroids in the size distribution is increased, which leads to larger meteors, that are more easily observed by eye from the ground.

Meteoroids have a large span of size and mass distribution ranging from sub-µm particles to huge impactors with several 10 meters diameters. The mass distribution has two maxima, with a first maximum for particles with a diameter of roughly 220 µm and of 10^{-8} kg (see, e.g., Love and Brownlee, 1993). The second maximum in the mass distribution is for huge impactors with masses of gigatons (see, e.g., Plane, 2012; Murad and Williams, 2002; Kyte and Wasson, 1986), which corresponds to a diameter of roughly 120 m. Different meteoroids in the mass and size distribution hit the Earth at very different time intervals and, furthermore, interact very differently with the Earth's atmosphere. Therefore, there is no single instrumental technique, which can cover the whole span of the particle size distribution. Most techniques have to work with further assumptions to estimate the daily input rate of material. The available studies on meteoric input show a range from roughly 2 to 300 tons per day (see, e.g., Plane, 2012), which is roughly a factor of 100 difference.

When the meteoroids enter the Earth's atmosphere, they collide with air molecules. This meteoroid-Earth interaction is a ballistic collision rather than one initiated by gravitational capture. Therefore, the entry velocities are typically very high and range between 11 and 72 km per second. The collision with air molecules leads to frictional heating and deceleration of the meteoroid and, in part, also to ionization of the meteoric material. The process, which the meteoroid undergoes in the atmosphere, depends on the size, mass, velocity, and entry angle of the meteoroid, as well as its composition, and the boiling points of its ingredients.

Small meteoroids decelerate fast enough that they do not melt and can reach the Earth's surface without any signs of ablation. This type of meteoroid is therefore hardly distinguishable from Earth material. However, this part of the mass-distribution can be estimated from impact crater experiments in space (see, e.g., Love and Brownlee, 1993).

Larger meteoroids ablate partly in the mesosphere and lower thermosphere, which leads to meteors, that are luminous phenomena created by emission of radiation from the ablated gaseous trail of the core meteoroid. Beside electromagnetic radiation, the meteoroid also produces shock waves when it moves with supersonic speeds.

If the meteoroid partly survives the heating phase, it cools down quickly and continues a dark flight until it hits the ground. The parts of the meteoroids that reach the ground are called meteorites. These meteorites typically have a melted surface. Furthermore, some meteorites can be identified by the structures of their metal composition, as they show patterns, e.g., so called Widmannstätten patterns, that can only be formed under extreme conditions that are not present on Earth. Although meteorites are distributed all over the world, they are typically found in regions with reduced weathering, e.g., deserts and polar ice regions.

One most recent well-documented large meteoroid-event occurred close to the city of Chelyabinsk, Russia, on February 15, 2013. Plenty of photo and video footage of the meteor and its smoke trail exists, which could be seen from distances of several hundreds kilometers. The original meteoroid exploded in an air burst at around 30 km altitude and the shock waves caused broken windows in several nearby cities.

The meteoroid burst into many fragments, which entered dark flight, before they hit the ground. It is estimated, that the diameter of the meteoroid was on the order of 20 m and the mass on the order of 10,000 tons. The largest fragment was found in a lake and had a mass of 654 kg.

Remnants of even larger meteoroids, which hit the Earth, can be found at certain locations on Earth. Very large meteoroids decelerate slowly enough that they hit the ground with supersonic velocities. This leads to the formation of craters, which, if large enough, have a central mountain, which is formed by the oscillation of the ground after the impact. Depending on the time and weathering conditions in the crater region, these craters can still be easily seen from aerial pictures. Table 3.2 shows a list of confirmed large meteoric impact craters on Earth, which have diameters of several hundreds of kilometers, and the Earth Impact database (EID, 2014) lists ≈ 200 confirmed impact sites.

Name	Country	Coordinates	Est. Age in mill. y	diameter in km
Vredefort	South Africa	27°S 27°E	$2000 - 3500$	300
Sudbury	Canada	46°N 81°W	1850	250
Chicxulub	Mexico	21°N 89°W	65	180
Kara	Russia	69°N 64°E	70	120
Manicougan	Quebec	51°N 68°W	215	100
Popigai	Russia	71°N 111°E	36	100

Table 3.2: List of large confirmed meteoric impact craters (Taken from Wikipedia on 28.2.2014).

3.2 Metals in the mesosphere and lower thermosphere

Meteoric metals are ablated in the mesosphere and lower thermosphere (a recent review was given by Plane et al., 2015). Model studies, e.g., by Vondrak et al. (2008) simulate the ablation process and predict the fraction of the ablated material as well as the altitude of the ablation. The injection rates for the different metal species can be calculated from these models, if they are combined with models/measurements on the distribution of the critical parameters of the meteoroids like velocity, size, density, composition, entry angle, etc. (a detailed description of this can be found, e.g., in Janches et al. (2006), Fentzke and Janches (2008) and Fentzke et al. (2009)). The vertical injection rate profiles for different metals have a maximum between $\approx$80-95 km with a FWHM of ≈ 10 km and metals with lower boiling points ablate at higher altitudes (see Fig. 3.1). The integrated vertical injection rate profile, or column injection rate, shows a seasonal cycle, which is stronger pronounced with higher latitude, with an autumn maximum (see Fig. 3.2). The injected metals undergo chemical and transport processes. The simulations by Vondrak et al. (2008) show, that a small fraction of the ablated metal atoms are ionized by hyperthermal collisions, but the larger fraction of injected metals are atoms. The upper edge of the metal's injection profile lies in the lower thermosphere, where ionizing solar radiation can penetrate the atmosphere. The metals there can either be directly ionized by radiation, or more likely undergo charge exchange reactions with the

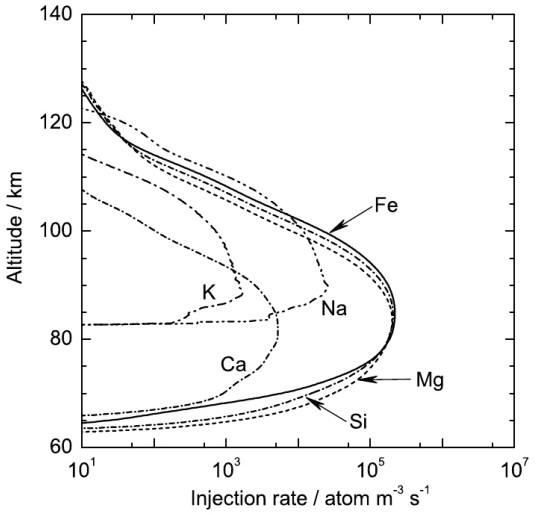

Fig. 11. Injection rates of individual elements, integrated over the LDEF distributions of meteoroid mass (5×10^{-18}–5×10^{-3} g) and velocity (11.5–72.5 km s^1). Note that these injection rates should be scaled by a factor of 0.34 if the total mass flux is 30 t d^{-1} (see text).

Fig. 3.1: Copied from Vondrak et al. (2008). LDEF is the Long Duration Exposure Facility, an orbital impact detector placed on a spacecraft for several years.

dominant ion species NO^+ and O_2^+ (see, e.g., Plane and Helmer, 1995). The metal atoms and ions quickly react with the main reactive species in this region, i.e., O and O_3. In further reactions, for example with H_2O and CO_2, the metal oxides react to form more stable metal-hydrates, -carbonates and, -hydrogen-carbonates. Figure 3.3 summarizes the chemical reaction pattern for Mg. Similar schemata, for other metals are, e.g., presented by Plane (2003) (Fig. 14 for Na and Fig. 15 for Fe). Tables of chemical reactions and reaction rates are, e.g., presented in Langowski et al. (2014b) for Mg and in Feng et al. (2013) for Fe.

The equilibrium between cosmic metal atom injection and chemical loss reactions leads to the formation of stable metal atom layers at around 80 to 95 km altitude and metal ion layers 5-15 km above the neutral metal layers. The metal atoms and ions are studied through remote sensing methods. However, the metal molecules can hardly be accessed, because their emissivities are much smaller, and the metal molecule densities are only known from models. Beside the stable metal atom and ion layers sometimes sporadic metal layers are formed. These layers typically form at slightly higher altitudes than the stable layer and in extreme cases may drastically exceed the peak densities of the stable metal layer. The sporadic metal ion layers, which often occur during meteor storms, are very closely related to the sporadic E-layers (E_s) discussed in Sect. 2.6. Although the total number densities of the meteoric metal species are only in the order of several 10000 particles per cubic centimeter, they provide valuable information as tracers for mesospheric transport processes and also have significant influence on other middle atmospheric phenomena.

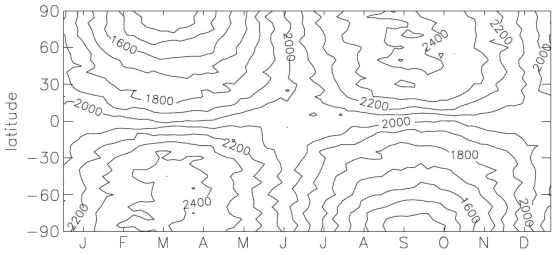

Figure 1. Seasonal and latitudinal distribution of meteoric Na column injection rates (atoms cm^{-2} s^{-1}) used in the WACCM simulation.

Fig. 3.2: Copied from Marsh et al. (2013a).

Meteoric metals are observed with the detection and ranging technique from the ground using either radio (radar) or visible (lidar) radiation, which is emitted by a transmitter and the response signal is collected. The distance of a reflecting body is determined by the delay of the signal, while the strength of the response signal is related to the extension or density of the reflector. The difference between using long-range radio or shorter visible radiation is, that the radar rather detects the backscattered signal from larger bodies (like meteors, ionization trails, airplanes etc.), while the visible light stimulates resonance fluorescence of single atom, ion or molecules and is more sensitive for the individual different constituents of the target. Radar observations of meteors, which are available since the 1940s, provide, e.g., valuable information on mesospheric wind speeds, temperatures and momentum fluxes. The airglow from resonance fluorescence of the sodium D lines at 589 nm has been observed with photometers since 1929 (see, e.g., Slipher, 1929) and metals were quantitatively investigated beginning in the 1950s; this technique was superseded by lidar techniques in the 1970s (see, e.g., Plane, 2003). The big advantage of the lidar technique is the active, well known and tunable light source for the resonance fluorescence of the metals, which enables the observation of the metal layers at night and also does not have other disadvantages of natural light sources (see, e.g., Sect. 6.2). Lidar observations of specific metal species provide, e.g., metal number densities, temperatures and wind speeds. Note, that different instruments may only provide a subset of this information. Ground-based measurements have a good vertical and temporal coverage, and are well suited to investigate short time wave phenomena and other phenomena like polar mesospheric clouds, sporadic metal layers etc.. A disadvantage of the ground-based measurements is, that they are very local in latitude and longitude. In contrast to the ground-based observations, satellite observations, like those presented in this thesis (see Sect. 4 for the satellite description), provide nearly a daily global latitudinal and longitude coverage, but have worse vertical and temporal resolution as a tradeoff. The discussion of metal species retrieval and the analysis of the retrieved data is the largest part of this thesis (Sects. 6 to 10).

The metal atoms and ion layers in the mesosphere and lower thermosphere are observed. The metal molecules, however, can hardly be accessed, as their emissivities of radiation are much lower than the emissivities of atoms and ions. Through

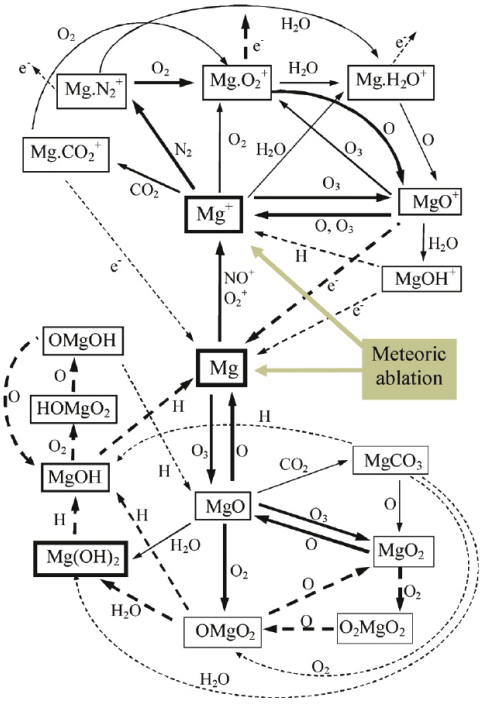

Figure 9. Schematic diagram of magnesium chemistry in the upper mesosphere/lower thermosphere region. Major magnesium species are shown in boxes with bold outlines. Important reaction pathways are indicated with thicker arrows. Reactions with measured rate coefficients are indicated with solid arrows; broken arrows indicate pathways for which rate coefficients are estimated or fitted. Note the role of O_2 in the MgO_2–O_2MgO_2–$OMgO_2$ and MgOH–$HOMgO_2$–OMgOH holding cycles.

Fig. 3.3: Copied from Plane and Whalley (2012).

further reactions, the metal molecules form mesospheric smoke particles (MSP) (see, e.g., Plane, 2011; Saunders et al., 2012). The MSP act as nucleation nuclei for heterogeneous nucleation in clouds and can be observed in NLCs (see, e.g., Hervig et al., 2012).

In brief, cloud droplets or chrystals (which is the case for NLCs) need a certain minimum diameter to grow, as the binding energy, which comes from the volume of the droplet/chrystal must be bigger as the surface tension energy, so that effectively small droplets/chrystals shrink, while larger droplets/chrystals grow. If only water molecules are involved the process is called homogeneous condensation/nucleation. A necessary condition of the formation of clouds is a supersaturation of water, which, following the Clausius-Clapeyron-equation, leads to the phase transition from gaseous to liquid water. However, the Clausius-Clapeyron-equation does not state how fast the phase transition occurs, and for a quick phase transition a large supersaturation is necessary.

A nucleation nucleus, like an ion or a highly polar molecule, which attracts water molecules, increases the diameter of the initial droplet, which is called heterogeneous nucleation. This heterogeneous nucleation needs much less supersaturation of the water vapor to quickly form clouds. MSP not only directly increase the for-

mation rate of clouds, but also indirectly by forming other aerosol particles in the stratosphere, which are condensation nuclei for heterogeneous condensation itself. Thus, meteoric metals may play a significant role in the formation of polar stratospheric clouds, which play an important role in the ozone chemistry (see Sect. 2.3). The role of the meteoric metals needs to be quantified. Figure 3.4 summarizes the impact of meteoric metals on the Earth's atmosphere.

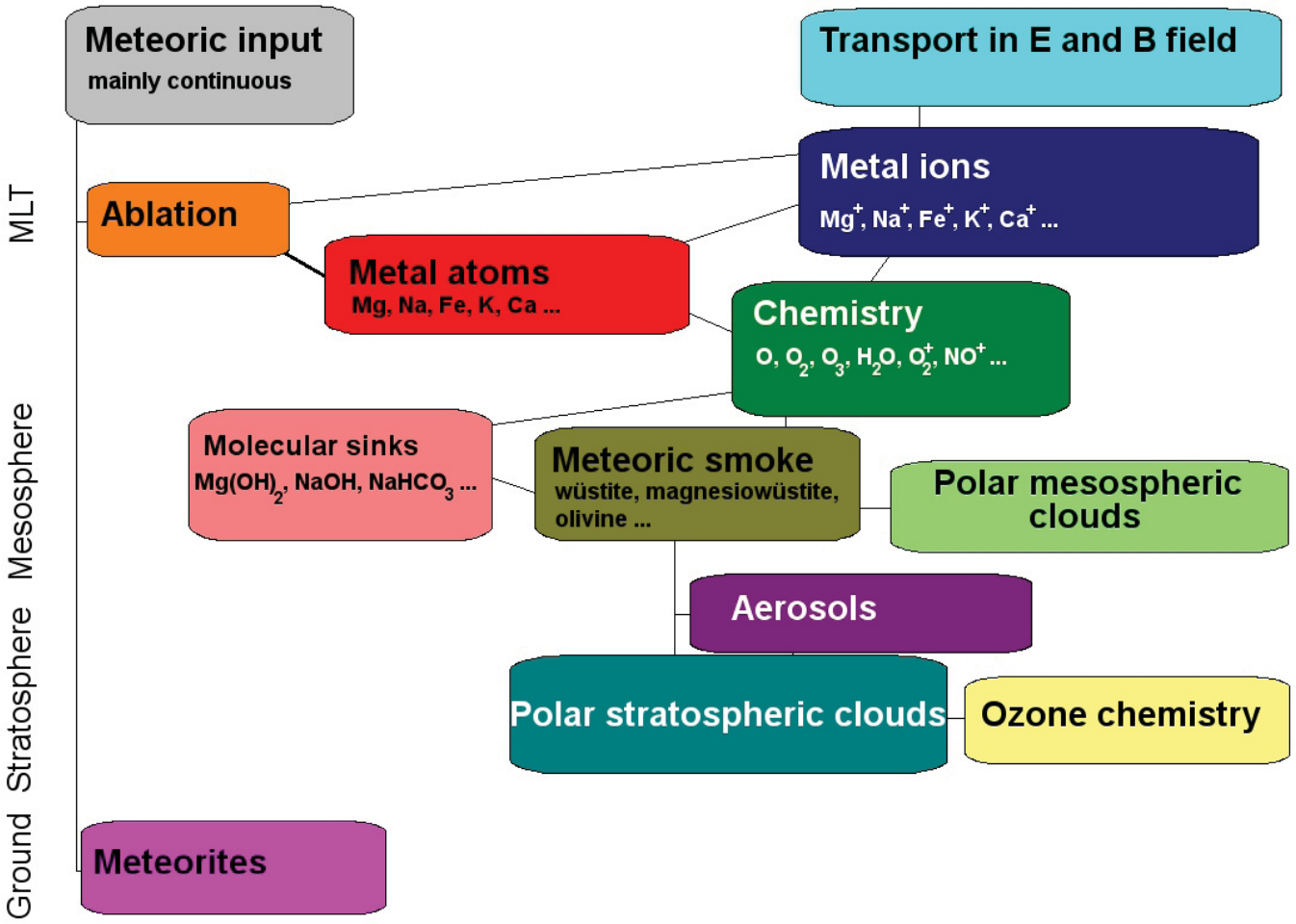

Fig. 3.4: Sketch of the role of meteoric metals in the Earth's atmosphere.

4 SCIAMACHY/Envisat

The SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY, SCIAMACHY is a grating spectrometer on board the European research satellite Envisat (see, e.g., Burrows et al., 1995; Bovensmann et al., 1999). Envisat was launched on February 28, 2002 from Kourou, French Guiana, aboard an Ariane 5 rocket into a Sun-synchronous orbit at an altitude of about 800 km with an equator crossing time for the descending node (moving from north to south over the equator) of 10 : 00 a.m. and 10 : 00 p.m. for the ascending node. The local time, shown in Fig. 4.1, varies only strongly in the polar regions, where morning times are passed in the Southern Hemisphere and evening times are passed in the Northern Hemisphere. SCIAMACHY needs roughly 100 minutes for one orbit around Earth, which results

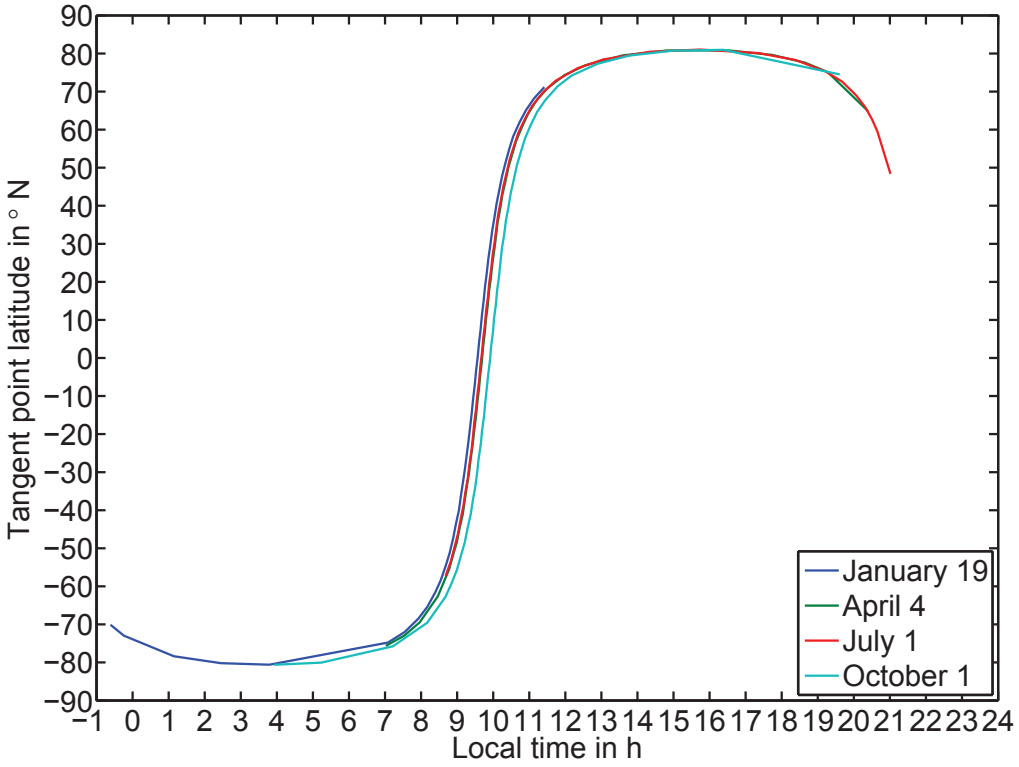

Fig. 4.1: Local time at the tangent point of SCIAMACHY limb scans.

in 14-15 orbits per day. In late October 2010 a last orbital maneuver took place, which shifted the satellite slightly away from its Sun synchronous orbit to reduce the risk of collisions on this strongly populated orbit. However, even after this maneuver the orbit was still close to Sun synchronous. After performing very well for a decade without any big gaps in data coverage, the contact to Envisat was lost on April 8, 2012 and ESA formally announced the end of the mission on May 9, 2012.

SCIAMACHY consists of 8 single spectrometer channels, each with an individual wavelength range and spectral resolution shown in table 4.3. The sampling rate, which is the number of pixels sampled in the resolution width, is about 2. This means, e.g., that the distance of neighboring pixels in channel 1, where the resolu-

Channel	Spectral range (nm)	Resolution(nm)
1	214-334	0.22
2	300-412	0.26
3	383-628	0.44
4	595-812	0.48
5	773-1063	0.54
6	971-1773	1.48
7	1934-2044	0.22
8	2259-2386	0.26

Table 4.3: Wavelength range and resolution of SCIAMACHY science channels. This includes also regions at the edges of each channel that can only be used for monitoring purposes.

tion is 0.22 nm, is 0.11 nm. SCIAMACHY has three different observation-geometry modes: limb, nadir and occultation. In the occultation mode the instrument points directly towards a light source like the Sun or the Moon. In nadir the instrument points nearly downwards towards Earth, while in limb mode the instrument points tangentially to the Earth's surface at different tangent altitudes. The nadir mode has the advantage of measuring total columns with a very good latitudinal and longitudinal resolution, while it provides nearly no vertical resolution (beside for O_3, where vertical resolution is retrieved from different wavelength regions, which are simultaneously observed). Good vertical resolution is the big advantage of the limb mode. However, the vertical resolution is only good between the lowest and the highest tangent heights of a series of consecutive limb measurements. This series of typically 31 consecutive limb measurements, which include one dark signal measurement, with SCIAMACHY is abbreviated with the term "limb state" or "limb scan".

At the beginning of the SCIAMACHY data set in 2002, there are limb scans available from 0 km up to about 105 km. However, the maximum tangent altitude was later changed to about 91 km altitude. In summary, daily limb measurements are available from 2002 to 2012 from 0 to 90 km in 30 consecutive 3.3 km steps. After the orbital maneuver in October 2010 the nominal limb mode changed to 29 steps, omitting the lower most step, as the orbit time was shorter. One additional measurement is done at 250 km altitude. At this altitude no signal is expected anymore and the signal at this altitude is subtracted from the other limb measurements as a part of the dark signal correction. Spectral structures with similar shape as the solar irradiance spectrum can be found in all dark signal measurements, however, typically in the first three to four states after sunrise this solar stray light contamination is extremely strong. The reason for this is most probably, that the Sun is in the field of view for these states.

Since mid 2008 a new limb mode for mesosphere and lower thermosphere (MLT) observation was performed, which ranges from 53 km to 150 km in 30 consecutive 3.3 km steps for the full remaining period until 2012. The dark signal measurement for the MLT mode was shifted to 350 km tangent altitude. The MLT mode

was performed nearly every two weeks for 15 consecutive orbits, which is roughly one day of observations. This resulted in 84 single days of MLT data. Example spectra from the MLT observations will be shown in the next section.

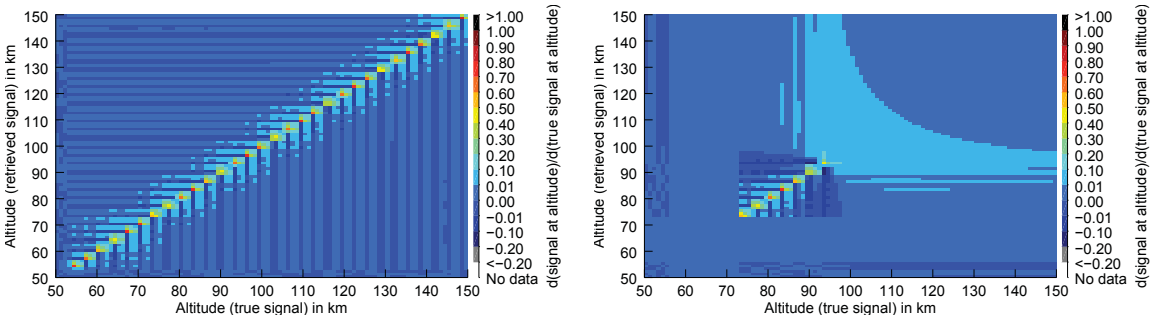

Fig. 4.2: left: Resolution matrix (also known as averaging kernel matrix) for a vertical retrieval grid from 50 km and 150 km and 30 limb measurements at tangent altitudes between 53 km and 150 km. right: Resolution matrix for the same retrieval grid, but only using 7 limb measurements at tangent altitudes between 72 km and 92 km. For the calculation of the shown matrices a singular value decomposition was used.

The advantage of the MLT-mode compared to the nominal mode in the altitude region of the metal and metal ion layers around 80 km to 110 km is illustrated in Fig. 4.2. The left panel in Fig. 4.2 shows the resolution matrix (which is also known as the averaging kernel matrix) for a retrieval of a typical set of 30 consecutive MLT-measurements. The y-axis represents the altitude at which the density is retrieved, while the x-axis is the true altitude of the signal. The ratio $\frac{\text{retrieved radiance at altitude}}{\text{radiance at true altitude}}$ shows from which actual slant column densities at certain altitudes the retrieved slant column density at an altitude is composed (For example a value of 0.5 at $x = 90$ km and 0.5 at $x = 91$ km and $y = 95$ km means that the signal retrieved at 95 km actually comes from the region between 90 km and 91 km). For a retrieval with a finer vertical sampling of 1 km for measurements than the tangent altitude difference of around 3.3 km the resolution matrix shows nearly only large entries on the diagonal elements, with a better match close to the actually measured tangent altitudes. This means, that the retrieved signal at an altitude also comes from this altitude region, but there is some vertical smoothing of the profile, because the resolution can not be better than the vertical distance between two consecutive limb measurements. Note, that due to the field of view of the instrument, which is not considered here, the retrieved profile is additionally smoothed.

The right panel of Fig. 4.2 shows the resolution matrix, when only the 7 measurements at tangent altitudes between 72 km and 92 km are used. Altitude regions below the lowest tangent altitude can not be resolved at all, and although each altitude above the highest tangent point has some contribution to the line of sight, the resolution for higher altitudes than the highest tangent point's altitude vanishes quickly with increasing altitude. Also note, that if the retrieval grid was chosen the same as the actual tangent altitude grid, the matrix was diagonal and the diagonal

elements were very close to 1 (tested, but not shown here). However, the finer resolution brings less numerical errors in the calculation of the absorption processes, which are exponentially depending on the path length in each cell and can be better approximated with small grids.

5 Emission lines in the SCIAMACHY spectrum

In this section the SCIAMACHY limb spectrum is investigated for emission lines at altitudes from 70 to 150 km. It should be noted, that a large part of the found spectral features in this section look similar to emission features, but more likely can be explained by other features that look very similar to emission features, and are therefore hard to separate. Nevertheless, these characteristic features in the spectra may be exploited for different analysis than emission, so it might still be advantageous to know these features.

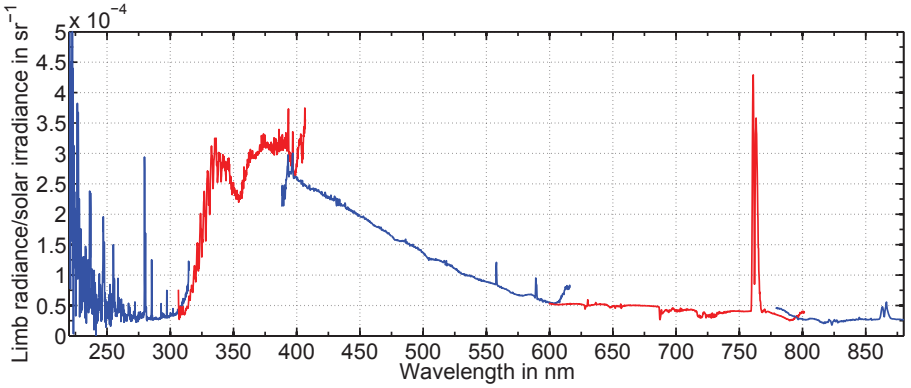

Fig. 5.1: SCIAMACHY limb radiance divided by solar irradiance spectrum between 220 and 880 nm wavelength at tangent point altitude of 90 km and equatorial latitudes. The limb data is averaged over the whole period of limb mesosphere and lower thermosphere measurements.

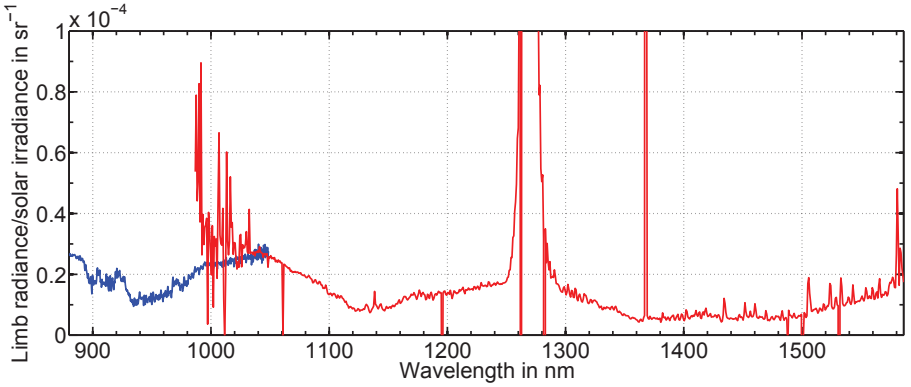

Fig. 5.2: SCIAMACHY limb radiance divided by solar irradiance spectrum between 880 and 1586 nm wavelength at tangent point altitude of 90 km and equatorial latitudes. The limb data is averaged over the whole period of limb mesosphere and lower thermosphere measurements. The figure is cut off at the upper end, to make weaker emission features better visible.

Figures 5.1 and 5.2 show the averaged limb radiance spectra at 90 km altitude and at equatorial latitudes divided by the SCIAMACHY solar irradiance spectrum

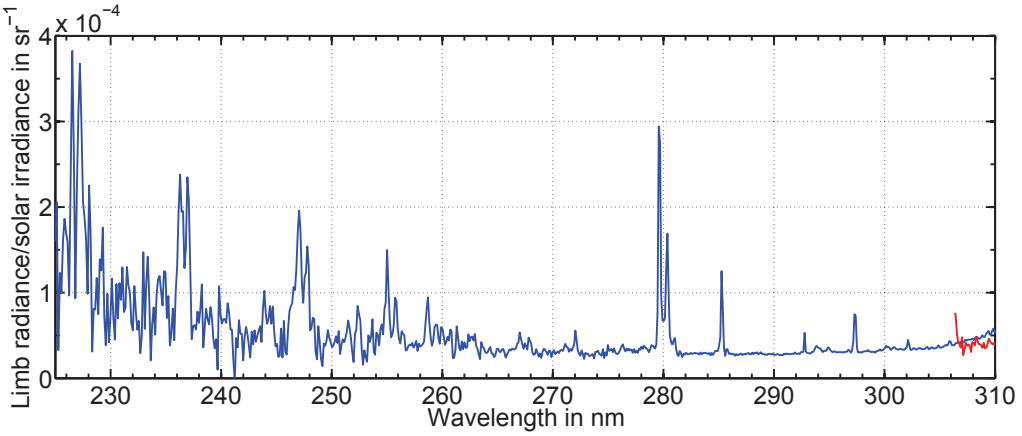

Fig. 5.3: SCIAMACHY limb radiance divided by solar irradiance spectrum between 240 and 310 nm wavelength at tangent point altitude of 90 km.

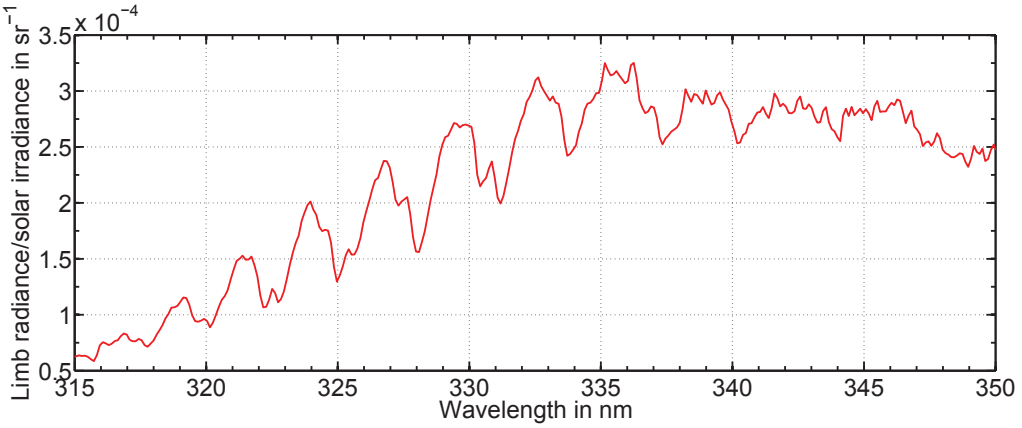

Fig. 5.4: SCIAMACHY limb radiance divided by solar irradiance spectrum between 315 and 350 nm wavelength at tangent point altitude of 90 km.

for the period of mesosphere and lower thermosphere (MLT) measurements from 2008 to 2012. The same altitude and latitude region is used for the other plots in this section, if not otherwise stated. In the figures, the different channels of SCIAMACHY alternate in color to indicate the edge regions of the channels, which do not show reasonable signals. The limb spectrum is mainly formed by Rayleigh scattering by air atoms and molecules which leads to a variation of the ratio spectrum for wavelength λ of a power of λ^{-4}. For small wavelength windows, e.g., of 1 nm width at around 300 nm this variation is rather weak compared to quickly varying signatures, which makes both signatures in the spectrum separable.

The the bulk of the quickly varying signatures can be reasoned by one of the following effects:

1. emission by atoms or molecules
2. absorption by atoms or molecules
3. inelastic Raman-scattering in Fraunhofer lines (Ring effect) (see Sect. 6.2.2 for

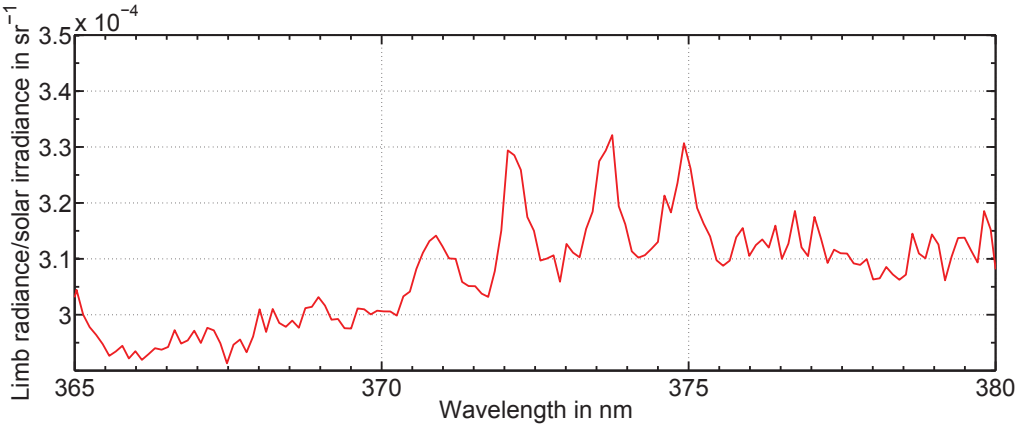

Fig. 5.5: SCIAMACHY limb radiance divided by solar irradiance spectrum between 365 and 380 nm wavelength at tangent point altitude of 90 km.

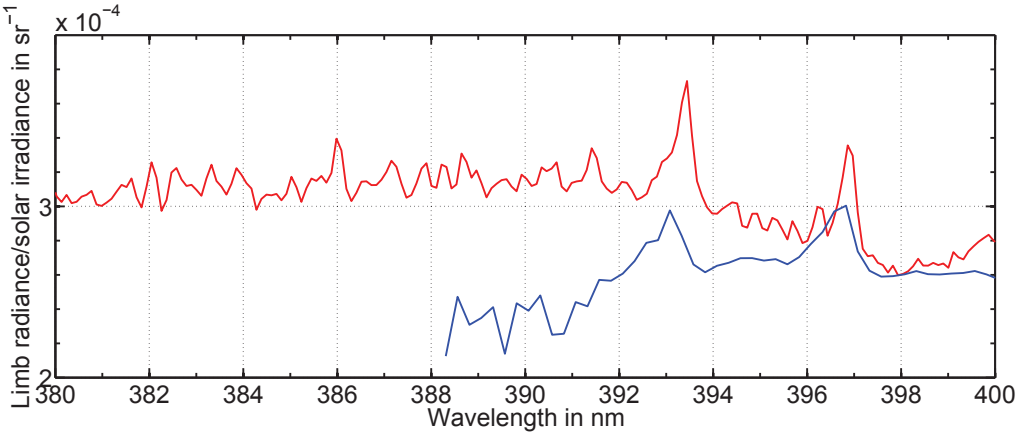

Fig. 5.6: SCIAMACHY limb radiance divided by solar irradiance spectrum between 380 and 400 nm wavelength at tangent point altitude of 90 km.

more details on this)
4. small wavelength shifts between the limb and the solar irradiance spectra
5. effects of folding the spectrum with the instruments slit function at a region with strong spectral gradient, e.g., a Fraunhofer line (so called I_0 effect)
6. stray light contaminations from different altitude regions of the atmosphere
7. multiple scattered radiation above 300 nm wavelength

The mesospheric metal emission line wavelengths typically lie in the wavelength range of solar Fraunhofer lines, because the solar atmosphere includes the same metal species that can be found in the Earth's mesosphere. Therefore, the weighting of process 1 and 3 to the total spectral signature must always be considered for metal emission lines.

Fig. 5.3 shows the spectrum in the UV region between 220 and 310 nm, where O_3 absorption in the stratosphere is strong enough that a single scattering approach for the emission process of the emission lines in this regions can be used

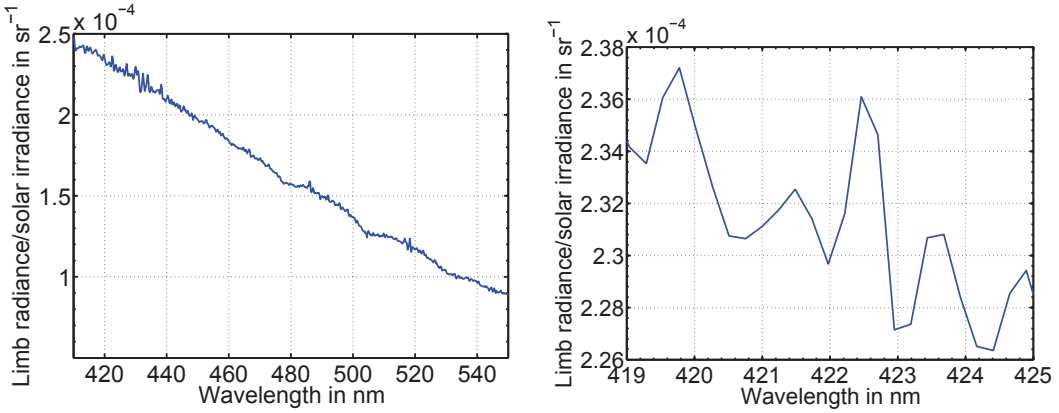

Fig. 5.7: SCIAMACHY limb radiance divided by solar irradiance spectrum between 410 and 550 nm wavelength at tangent point altitude of 90 km. The right figure is a zoom in for the Ca line at 422 nm.

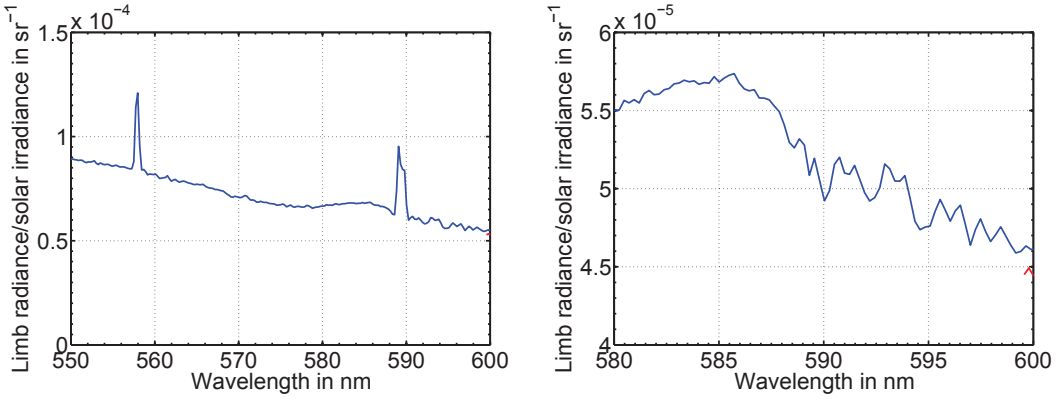

Fig. 5.8: left: SCIAMACHY limb radiance divided by solar irradiance spectrum between 550 and 600 nm wavelength at tangent point altitude of 90 km.
right: Wavelength region between 580 and 600 at tangent point altitude of 115 km

(see Sect. 6.2.1). Unfortunately the incoming solar radiance has a low intensity in this spectral region and SCIAMACHY has a strong instrumental degradation there, so that the signal to noise ratio is low for shorter averaging periods of the MLT measurements (the nominal limb measurements are performed more frequently). The strongest emission lines in this wavelength region are the NO-γ-band lines and the Mg/Mg$^+$ lines.

The NO-γ-band lines are formed by the electronic transition from the first excited state ($A^2\Sigma^+$) to the ground state ($X^2\Pi$) of NO. The different resolvable double lines correspond to transitions between different vibrational states, while the rotational structure is not fully resolved with SCIAMACHY resolution. The nomenclature for vibrational transitions is (upper vibrational state, lower vibrational state). The following NO-γ-band lines show clear emission features in Fig 5.3: (2,2) at 222 nm, (0,0) at 227 nm, (0,1) at 237 nm, (0,2) at 247 nm, (1,4) at 255 nm, (0,3) at 258 nm, (1,5) at 268 nm, (0,4) at 272 nm, (1,7) at 295 nm and at 301 nm. There

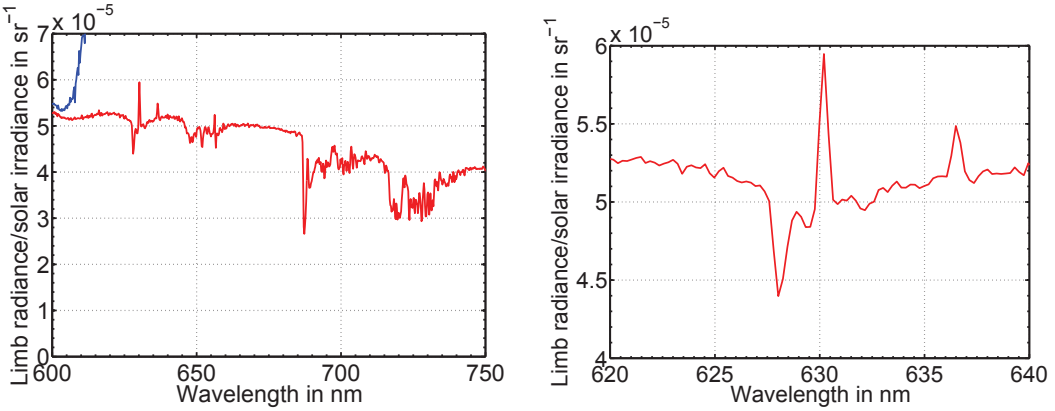

Fig. 5.9: SCIAMACHY limb radiance divided by solar irradiance spectrum between 600 and 750 nm wavelength at tangent point altitude of 90 km.

are also relatively weak NO lines at 280 nm (1,6) and 285 nm (0,5) at the edge of the stronger Mg/Mg$^+$ emission lines. Bender et al. (2013) used the (0,2), (1,4) and (1,5) transition lines to successfully retrieve NO densities from the SCIAMACHY MLT data set.

The Mg/Mg$^+$ lines at 280 and 285 will be comprehensively discussed in Sect. 6 (see, e.g., Fig. 6.3). Beside the NO and the Mg lines, other weaker lines can be found in Fig. 5.3. Several potential Fe lines can be found in this wavelength region, e.g., at 248 nm, 252 nm, 272 nm and 302 nm, however, these lines either overlap with nearby stronger NO lines or have a very low signal to background ratio, which introduces issues with the Ring effect correction (see Sect. 6.2.2). Other non metal emission lines in this wavelength region are the O line at 297 nm, as well as the OH line at 308 nm (The line at 292 nm, can only be seen at 90 km and is not real.). The O line peaks above the highest tangent altitude of 150 km and, therefore, it can hardly be investigated. The nominal limb measurements are more suited for an investigation of the 308 nm emission line, because significant OH emissions can only be found below 90 km.

The wavelength region between 300 and 400 nm is covered by SCIAMACHY channel 2. For a large wavelength window around 350 nm a dip in the limb radiance to solar irradiance ratio is observed. This is most likely due to an issue with the polarisation correction at this wavelength. The spectra in channel 2 seem to be very noisy. However, most structures that can be seen in the 4 year average can also be seen in daily averages, so they seem to be rather systematic and are related to Fraunhofer-structure that does not fully cancel out when forming the ratio spectrum. Small wavelength shifts of the limb spectrum to the solar spectrum, however, reduce structures, which is a further sign that the variations in the spectrum are not just noise. These small constant shifts work only for small wavelength windows. The most striking features in this wavelength regions are the Huggins ozone-absorption bands between 315 and 350 nm shown in Fig. 5.4 (compare to ozone absorption cross section in Fig. 6.1). It was recently reported by Martus and Collins (2013), that a Ni emission line at 337 nm was found in lidar observations. However, the line strength

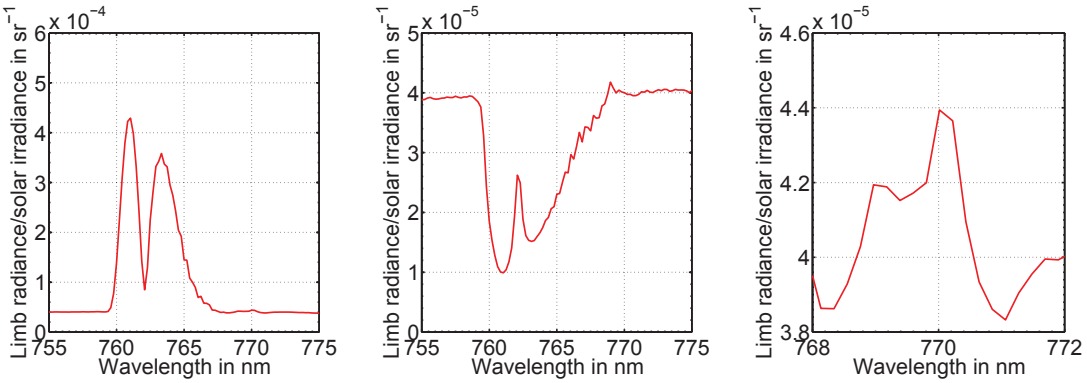

Fig. 5.10: Left panel: SCIAMACHY limb radiance divided by solar irradiance spectrum between 755 and 775 nm wavelength at tangent point altitude of 90 km.
Middle panel: SCIAMACHY limb radiance divided by solar irradiance spectrum between 755 and 775 nm wavelength at tangent point altitude of 150 km.
Right panel: SCIAMACHY limb radiance divided by solar irradiance spectrum between 768 and 772 nm wavelength at tangent point altitude of 90 km.

defining parameters are rather weak compared to other metal lines and even if the peak at 337 nm can be assigned to Ni it is rather complicated to separate the Ni line from the background.

Fe and Ca^+ lines can also be found in this wavelength region. Fig. 5.5 shows the 372 nm Fe line. However, beside the 372 nm line there are 3 equally strong Fe lines at 371, 374 and 375 nm. These lines are not ground state transitions and the lower states of these transitions should not be populated strongly enough at mesospheric temperatures, that these emission lines can be seen at all. The 372 nm and 374 nm line are used in lidar observations to determine temperatures and the density of the excited state retrieved from the 374 nm line is much lower than the one of the ground state. One hypothesis to explain this spectrum is, that these lines are mostly formed by inelastic rotational Raman scattering, i.e., the Ring effect, rather than by emission. The O_3 cross section furthermore has one of the lowest values in the UV/Vis range at 370 nm and there are no other strong absorbers at this wavelength, so that multiple scattering is strong there. The same issues most probably also apply for the Ca^+ lines at 393 nm and 396 nm.

Few emission features can be observed between 400 nm and 550 nm shown in Fig. 5.7, and the Ca line at 422 nm shows a low signal to background ratio. There are strong emission lines between 550 nm and 600 nm, shown in Fig. 5.8. There is the very strong green atomic O line at 558 nm and the Na doublet lines at 589 nm, which will be discussed in Sect. 6.5. Between 585 nm and 600 nm 4 doublet "emission" lines can be observed. This structure, however, is most likely coming from multiple scattering originating from the lower atmosphere.

Two red atomic oxygen lines are shown in Fig. 5.9 at 630 nm and 636 nm, but no further strong emission lines can be found up to 750 nm. The strongest atomic Li line is located at 670 nm, but is not identified in the SCIAMACHY spectrum. This is, because the emissivity of this Li line is similar to the emissivities of the Na D

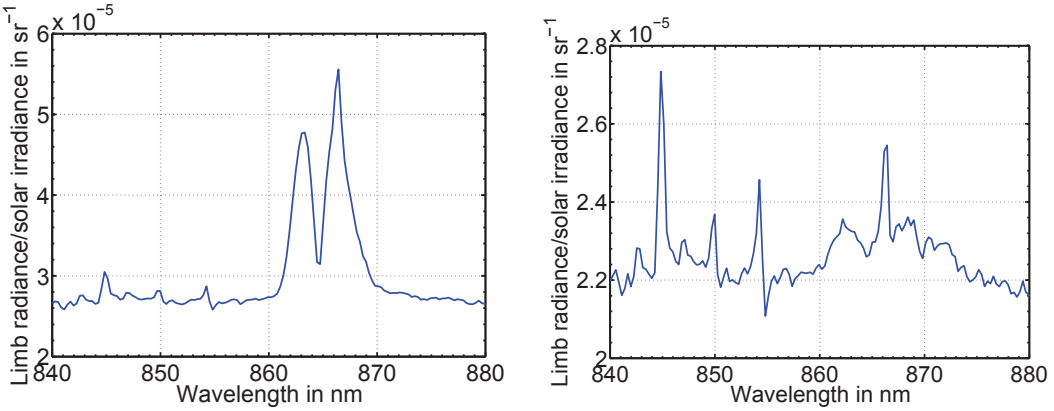

Fig. 5.11: SCIAMACHY limb radiance divided by solar irradiance spectrum between 840 and 880 nm wavelength at tangent point altitude of 90 km (left) and 122 km (right).

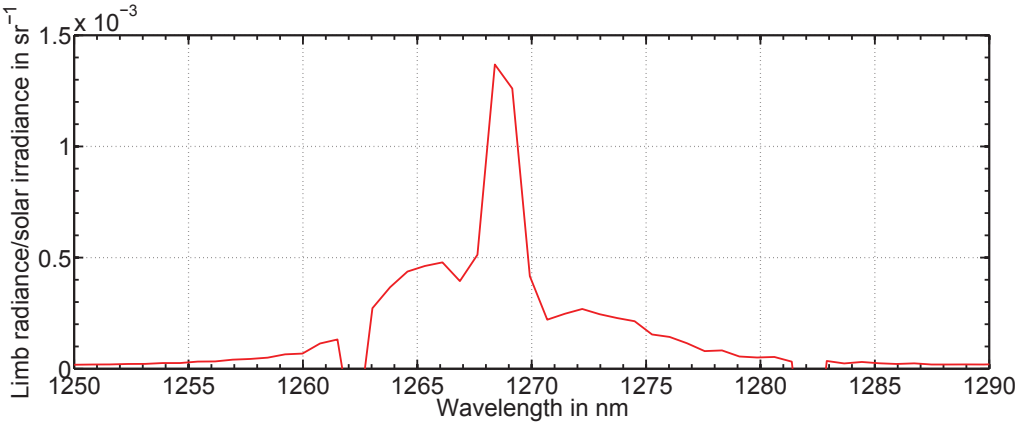

Fig. 5.12: SCIAMACHY limb radiance divided by solar irradiance spectrum between 1250 and 1270 nm wavelength at tangent point altitude of 90 km.

lines at 590 nm, but the densities are at least a factor 10^3 smaller than for Na and, therefore, the signal to noise ratio is too low to identify this line.

Very interesting features are observed in Fig. 5.10 between 755 nm and 775 nm. One of the strongest emission lines, the O_2 A-band is located here. It is very interesting that this band shows emission features in the mesosphere, but change this role from emission to absorption in the thermosphere, and show absorption features at 150 km altitude. However, it is highly likely that this behavior comes from stray light radiance from the lower atmosphere, that is larger than the actual radiance at 150 km (this effect is ,e.g., also shown in Fig. 6.36). Nearby to the comparably huge O_2 emission lines is the small K emission line at 770 nm. Dawkins et al. (2014) showed that it is possible to retrieve reasonable K profiles from this line.

Another O_2-band emission, the $O_2(0,1)$ band emission, can be found at 865 nm, which is shown in Fig. 5.11. This band disappears at higher altitudes and, in contrast to the O_2 A-bands, does not show absorption at higher altitudes. Atomic O at

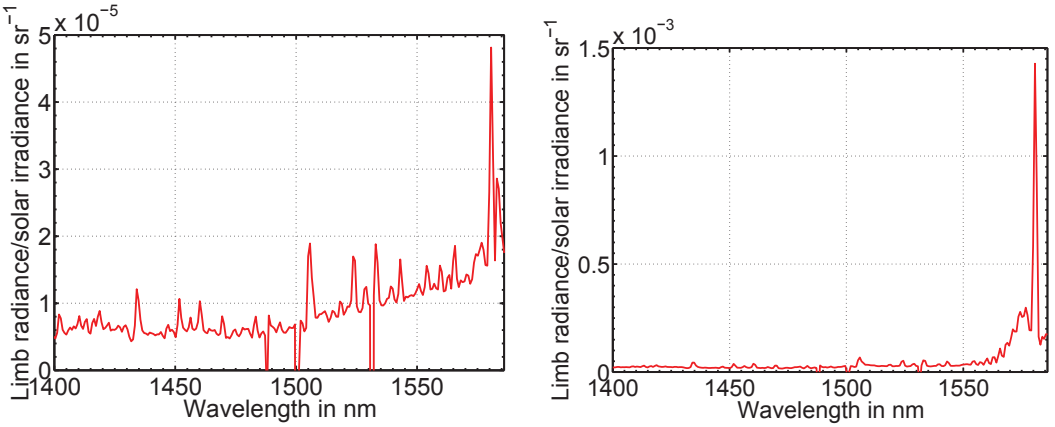

Fig. 5.13: SCIAMACHY limb radiance divided by solar irradiance spectrum between 1400 and 1586 nm wavelength at tangent point altitude of 90 km (left) and 53 km (right).

845 nm and Ca^+ emissions at 850 nm, 855 nm and 866 nm are also present in this wavelength region.

The longer wavelength region is dominated by molecular transition lines. Figure 5.12 shows the strong emission line at 1270 nm with a smaller broad band emission feature between 1265 and 1275 nm. This emission comes from an O_2 singlet transition, with the signature at 1270 nm being the Q-branch, and can also be seen at the highest tangent altitude of 150 km. The strong line in Fig. 5.2 at 1368 nm is an artifact in the solar irradiance spectrum (bad pixel).

Fig. 5.13 shows the remaining wavelength part of the SCIAMACHY spectrum. The most striking feature in this figure is the O2 $^1\Delta_g(0-1)$ band band emission between 1570 nm and 1580 nm which overlaps with a CO_2 band, which however, only shows significant absorption features at lower altitudes, than shown here. For longer wavelengths in SCIAMACHY channel 6 and in channel 7 and 8 there are too many bad pixels in the spectra of the averaged MLT data set to perform a survey of emission lines.

Part II
Number density retrieval algorithm

6 Two dimensional Retrieval algorithm

A comprehensive summary of the 2D retrieval algorithm for metal number densities in the mesosphere from satellite spectra, which has been developed and used for this thesis, has been published by Langowski et al. (2014a). This paper will be used in the following description of the retrieval algorithm. In addition, this chapter includes more details, which have been omitted in Langowski et al. (2014a), because of limited space. The algorithm was first developed for the Mg and Mg^+ lines in the UV region. Therefore, the description of the algorithm for these species is presented first. Afterwards, the extension of the algorithm for Na lines in the visible region is discussed. This discussion for Na was concurrently worked out to this work in Langowski et al. (2015a).

6.1 Outline

The retrieval algorithm used for this work is based on the retrieval algorithm developed by Scharringhausen (2007) in Fortran. The new program was first rewritten in C++ and reproduced the previous results very well for test data. Furthermore, several important features have been found and implemented to improve the retrieval algorithm. The latitude and altitude grid has been adjusted to the geometry of the limb MLT measurements. Only altitude regions in between the highest and lowest tangent altitude of the measurements are considered now, and the vertical grid cells are equidistant (for the nominal states additional grid elements had to be added above the highest tangent altitude, because the region of the highest densities for Mg^+ is above the highest tangent altitude). Furthermore, additional grid elements for the high latitudes have been added to separate grid elements, which have the same latitude but different longitudes, local times, solar zenith angles and scattering angles. A new algorithm to calculate path lengths within the grid elements has been implemented, which calculates the path lengths more directly and requires less calculation steps (and is therefore faster) than the ray-tracing approach, previously used.

Other more important new-implemented features have a significant influence on the actual results and physics. A Ring effect (see, e.g., Grainger and Ring, 1962) correction has been implemented, which has a significant influence on the vertical structure of the retrieved Mg layer at the lower edge. Self absorption of the radiation emitted by the metals, which is absorbed by the same metals (not the same individual particles) is now considered, which effectively increases the retrieved densities. The resolution of SCIAMACHY of 0.22 nm in channel 1 is a factor 10 to 1000 worse than the actual width of the solar absorption and emission features in the used solar spectra, as well as the emission lines in the mesosphere. Several calculations considering the high resolution feature have been added to the algorithm. High resolution solar spectra are now used for the calculation of the emissivity. In case of solar absorption lines, e.g., Mg at 285.2 nm, this leads to larger densities, because in the low resolution spectrum the true solar irradiance spectrum is smoothed and the solar irradiance, and with it the emissivity of the spectral line, is overestimated. For solar emission lines, like Mg^+ at 279.6 nm and 280.4 nm, using high resolution solar

spectra leads to lower densities. Several spectral shifts caused by Doppler effect, isotope effects, and hyper-fine-splitting are considered now. The isotope shifts have the strongest impact on Mg and Mg^+, while Doppler and hyper-fine-splittings mainly influence Na.

The retrieval algorithm uses all limb measurements of an orbit simultaneously to retrieve densities on a 2D latitudinal and vertical grid. For averaged data the geolocation data of a reference orbit, formed from the measurements used for the average is used. The retrieval algorithm can be separated into two steps. In the first step the slant column emission signal is determined from the spectra of each measurement. To do this, the emission spectrum has to be separated from the background signal. The second step is the calculation of the densities in each grid element from the slant column emissions. This is done by inverting a forward model for the emission signal. The used forward model is a single scattering radiative transfer model. The reason why a single scattering model can be used for the Mg^+ and Mg lines will be explained in the next section. For other metal lines like Na at 590 nm a multiple scattering and a surface reflection contribution has to be considered and the necessary extensions of the algorithm are described in Sect. 6.5.

Assuming single scattering of light from the Sun into the line of sight (LOS) of the instrument (emission path s_e), for a single measurement (one limb spectrum at one altitude) the forward model for the radiance can be expressed as follows:

$$4\pi I = \int\limits_{\mathrm{LOS}} \gamma n(s_\mathrm{e}) f \left(\int n(s_\mathrm{a}) \mathrm{d}s_\mathrm{a} \right) \mathrm{d}s_\mathrm{e} \tag{6.1}$$

with emissivity γ, density n and an absorption part – along the LOS and the line from Sun (LFS) (s_a stands for both absorption paths) – f. Incoming photons scatter in the form of resonance fluorescence, a certain sequence of absorption and re-emission, with the investigated particle species. The resulting radiation is proportional to the density n, with the emissivity γ being the constant of proportionality. The photons are partly absorbed before the scattering process along the LFS, or after the scattering process along the LOS. The measured radiance is the integrated radiation coming from these processes along the LOS. In the first step of the retrieval the left hand side of Eq. (6.1) is determined, while in the second step the right hand side of Eq. (6.1) is inverted to obtain the density n. More detailed information on γ and f is provided in Sects. 6.2.3 and 6.3.4, respectively.

6.2 Determination of slant column densities

6.2.1 Determination of the background signal in the UV

As a first step in the determination of the pure emission signal, it has to be separated from the background signal. The background signal is mainly formed by Rayleigh and Raman scattering of solar electromagnetic radiation by air molecules. Above 70 km the air is thin enough and scattering path lengths are long enough that only single scattering has to be considered. Below 300 nm, ozone in the stratosphere absorbs the main part of the incoming solar electromagnetic radiation, so that no multiple scattering or surface reflection contributions have to be considered.

Figure 6.1 shows recent measurement results of ozone absorption cross section for typical atmospheric temperatures by Gorshelev et al. (2013) and Serdyuchenko et al. (2013). The ozone absorption cross section between 200 nm and 800 nm wavelength

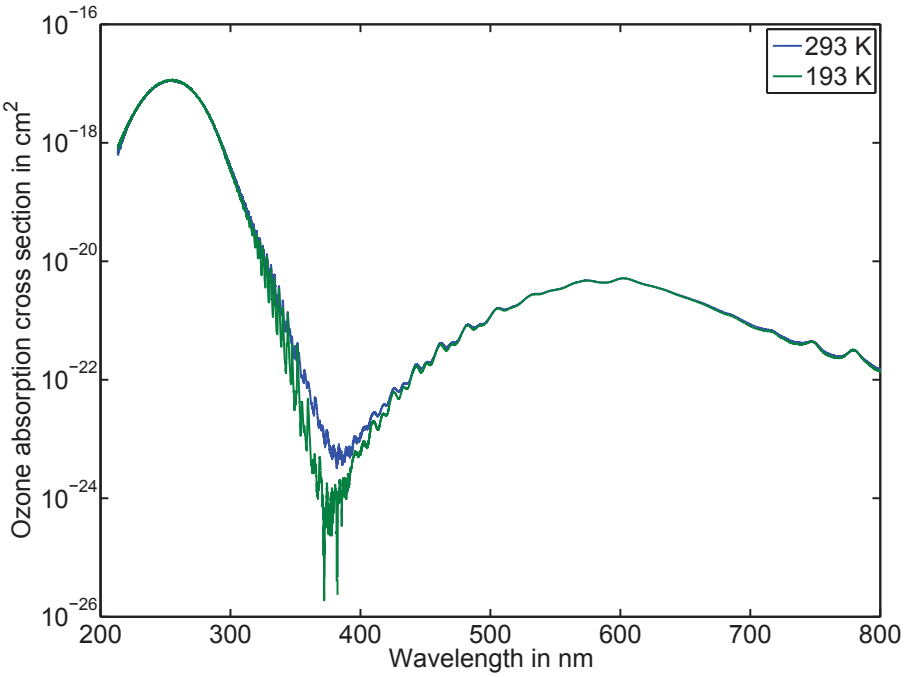

Fig. 6.1: O_3 cross section for the 200 nm to 800 nm range from Gorshelev et al. (2013); Serdyuchenko et al. (2013). For a typical ozone profile at mid latitudes and a solar zenith angle of SZA = 0°, 96 % of the light, incoming from the top of the atmosphere, at 600 nm and at 330 nm is left at 10 km altitude, while at 300 nm only 5 % is left. For the calculated remaining light at 285 nm and 280 nm values of the order of 10^{-9} and 10^{-14}, respectively, are obtained. For this study, only the ozone absorption in the stratosphere is relevant: for a typical SCIAMACHY MLT limb state even at 280 nm the remaining signal is still more than 99 % of its original value after passing the whole line of sight for tangent altitudes above 68 km. Therefore, ozone absorption in the mesosphere can be and is neglected in the radiative transfer model.

has three distinguished strong absorption bands. The Hartley band with a maximum absorption cross section at 255 nm shows the strongest absorption. The other bands are the Huggins band between 320 and 360 nm, which overlaps the edge of the Hartley band and the Chappuis band with a maximum absorption cross section at 600 nm. In the mesosphere the ozone density is small enough, so that less than 1% of the incoming solar electromagnetic radiation at 280 nm is attenuated along a line of sight with a tangent altitude of 70 km. However, for wavelengths closer to 255 nm the ozone absorption is important. Below 70 km altitude the Ring effect, described in the next section, is a larger error source than ozone absorption.

In the stratosphere, the ozone layer is thick enough that far less than 1% of the UV radiation below 295 nm reaches the ground. Therefore, multiple scattering

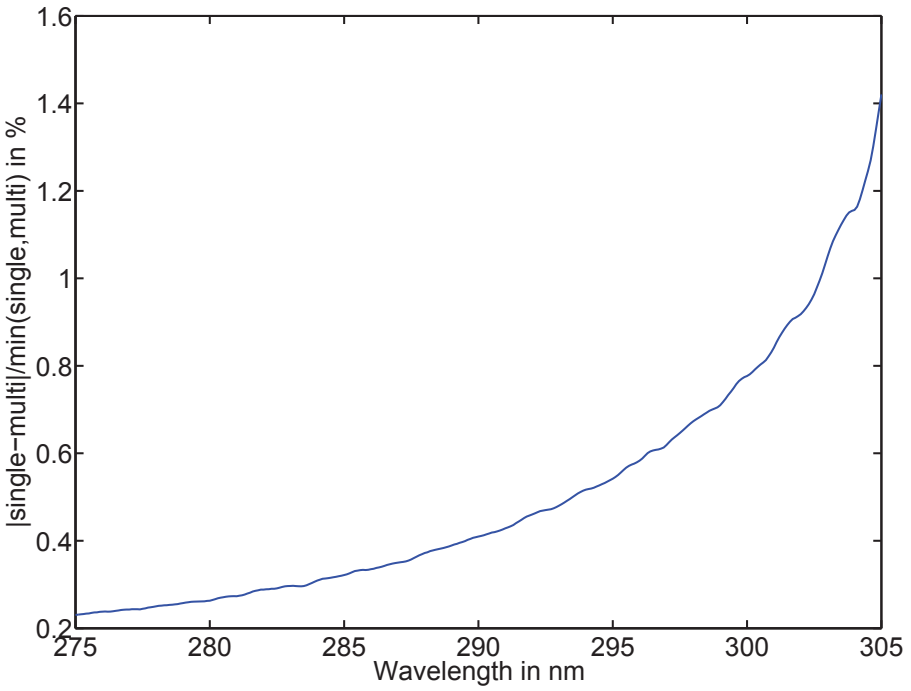

Fig. 6.2: Relative difference $\frac{|\text{multiple scattering}-\text{single scattering}|}{\min(\text{multiple scat., single scat.})}$ of simulated spectra for multiple scattering for the wavelength region of the Mg/Mg^+ lines. The simulations were performed with SCIATRAN for an equatorial scenario. This and other scenarios with, e.g., a total ozone column density of 150 Dobson units (1 Dobson unit= $2.687 \times 10^{16}\,\mathrm{cm}^{-2}$) show only minor differences of less than 1%. Multiple scattering thus can be neglected for the retrieval of Mg/Mg^+ densities in the MLT.

from lower altitudes is negligible at 280 and 285 nm. Figure 6.2 shows the relative difference between singly scattered and multiply scattered electromagnetic radiation for a typical limb scan. This was simulated by Vladimir Rozanov with the more sophisticated radiative transfer model SCIATRAN (see, e.g., Rozanov et al., 2005) and shows less than 1% differences between single and multiple scattering at these wavelengths.

The cross section for single Rayleigh scattering for a refractive index of $n = 1$ is given by the following equation:

$$\sigma(\lambda) = \frac{128\pi^5\alpha^2}{3} \cdot \frac{1}{\lambda^4} \tag{6.2}$$

with the polarizability volume α and the wavelength λ. The single Rayleigh scattered spectrum, therefore, is an attenuated copy of the solar spectrum with an additional wavelength modulation of λ^{-4}. For wavelength windows with a much smaller width than the mean wavelength, e.g., a 2 nm window at 280 nm, the ratio of Rayleigh scattered electromagnetic radiation and incoming solar radiation can be well approximated by a linear function.

Figure 6.3 shows the measured ratio of limb radiance and solar irradiance between 275 and 290 nm. The strong peaks in this ratio are clearly identified as

emission signals or inelastic Raman scattering.

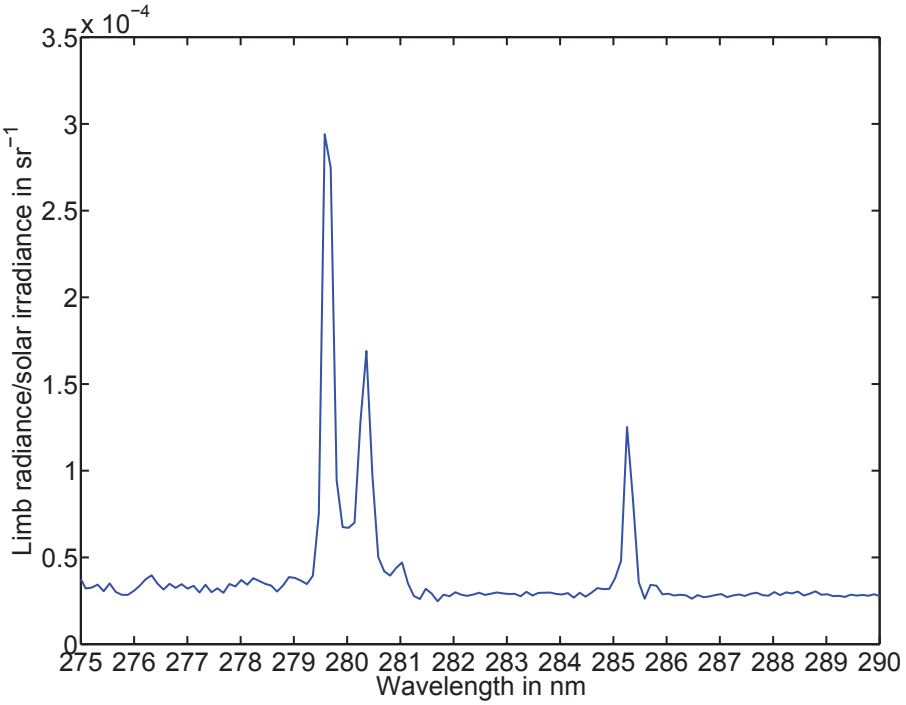

Fig. 6.3: Limb radiance divided by solar irradiance (average 2008–2012, equator, tangent altitude 89.3 km). Division of both spectra leads to a smooth signal, beside emission lines that are clearly identified. The background signal is separated from the emission lines through a linear fit. Note, that beside the strong Mg^+ at 279.6 and 280.4 nm and the strong Mg line at 285.2 nm very weak NO lines can be observed at the longer wavelength edge of the Mg^+ lines at 281 nm. However, this influence is negligibly small.

To exclude influence of the inelastic rotational Raman scattering a Ring effect correction, which is described in the next section, is performed before forming the ratio of limb radiance and solar irradiance. The Rayleigh background can be fitted as a straight line of this ratio for a small window around the emission lines, but excluding the emission lines. After subtraction of the Rayleigh background, the ratio is multiplied with the solar irradiance again to obtain the pure emission spectrum. The result is fitted with the slit function of the instrument to obtain the slant column emission (SCE) signal. These steps are illustrated in Fig. 6.4. For a good fit of the slant column emission signal it is advantageous to have a good signal to noise ratio, low influence of Ring effect and strong emission signals compared to other background signals and no overlap with other nearby emission lines. Figure 6.3 shows a small emission line at 281 nm from the overlapping NO $\gamma(1,6)$ band. However, this emission is negligibly small.

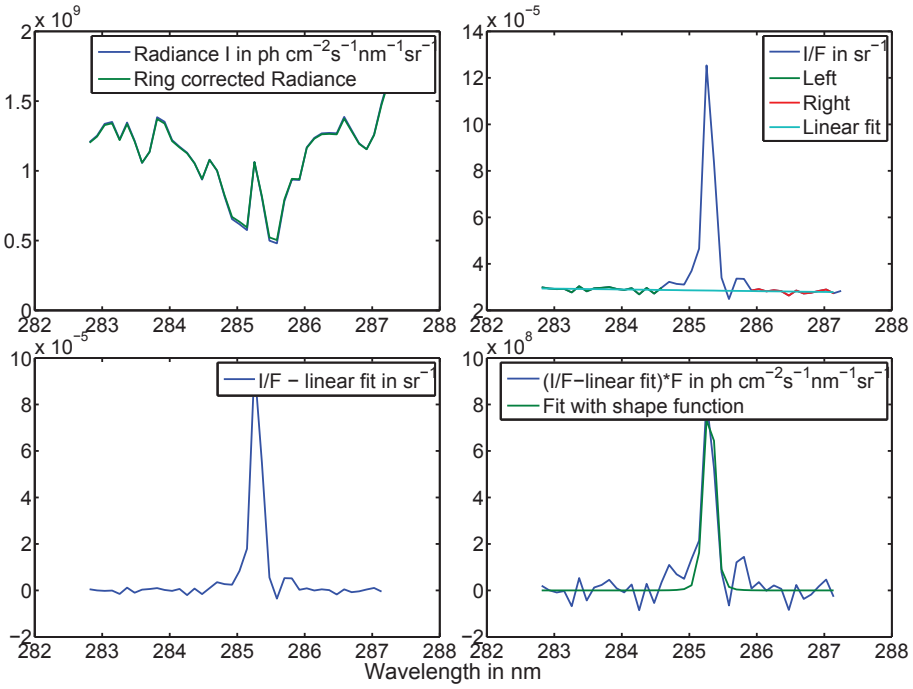

Fig. 6.4: Slant column emission (SCE) determination. First, the Ring effect correction is applied to the limb radiance I (top left panel). The limb radiance is divided by the solar irradiance F and the wavelength region left and right to the line are used for a linear fit of the background radiation (top right panel). The linear fit is subtracted from the ratio I/F (bottom left panel) and the spectrum is again multiplied with the solar irradiance F (bottom right panel). Finally, the area of the emission line is fitted with the slit function of the instrument.

6.2.2 Ring effect correction

The metal atoms and ions, which cause the emissions in the Earth's atmosphere, are also present in the solar atmosphere. The presence of the metals leads to atmospheric emission and absorption features in the atmosphere of both the Earth and the Sun. The resulting solar spectrum features introduce some challenges into the investigation of metals in the Earth's mesosphere. The absorption lines in the solar spectrum, which partly also show emission lines as a finer structure in the absorption line, are also called Fraunhofer lines.

Three important issues arise for the retrieval of metal densities from emission lines of the same wavelengths as Fraunhofer lines. First, a solar spectrum from the same day as the limb observations has to be used, to separate solar and mesospheric effects. Second, the high resolution structure of the solar Fraunhofer lines has to be known, because instruments with too low resolution smooth the solar spectrum leading to too high values for the solar irradiance in absorption lines and too small values of the solar irradiance in emission lines. The third issue is inelastic scattering by air molecules in the Earth's atmosphere. In contrast to atoms, molecules have additional degrees of freedom for rotational and vibrational energy levels. The incoming electromagnetic radiation may stimulate transitions between these energy levels, leading to the emission of electromagnetic radiation with lower or higher energy compared to the incoming electromagnetic radiation. Scattering is stronger, where the incoming electromagnetic radiation field is stronger. Therefore, there is effectively more scattering from a wavelength region with high intensity to nearby wavelength regions with lower intensity than vice versa. This leads to a redistribution of radiation in the vicinity of Fraunhofer lines. For the typical width of Fraunhofer lines only rotational Raman scattering is relevant for the redistribution.

Roughly 4% of the total scattered light in the wavelength region of the Mg and Mg^+ line comes from rotational Raman scattering, while the majority of the scattered light is Rayleigh scattered. The inelastic scattering part effectively leads to a filling-in effect of Fraunhofer lines in the Earth's atmosphere, where light from the higher intensity edges is shifted to the low intensity center of the line. This filling-in effect is known as the Ring effect found by Grainger and Ring (1962). The applied Ring-effect correction is based on the assumption, that the Raman scattering part of the total scattered light is much smaller than the Rayleigh scattered part, and the Ring-effect can be linearized. This means that a second application of the Ring effect to the spectrum leads to the same changes as the first application. Under this assumption the Ring effect is artificially applied to the Ring effect affected spectrum I_1 to yield I_2 and the difference $I_1 - I_2$ needed for a correction of the Ring effect and to retrieve the non Ring effect affected spectrum I_0:

$$I_2(\lambda) = I_1(\lambda) + \underbrace{\Delta I_{\mathrm{R}2}(\lambda)}_{I_2(\lambda)-I_1(\lambda)} \tag{6.3}$$

$$I_1(\lambda) = I_0(\lambda) + \underbrace{\Delta I_{\mathrm{R}1}(\lambda)}_{I_1(\lambda)-I_0(\lambda)} . \tag{6.4}$$

Under the assumption $\Delta I_{\mathrm{R}}(\lambda) = \Delta I_{\mathrm{R}2}(\lambda) = \Delta I_{\mathrm{R}1}(\lambda)$ it follows that

$$I_0(\lambda) = I_1(\lambda) - \Delta I_{\mathrm{R}} = 2I_1(\lambda) - I_2(\lambda). \tag{6.5}$$

To test whether this linearization approach works the Ring effect operator, which is described in the next paragraph, is applied to I_0, which results in I_1, if the correction works correctly. Figure 6.5 shows I_0, I_1, I_2 and I_1 recalculated from I_0 for a sample spectrum and both I_1 in Fig. 6.5 are nearly identical.

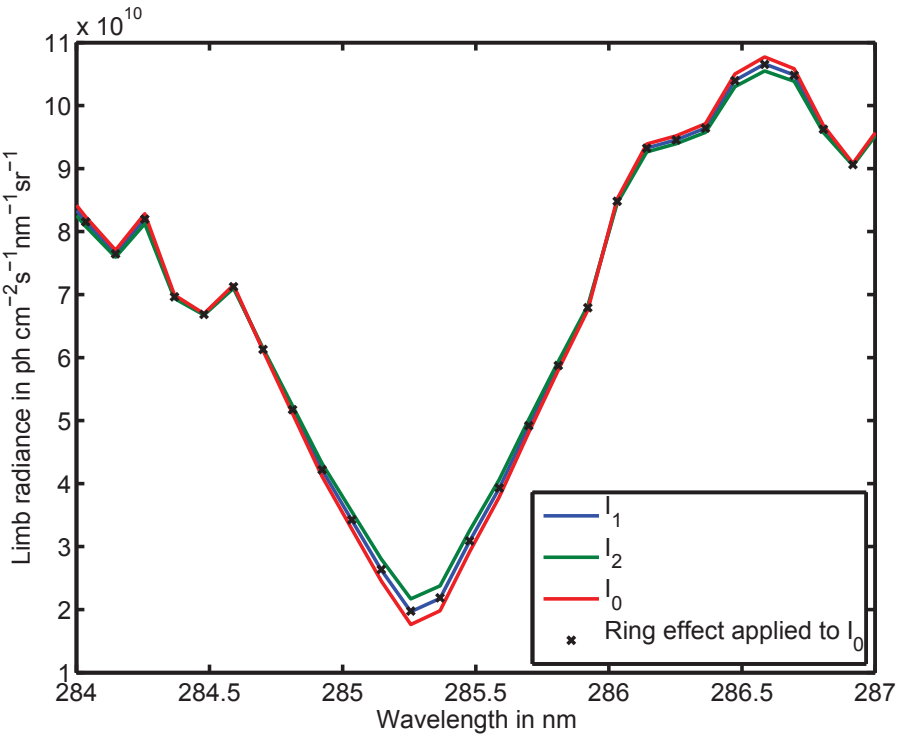

Fig. 6.5: Ring effect correction for the Mg 285.2 nm line for a limb measurement with a tangent altitude of 53.5 km. The Ring effect is forward simulated using the measured spectrum $I_1(\lambda)$ to obtain $I_2(\lambda)$. The filling-in effect is assumed to be small and differences between the first and the second application are, therefore, also small. Under this assumption the differences are nearly the same $I_0(\lambda) - I_1(\lambda) \approx I_1(\lambda) - I_2(\lambda)$. And by adding $I_1(\lambda)$ the corrected spectrum $I_0(\lambda)$ is obtained as $I_0(\lambda) = 2I_1(\lambda) - I_2(\lambda)$.

A 3.3 nm wide boxcar function with a central peak at $\Delta\lambda = 0$ shown in Fig. 6.6 is used as the Ring effect operator. The central peak represents the 96% Rayleigh scattering, while the boxcar function represents the Raman scattered part. Note, that actually the Raman spectra of N_2 and O_2 show discrete narrow lines. However, when smoothing a Raman spectrum (e.g., from Penney et al., 1974) with the SCIAMACHY resolution of 0.22 nm it is very similar to a boxcar function. The boxcar function, therefore, is adequate for a rough estimation of the Ring effect as long as the correction is not so strong, that a more exact method is necessary.

The fraction of Raman scattering is estimated to be between 3 and 6%. To test whether the Ring effect correction works correctly, nearby Fraunhofer lines of metals with very low emission signals in the mesosphere, e.g., the Si line at 288 nm, are used. It is assumed, that these apparent "emission lines" in the limb radiance

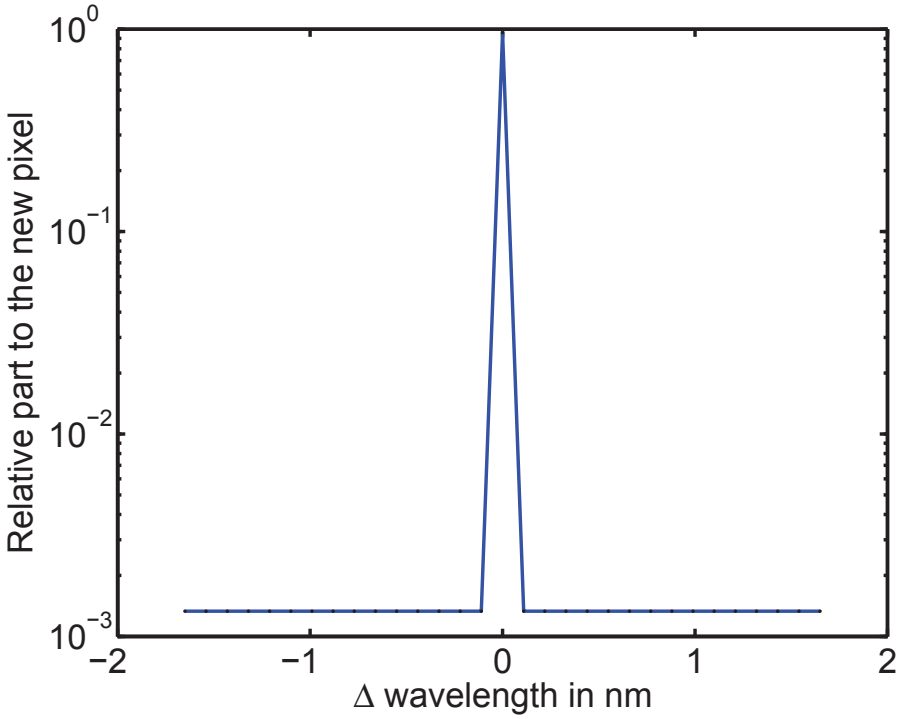

Fig. 6.6: Ring effect smoothing function. 96 % of the background light is Rayleigh scattered. The Raman scattered part is approximated with a boxcar for the remaining 4 %.

to solar irradiance ratio result only from Raman scattering and should disappear if the correct percentage of Raman scattering is chosen. If the line is still an apparent emission line after the correction, the correction is too weak, while if the line is an apparent absorption line after the correction the correction is too strong.

The Ring effect is especially strong for Mg compared to Mg^+ as the Mg line is a single Fraunhofer line and can be filled in from both the longer and shorter wavelength edge, while the Ring effect is much weaker for the Mg^+ double line, where the line is wide enough that filling-in comes only from one edge of the spectrum. Ratio spectra of Ring effect corrected limb radiances and solar irradiance for different percentages of Raman scattering and different altitudes are shown in Figs. 6.7 and 6.8. The Ring effect correction influence on the vertical slant column emission profile is shown in Fig. 6.9.

For the Mg 285.2 nm line the Ring effect influence is nearly negligibly small above 90 km. Below 90 km the Ring effect influence steeply increases and the ratio of contribution to the spectral peak from the Ring effect and the actual emission becomes larger than 1 between 80 and 70 km. Below 70 km the Ring effect contribution to the peak in the spectrum dominates the emission signal. Therefore, only measurements with tangent altitudes above 70 km are used for the Mg and Mg^+ retrieval.

The situation is much better for the Mg^+ double lines, where the Ring effect can nearly be neglected above 70 km. Furthermore, the Ring effect correction algorithm

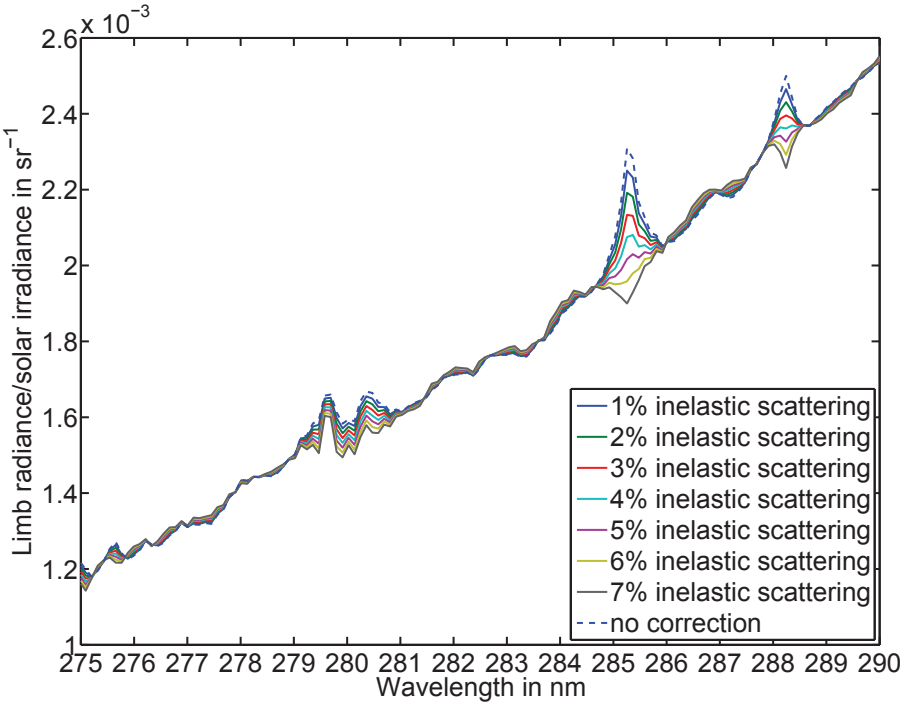

Fig. 6.7: Result of the Ring effect correction for a tangent altitude of 53.5 km for the ratio I/F. Shown are the uncorrected spectrum and the corrected spectra with different percentages of Raman scattering in the applied method. The emission signals become smaller with this correction and the Mg line at 285.2 nm, as well as the Si line at 288 nm, which cannot be observed in the region between 90 to 100 km and originates probably purely from Ring effect, nearly vanish for the Ring correction with 4 % of inelastic scattering in the background signal.

produces a small systematic error in the peak altitude region of Mg^+. This is, because the signal to background ratio is high there. For the case of a pure emission spectrum the Ring effect smoothing reduces the peak value of the emission by 4%. Therefore, it is better to not to do the Ring effect correction at all for the Mg^+ lines, but also neglect altitudes below 70 km (Not doing the Ring effect correction above a certain altitude or including the emission feature into the consideration of the Ring effect correction was another option here).

The Ring effect correction was also compared to the more sophisticated and exact calculations of SCIATRAN (see, e.g., Rozanov et al., 2014), showing very similar results, so that the use of the simpler approach is justified above 70 km. SCIATRAN allows to forward model radiative transfer with and without inelastic scattering. However, when comparing these to measured spectra, there are small differences in this wavelength region. These differences may originate partly from simplifications in the models on the one hand side and spectrometer calibration issues on the other hand side. Whatever the reason for these differences is, it makes it hard to apply the SCIATRAN corrections directly to the measured spectra without making significant systematic errors. Therefore, the simpler approach of the Ring effect correction is not just simpler, but also not worse than using the more

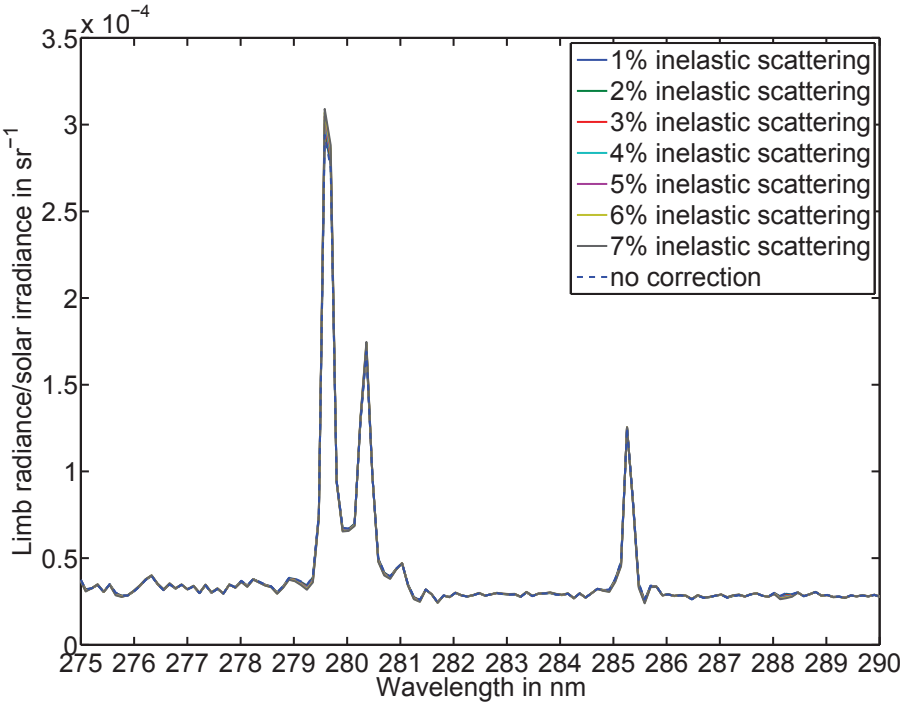

Fig. 6.8: Result of the Ring effect correction for a tangent altitude of 90 km for the ratio I/F. The background signal is low at this altitude and the Ring effect correction is only very small.

sophisticated method, at this point.

6.2.3 γ factors

There are several constant factors on the right-hand side of Eq. (6.1), which is repeated here:

$$4\pi I = \int \gamma n(s_\mathrm{e}) f \left(\int n(s_\mathrm{a}) \mathrm{d}s_\mathrm{a} \right) \mathrm{d}s_\mathrm{e}, \tag{6.6}$$

with γ being the emissivity,

$$\underbrace{\gamma}_{\text{photons}\,\mathrm{s}^{-1}} = \underbrace{P(\theta)}_{\text{Phase function}} \times \underbrace{\pi F_{\lambda_0}}_{\text{photons}\,\mathrm{s}^{-1}\,\mathrm{cm}^{-2}\,\mathrm{nm}^{-1}} \times \underbrace{\frac{1}{4\pi\epsilon_0}\frac{\pi \mathrm{e}^2}{mc^2} f_{ij} \lambda_{ij}^2}_{\text{integrated abs. cross sect. in nm cm}^2} \times \underbrace{\frac{A_{ji}}{\sum_k A_{jk}}}_{\text{rel. Einstein coeff.}} . \tag{6.7}$$

The physical process for this emission is resonance fluorescence (see, e.g., Weisskopf, 1931), i.e., absorption and re-emission of electromagnetic radiation by the atom/ion. The incoming radiation field πF_{λ_0} excites the atom/ion from the initial state i to the excited state j. Therefore, the electromagnetic radiation first must be absorbed.

Only the wavelength integrated emission over the whole line is detected, because the SCIAMACHY instrument has a roughly 1000 times worse resolution than needed to resolve the pm wide structure of the emission line. Strong metal emission

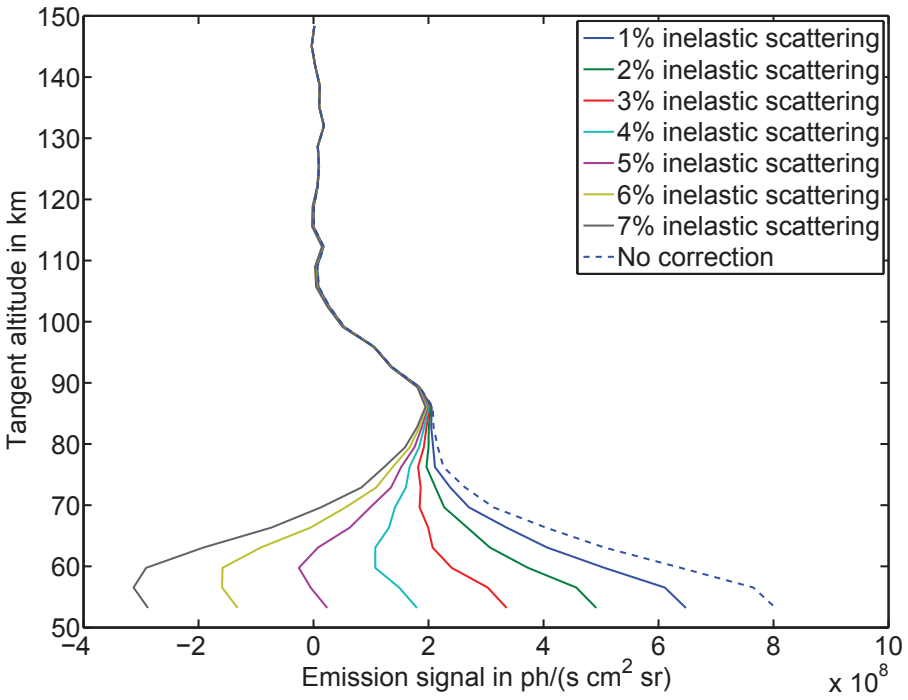

Fig. 6.9: Vertical slant column emission (SCE) profile of Mg with different percentages of inelastic scattering in the Ring effect correction. The difference between the corrected and uncorrected profile increases with decreasing tangent altitude below 90 km. Note, that the SCE of Mg at 53.5 km is still in the order of 50 % of the SCE at 90 km for typical recalculations of the forward model from retrieval results.

lines typically have Einstein coefficients for spontaneous emission on the order of 10^8 s^{-1}, which results in immediate re-emission.

There may be other non resonant emission from the excited state to lower states, which depopulate the excited state. This effect is considered in the relative Einstein coefficient. Not considered are population increases of the excited state through other transitions, which are however considered to be negligibly small. As for Mg the considered transitions are the lowest allowed transitions from excited states to the ground state, the relative Einstein coefficient is very close to 1.

The direction of the re-emitted photon is considered in the phase function, which is normalized to 4π. The Phase function is given by

$$P(\theta) = \frac{3}{4}E_1(\cos^2(\theta) + 1) + E_2. \tag{6.8}$$

The factors E_1 and E_2 depend on the total angular momentum j of the lower state and the change in total angular momentum Δj and are taken from Chandrasekhar (1960) (see Table 6.4). The first part of the phase function is similar to the Phase function of Rayleigh scattering, while the second part is isotropic. Depending on E_1 and E_2, the phase function is a linear combination of both individual phase functions. The factors E_1 and E_2 for the different metal emission lines are shown in Table 6.5. An interesting feature of the Mg^+ doublet line is, that both lines

Table 6.4: E_1 and E_2 depend on the angular momentum j of the lower state and the change of angular momentum Δj (from Chandrasekhar, 1960, table II, page 52).

Δj	E_1	E_2
$+1$	$\frac{(2j+5)(j+2)}{10(j+1)(2j+1)}$	$\frac{3j(6j+7)}{10(j+1)(2j+1)}$
0	$\frac{(2j-1)(2j+3)}{10j(j+1)}$	$\frac{3(2j^2+2j+1)}{10j(j+1)}$
-1	$\frac{(2j-3)(j-1)}{10j(2j+1)}$	$\frac{3(6j^2+5j-1)}{10j(2j+1)}$

have a different E_1 and E_2, which not only has implications on the phase function, but also on the polarization of the emitted light. Since grating spectrometers like SCIAMACHY are sensitive to polarization, this must be considered in the calibration of the spectra. The comparison of the densities retrieved from both Mg^+ lines can be used as an indicator, how good this polarization correction works, which is discussed in Sect. 6.5.6.

Table 6.5: E_1 and E_2 for different metal lines.

species	wavelength in nm	E_1	E_2
Mg	285.2	1	0
Mg^+	279.6	0.5	0.5
Mg^+	280.4	0	1
Na	589.2 (D2)	0.5	0.5
Na	589.8 (D1)	0	1

The wavelength-integrated cross section depends only on the transition wavelength λ_{ij} and the absorption oscillator strength f_{ij} of the transition. Both values are provided, e.g., in the NIST atomic spectra database (Kramida et al., 2012), which is used for the calculations. The integrated cross section has to be spectrally distributed over the shape of the line for mesospheric conditions, i.e., a Doppler-broadened Gaussian line shape for each individual line with a FWHM of $\sqrt{\frac{8\,RT\,\ln(2)}{Mc^2}}\lambda_0$ – with the gas constant R, absolute temperature T, molar mass of the atom M and speed of light c – which is roughly 0.6 pm for all three lines at typical mesospheric temperatures.

Furthermore, effects that cause non-negligible splitting and spectral shifts of the lines have to be taken into account. Mg has three stable isotopes $^{24}_{12}$Mg, $^{25}_{12}$Mg and $^{26}_{12}$Mg with natural relative abundances of 78.99, 10.00 and 11.01 %. Only $^{25}_{12}$Mg shows hyper-fine-structure-splitting, which is, however, negligibly small and so can be ignored. More important for Mg and Mg^+ are isotope shifts that are comparable to the FWHM. For the Mg line at 285.2 nm, the shift from $^{24}_{12}$Mg to $^{25}_{12}$Mg is 0.195 pm and for $^{24}_{12}$Mg to $^{26}_{12}$Mg it is 0.377 pm (taken from Le Boiteux et al., 1988). For the Mg^+ line at 279.6 nm, the shift from $^{24}_{12}Mg^+$ to $^{25}_{12}Mg^+$ is 0.423 pm and for $^{24}_{12}Mg^+$ to $^{26}_{12}Mg^+$ it is 0.805 pm (taken from Batteiger et al., 2009). For the Mg^+ line at

280.4 nm, the shift from $^{24}_{12}Mg^+$ to $^{25}_{12}Mg^+$ is 0.425 pm and for $^{24}_{12}Mg^+$ to $^{26}_{12}Mg^+$ it is 0.809 pm (taken from Batteiger et al., 2009).

Not just the cross section, but also the solar irradiance πF_{λ_0} has to be known with high resolution, as actually the integrated product of both is needed. It is sufficient to only know the integrated values, as long as only emission is considered. However, if the metal density is large enough, absorption of other metals from the same species along the line of sight of the instrument and the line from Sun is not negligibly small anymore. This can be interpreted as a reduced emissivity $f\gamma$ with the attenuation factor f. A high resolution treatment has to be carried out to properly calculate f, which is discussed in Sect. 6.3.4 (see Fig. 6.19).

The metals present in the Earth's upper atmosphere are also present in the solar atmosphere. Mg even has the same isotopic ratios in the solar atmosphere as in the terrestrial atmosphere (see, e.g., Boyer et al., 1971). The occurrence of metal species in the solar atmosphere leads to the formation of Fraunhofer lines in the solar spectrum. As the majority of Fraunhofer lines are narrow, having line widths in the pm range, a high resolution solar spectrum is required for a proper treatment of radiative transfer effects relevant for this study. SCIAMACHY's spectral resolution of about 0.22 nm in the 280–285 nm spectral range is too poor to resolve the individual atomic lines. Ground-based instruments can have the required spectral resolution but are incapable of observing the Mg/Mg^+ lines, because the stratosphere is optically thick below about 300 nm. Therefore, the balloon-borne measurements of Anderson and Hall (1989) and Hall and Anderson (1991) carried out in 1978 and displayed in Fig. 6.10 are employed. The spectral resolution of these solar irradiance measurements is 0.01 nm. This spectral resolution is sufficient to resolve the Mg and the Mg^+ lines. However, the Mg^+ emission lines show an inner minimum that is not fully resolved. The solar Fraunhofer lines are also much broader than expected from pure Doppler broadening and this broadening is assumed to be mainly pressure broadening. For later considerations in Sect. 6.3.4 a spectrally constant solar irradiance is assumed in the considered wavelength range in the mesosphere, because the Doppler broadened cross section profiles have only a width of several pm, which is 2 orders of magnitude smaller than the width of the solar lines.

Beside the high resolution of the spectrum, temporal variability of the spectrum has to be considered. The MgII index –i.e., the ratio between the irradiance of the chromospheric emission lines near the center of the MgII Fraunhofer line and the wings of the MgII Fraunhofer line – is one of the most commonly used solar proxies (see, e.g., Skupin et al., 2004; Snow et al., 2005). There is a variability of 20 % associated with the 11 year cycle and similar variability of up to 10 % associated with the 27 day cycle. Note, that in spectroscopy Mg is also referred as MgI and Mg^+ is referred as MgII.

The solar cycle variation is considered in the following way. First, the high resolution spectrum is scaled to the SCIAMACHY spectrum with a constant factor, to consider different degradation effects of the instrument. This works best when a smoothed version of the high resolution spectrum is used (see, e.g., Fig. 6.11) to find the best fitting factor first. As the next step the actual effect of the variability is simulated. Therefore, the spectrum is scaled only in the center of the emission lines by a constant factor, which simulates high or low solar activity. Note, that

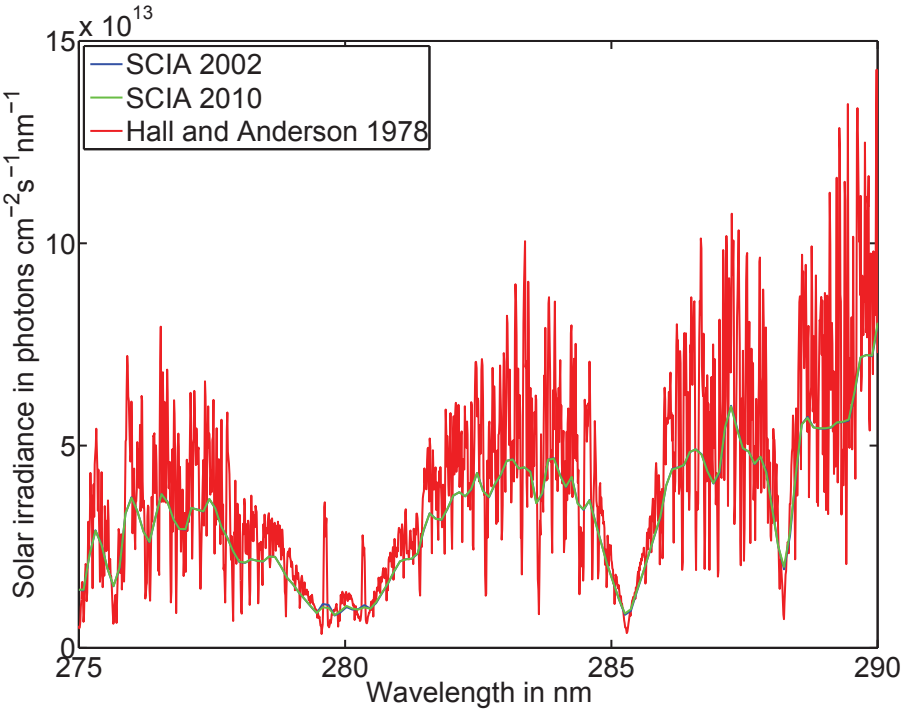

Fig. 6.10: High resolution solar spectrum from Hall and Anderson (1991) (data from 1978) in comparison to the SCIAMACHY solar irradiance spectrum during high solar activity (2002) and low solar activity (2010). There are especially large differences between the high resolution spectrum and the SCIAMACHY spectrum in the Fraunhofer lines at 279.6, 280.4, 285.2 and 288.2 nm. When smoothed and scaled, the high resolution spectrum is very similar to the SCIAMACHY spectrum (see Fig. 6.11).

for the inversion step from the smoothed lines to the highly resolved lines it is assumed, that the shape of the emission lines, which is not resolved in the smoothed spectrum, is the same as in the high resolution spectrum, which may not be true. The following instrumental effects of SCIAMACHY are applied to the scaled and peak-scaled high resolution spectrum. The high resolution spectrum is smoothed by convolving it with the SCIAMACHY channel 1 slit function, i.e., a hyperbolic function with a FWHM of 0.22 nm and a shape given by Eq. (6.9):

$$h(\lambda) = \frac{\mathrm{FWHM}^3}{4\pi\sqrt{2}} \frac{1}{(0.5\mathrm{FWHM})^4 + (\lambda - \lambda_0)^4}. \tag{6.9}$$

As a last step the spectrum is sampled just as the SCIAMACHY spectrum. This is done by interpolating the smoothed spectrum to the SCIAMACHY wavelength grid. This scaling, smoothing and sampling leads to a similarly shaped spectrum the SCIAMACHY spectrum and the resulting spectrum shown in Fig. 6.12 for Mg and in Fig. 6.13 for Mg^+. Smoothing and interpolating only slightly affect the integrated value over each single line.

Using different scaling factors for the emission line in order to model the solar variability results in different integrated values. This change can be described by

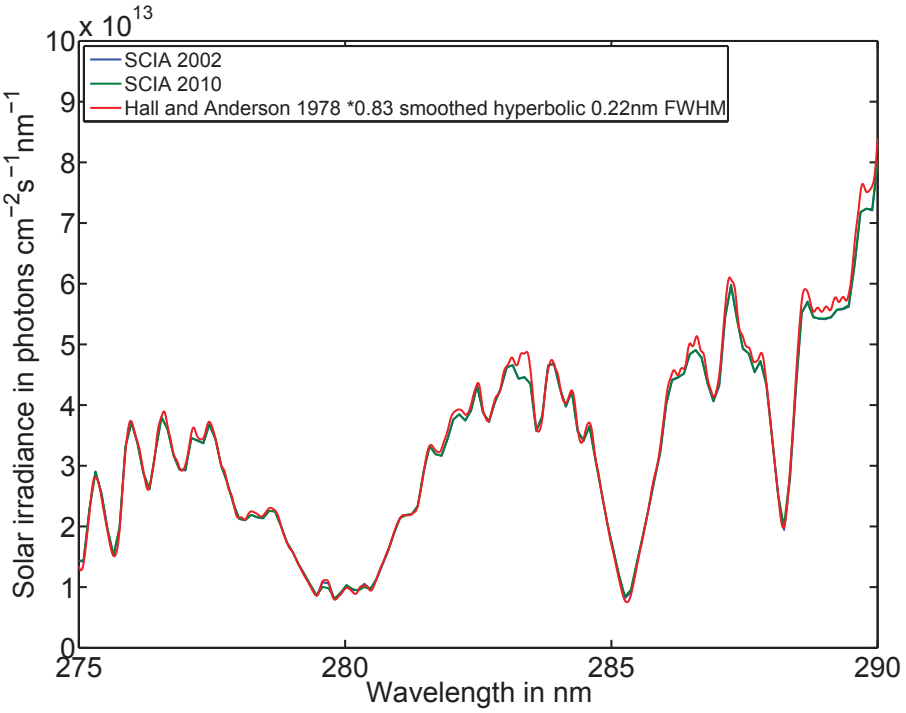

Fig. 6.11: The high resolution spectrum from Fig. 6.10 is smoothed and scaled to the SCIAMACHY spectrum by a factor of 0.83, so that the peak edges of the spectra are in agreement.

a linear function for each individual line. To consider the variations, this linear function is inverted for the integrated values of the daily measured SCIAMACHY spectrum to obtain the scaling factor that has to be applied to the emission lines. By doing this, the high resolution spectrum is used to obtain the correct line shape of the lines, while the daily SCIAMACHY spectrum is used to correct for daily variability.

To obtain the SCDs, the SCEs have to be multiplied by 4π and divided by γ. However, in the retrieval code the phase function part of γ is excluded from this division, because the phase function changes along the line of sight and is, therefore, treated at a later point.

6.3 Radiative transfer model

6.3.1 Discretization of the forward model

For the determination of number density distributions from SCDs, the forward model for calculating SCDs from number density distributions first has to be formulated, and then this forward model has to be inverted. For a single measurement, the forward model is formulated as Eq. (6.1). For numerical reasons, the forward model is discretized to a grid. All constant factors in Eq. (6.1) are combined to a factor c_1. The Phase function P is scattering angle dependent and therefore not constant along the LOS. The detected resonance fluorescence signal corresponds to the signal

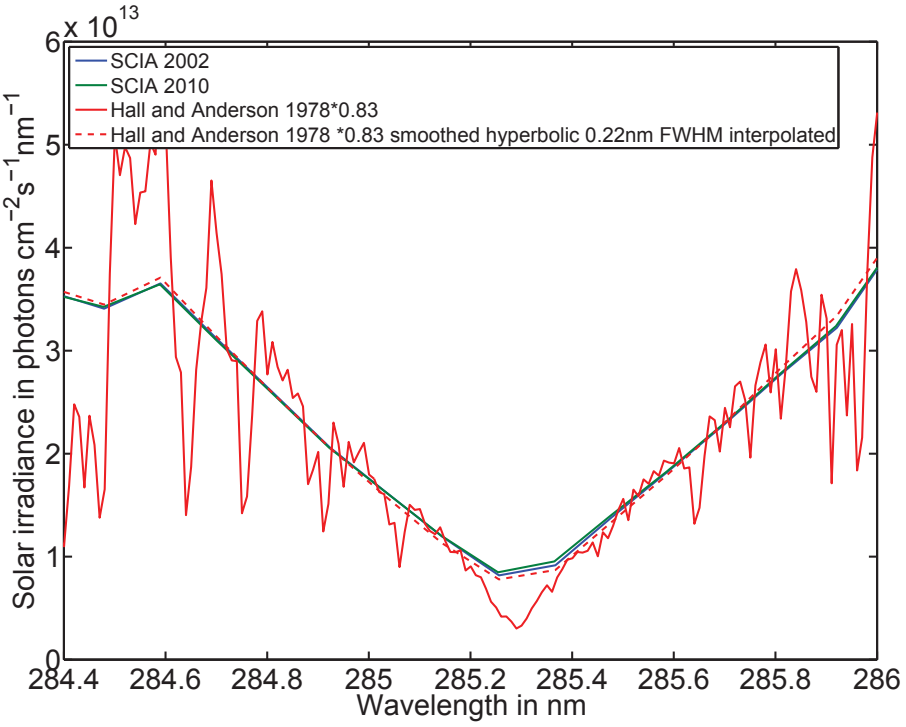

Fig. 6.12: SCIAMACHY solar spectrum in the Mg 285.2 nm line region for low (2010) and high (2002) solar activity in comparison with the scaled high resolution reference spectrum. When the high resolution spectrum is smoothed and sampled as the SCIAMACHY spectrum, there are only small differences left, mainly resulting from different phases of the solar cycle for the different spectra. Note, that a scaling factor of 0.9 instead of 0.83 is used for Mg in the retrieval because this fits the line center better.

integrated along the line of sight (LOS). Before the solar radiation is absorbed and re-emitted, it follows a line from Sun (LFS) to the center of the grid cell. Absorption occurs on both the LOS and the LFS. Self-absorption by the emitting species introduces non-linearity into the forward model. Self-absorption is assumed to be weak enough, that other multiple scattering effects, e.g., scattering into the LOS or LFS are negligible. Furthermore, absorption by other species, e.g., ozone is assumed to be negligible, as discussed in Sect. 6.2.1. Therefore, only self-absorption is considered in the forward model.

Let the forward model to be inverted be $\mathbf{K}\vec{x} = \vec{y}$, where $\vec{y}$ represents the vector of the SCDs. If there are n individual measurements, $\vec{y}$ is a $n \times 1$ matrix. $\vec{x}$ also has to be a vector and is a $m \times 1$ matrix. Then $\mathbf{K}$ has to be a $n \times m$ matrix. The solution $\vec{x}$ for the number densities on the latitude–altitude grid has to be a vector, so technically the $a \times b$ grid matrix (with a and b being the number of latitudes and altitudes, respectively) has to be stored into a $m \times 1$ vector.

Three different pathways have to be considered (see Fig. 6.14). A pathway for the emission along the whole LOS (1.). For each grid cell along the LOS, the path from the Sun to this grid cell (2.), as well as the path from the grid cell to

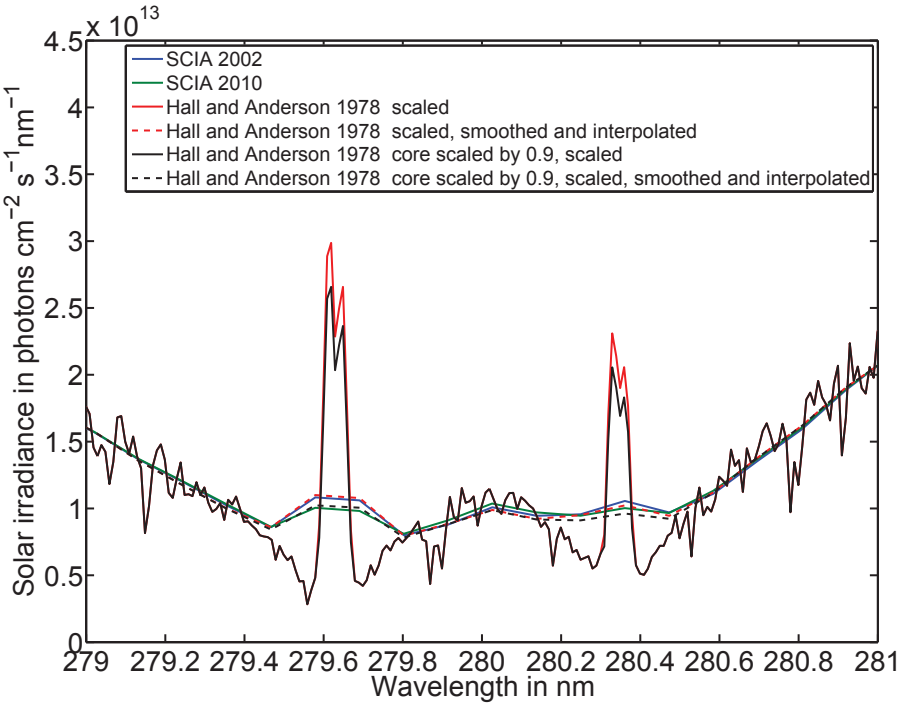

Fig. 6.13: SCIAMACHY solar spectrum in the Mg^+ double line region for low (2010) and high (2002) solar activity in comparison with the scaled high resolution reference spectrum. When the high resolution spectrum is smoothed and sampled as the SCIAMACHY spectrum, there are only small differences left between both, mainly resulting from different phases of the solar cycle for the different spectra. An adjustment for the solar phase is done by scaling the core of the emission lines. The adjustment has to be done for each line individually. Possible variations of the inner structure of the emission lines are not considered.

the satellite (3.), is needed for the absorption along both paths. For every grid cell i there is the pathway along the LOS $\Delta s_{\mathrm{LOS}_i}$ through this grid cell and the matrix (which is represented as a vector) of the pathway along the LOS, starting from the grid cell to the satellite $\Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,j}}}$, and the grid cell-specific pathways along the LFS $\Delta s_{\mathrm{LFS}_{\mathrm{gc}_{i,j}}}$. To clarify this, the difference of $\Delta s_{\mathrm{LOS}_i}$ and $\Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,j}}}$ is, that the first one includes the path length of all grid cells along the LOS from the satellite to the opposite side of the satellite, while for the second one only the grid cells j between the grid cell i and the satellite are considered and all path lengths in grid cells j along the LOS behind the grid cell i are therefore 0. The two special cases, $j = i$ and grid cell i that includes the tangent point are also considered.

The ray tracing for $\Delta s_{\mathrm{LOS}_i}$ is only done once, while the ray tracing for $\Delta s_{\mathrm{LFS}_{\mathrm{gc}_{i,j}}}$ and $\Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,j}}}$ is done once for every grid point. The discretized formula is as follows:

$$I = c_1 \sum_{i=1}^{m} P_i x_i \Delta s_{\mathrm{LOS}_i} f \left(\sum_{j=1}^{m} x_j \Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,j}}} + \sum_{j=1}^{m} x_j \Delta s_{\mathrm{LFS}_{\mathrm{gc}_{i,j}}} \right). \tag{6.10}$$

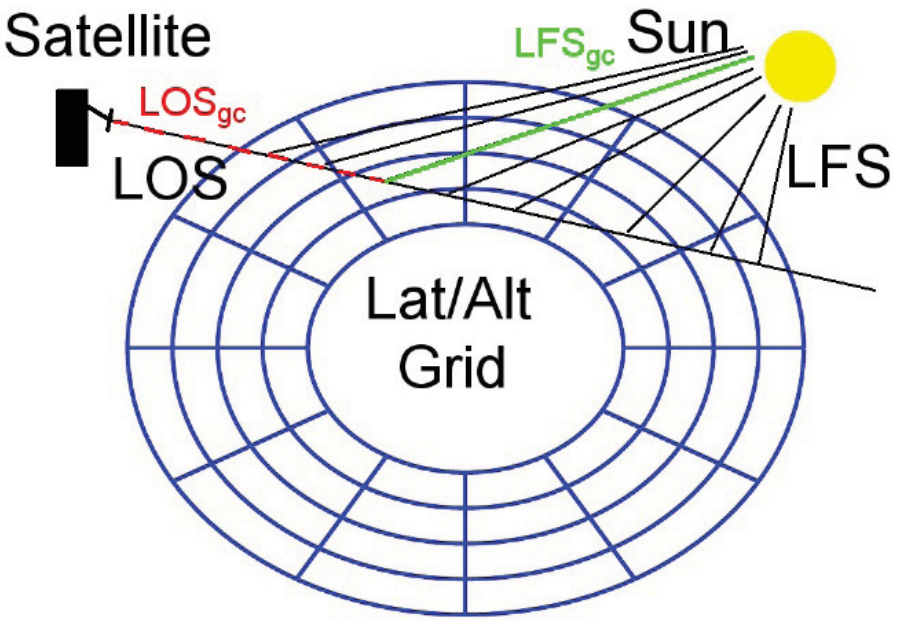

Fig. 6.14: Scheme of the altitude (radius) and latitude (sectors) grid and the line-of-sight (LOS) and line from Sun (LFS). Each grid element has a different LFS. Note, that this figure shows only a specific 2D projection of the atmosphere along a constant meridian so that latitude sectors are equally sized. Each grid cell on the LOS has its own matrices for the LFS from the Sun to the grid cell and the LOS from the grid cell to the satellite, which are needed for absorption calculations.

The argument of f is summarized to

$$g_i(x) = \sum_{j=1}^{m} x_j \Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,j}}} + \sum_{j=1}^{m} x_j \Delta s_{\mathrm{LFS}_{\mathrm{gc}_{i,j}}}. \tag{6.11}$$

The forward model for the measured radiance is expressed through a matrix equation $\mathbf{K}\vec{x} = \vec{y}$, with the state vector $\vec{x}$ containing the metal species densities to be retrieved for all latitudes and altitudes of the model atmosphere. The measurement vector $\vec{y}$ contains the measured radiances for all limb scans of one orbit, at the different tangent altitudes and latitudes. $\mathbf{K}$ is the matrix representing the forward model. For a linear problem $\mathbf{K}$ is independent of $\vec{x}$ and, therefore, $\mathbf{K}$ is identical to the Jacobian $\mathbf{J}$ ($\mathbf{J} = \mathbf{K}$) of $\mathbf{K}\vec{x}$. However, only the emission part is linear, while the absorption part is non-linear. To find the least squares of the residuum between model and measurements, the equation system $\mathbf{J}^T\mathbf{J}\vec{x} = \mathbf{J}^T\vec{y}$ has to be solved and, therefore, $\mathbf{J}$ is required. For a single wavelength the absorption part is exponential. However, several different wavelengths are involved, so that the exact calculation is a mixture of different exponentials. Therefore, the exact function f should be used, which depends on the optical depth along the LOS and LFS. The function f and its first derivative f' with respect to optical depth can be derived numerically (shown

in Sect. 6.3.4). Both are quite smooth and monotonic. To get **J** the derivative of Eq. (6.10) has to be formed.

6.3.2 Formation of the Jacobian J

The Jacobian **J** is built element-wise. To calculate one element the derivative of Eq. (6.10) has to be formed.

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 \frac{\mathrm{d}}{\mathrm{d}x_k} \sum_{i=1}^{m} P_i x_i \Delta s_{\mathrm{LOS}_i} f\big(g_i(x)\big) \tag{6.12}$$

Here the sum rule is applied to pull the derivation into the sum:

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 \sum_{i=1}^{m} \frac{\mathrm{d}}{\mathrm{d}x_k} \left(P_i x_i \Delta s_{\mathrm{LOS}_i} f\big(g_i(x)\big) \right). \tag{6.13}$$

The constants are pulled before the derivation:

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 \sum_{i=1}^{m} P_i \Delta s_{\mathrm{LOS}_i} \frac{\mathrm{d}}{\mathrm{d}x_k} \Big(x_i f\big(g_i(x)\big) \Big). \tag{6.14}$$

The product rule is applied:

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 \sum_{i=1}^{m} P_i \Delta s_{\mathrm{LOS}_i} \left(f\big(g_i(x)\big) \frac{\mathrm{d}}{\mathrm{d}x_k} x_i + x_i \frac{\mathrm{d}}{\mathrm{d}x_k} f\big(g_i(x)\big) \right). \tag{6.15}$$

The two sums with the derivations are treated separately:

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 \sum_{i=1}^{m} P_i \Delta s_{\mathrm{LOS}_i} f\big(g_i(x)\big) \frac{\mathrm{d}}{\mathrm{d}x_k} x_i + c_1 \sum_{i=1}^{m} P_i \Delta s_{\mathrm{LOS}_i} x_i \frac{\mathrm{d}}{\mathrm{d}x_k} f\big(g_i(x)\big). \tag{6.16}$$

In the first part $\frac{\mathrm{d}}{\mathrm{d}x_k} x_i$ is δ_{ik} and the first sum over i disappears:

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 P_k \Delta s_{\mathrm{LOS}_k} f\big(g_k(x)\big) + c_1 \sum_{i=1}^{m} P_i \Delta s_{\mathrm{LOS}_i} x_i \frac{\mathrm{d}}{\mathrm{d}x_k} f\big(g_i(x)\big). \tag{6.17}$$

Now the first part is in its final state and the second part is evaluated. For the derivation of f the chain rule is applied to get:

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 P_k \Delta s_{\mathrm{LOS}_k} f\big(g_k(x)\big) + c_1 \sum_{i=1}^{m} P_i \Delta s_{\mathrm{LOS}_i} x_i \frac{\partial f\big(g_i(x)\big)}{\partial g_i(x)} \frac{\partial g_i(x)}{\partial x_k}. \tag{6.18}$$

As f is derived numerically and f is a smooth function, $\frac{\partial f(g_i(x))}{\partial g_i(x)}$ can also be derived numerically and for simplicity $f' = \frac{\partial f(g_i(x))}{\partial g_i(x)}$ is introduced.

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 P_k \Delta s_{\mathrm{LOS}_k} f\big(g_k(x)\big) + c_1 \sum_{i=1}^{m} P_i \Delta s_{\mathrm{LOS}_i} x_i f'\big(g_i(x)\big) \frac{\partial g_i(x)}{\partial x_k}. \tag{6.19}$$

For the derivative of g the sum rule is applied to pull the derivation into the sums and get Kronecker-δ. Then the sums vanish:

$$\frac{\partial g_i(x)}{\partial x_k} = \Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,k}}} + \Delta s_{\mathrm{LFS}_{\mathrm{gc}_{i,k}}}. \tag{6.20}$$

One finally obtains:

$$\frac{\mathrm{d}}{\mathrm{d}x_k} I = c_1 P_k \Delta s_{\mathrm{LOS}_k} f\big(g_k(x)\big) + c_1 \sum_{i=1}^{m} P_i \Delta s_{\mathrm{LOS}_i} x_i f'\big(g_i(x)\big)\big(\Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,k}}} + \Delta s_{\mathrm{LFS}_{\mathrm{gc}_{i,k}}}\big). \tag{6.21}$$

Note, that if f would be independent of $\vec{x}$, then $f' = \frac{\partial f(g_i(x))}{\partial g_i(x)}$ would vanish and with it the second term. This finally would lead to $\mathbf{K} = \mathbf{J}$. However, as f depends on $\vec{x}$, a slightly different matrix has to be formed. Note, that for this special mathematical problem omitting the second addend in Eq. (6.21) reproduces synthetic model density profiles that are forward modeled with the forward model and then again inverted with the retrieval algorithm very well. This is easier and faster than solving the full equation. However, $\mathbf{K}$ still depends on $\vec{x}$. This and further consequences of the dependency of f on $\vec{x}$ are discussed in Sect. 6.3.5. Now $\mathbf{J}$ can be built elementwise with Eq. (6.21). For one measurement the Jacobi matrix is a $1 \times m$ matrix, if there is one measurement I_1 and m grid cells. The k_{th} element of this matrix is just $\frac{d}{dx_k} I$. So for one measurement it looks like this.

$$J = \begin{pmatrix} \frac{d}{dx_1} I_1 & \frac{d}{dx_2} I_1 & \dots & \frac{d}{dx_m} I_1 \end{pmatrix} \tag{6.22}$$

If there are more measurements (e.g., n), additional rows are added.

$$J = \begin{pmatrix} \frac{d}{dx_1} I_1 & \frac{d}{dx_2} I_1 & \dots & \frac{d}{dx_m} I_1 \\ \frac{d}{dx_1} I_2 & \frac{d}{dx_2} I_2 & \dots & \frac{d}{dx_m} I_2 \\ \dots & \dots & \dots & \dots \\ \frac{d}{dx_1} I_n & \frac{d}{dx_2} I_n & \dots & \frac{d}{dx_m} I_n \end{pmatrix} \tag{6.23}$$

6.3.3 Calculation of path matrices

To solve Eq. (6.21) the path matrices $\Delta s_{\mathrm{LOS}_i}$, $\Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,j}}}$ and $\Delta s_{\mathrm{LFS}_{\mathrm{gc}_{i,j}}}$ introduced in Sect. 6.3.1 are needed. These matrices are built once during the initialization of the program and are later used together with the initial density profile (introduced in Sect. 6.3.5) of the last iteration step to calculate $g_i(x)$ in Eq. (6.21).

The paths are calculated on a 2D grid of altitudes and latitudes. To overcome ambiguity at the poles, additional 82° of latitudes are added at each pole to separate dayside and nightside latitudes. Latitudes higher than 82° are ignored in both hemispheres as this is the highest latitude covered by SCIAMACHY measurements.

An example plot for path lengths in different grid cells for a typical LOS is shown in Fig. 6.15. All altitude intervals above the tangent altitude are passed twice by the LOS (from the satellite's point of view, downwards and upwards). To overcome ambiguities of different non-connected parts of the LOS that are, e.g., several 100 km apart, but are within the same altitude and latitude interval, the LOS

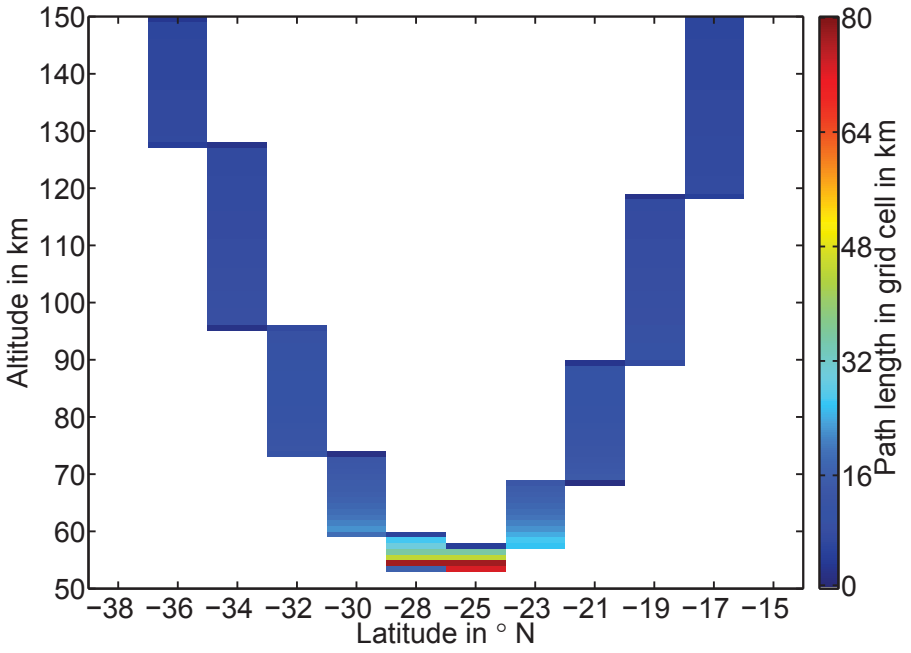

Fig. 6.15: Path lengths in different grid cells for a typical line-of-sight (LOS) of a limb measurement. The biggest part of the path lies in the tangent point altitude region, but higher altitudes are also passed. Note, that for the contribution of a grid element's emission to the total emission along the LOS, the path length in the grid interval, as well as the density of the emitters in the grid interval and the absorption of the emitted light along the LOS have to be considered. For measurements with tangent altitudes far below the metal layer peak, where the density is far smaller than in the peak region, the emission signal mainly comes from the region above the tangent altitude.

is separated at the tangent point into 2 parts. This separation is necessary, because when there is a high density region between both parts of the LOS within the same grid cell, then the optical depth and with it the absorption specific function f and f' are quite different for both parts. Therefore, the absorption relevant matrices are built up separately for each side of the tangent point of the LOS. However, in the end, after building up the separate Jacobians for both parts of the LOS, both Jacobians are added.

For each side on the LOS, $\Delta s_{\mathrm{LOS}_i}$ has to be computed once for each measurement and $\Delta s_{\mathrm{LOS}_{\mathrm{gc}_{i,j}}}$, which includes only the elements of $\Delta s_{\mathrm{LOS}_i}$ that are closer to the satellite than the grid cell i, is quickly built from $\Delta s_{\mathrm{LOS}_i}$ by copying the valid non-zero elements. For the far side of the LOS matrices, $\Delta s_{\mathrm{LOS}_i}$ for both parts of the LOS are needed.

There are only minor differences in the calculation of $\Delta s_{\mathrm{LFS}_{\mathrm{gc}_{i,j}}}$ to the calculation of $\Delta s_{\mathrm{LOS}_i}$, e.g., the additional finding of the tangent point of the straight line, which contains the LFS, and to find out whether this tangent point is also a part of the LFS. For SZA $< 75°$, LFS contributions are negligibly small. For SZA $\approx 90°$, the LFS contribution is as strong as the LOS contribution. Also, measurements with

a too high SZA and with a tangent point of the LFS in the lower atmosphere have to be excluded, because the model does not include lower atmospheric contributions. Note, that refraction of radiation, due to the dense atmosphere, can be neglected in the MLT region.

The actual calculation of the path length in each grid element is done as follows: for a pure vertical grid right-angled triangle algebra is used to derive the path length in each altitude interval. The right angled triangle is formed by the maximum altitude of the grid cell as the hypotenuse c and the tangent height as the adjacent side b to derive the angle $\alpha = \arccos(\frac{b}{c})$ and the opposite side a as $a = b \tan\alpha$ or $a = 0$ for $c < b$. The path length in each altitude level is then derived as the difference of the opposite sides a for neighboring altitude levels. This is illustrated in Fig. 6.16.

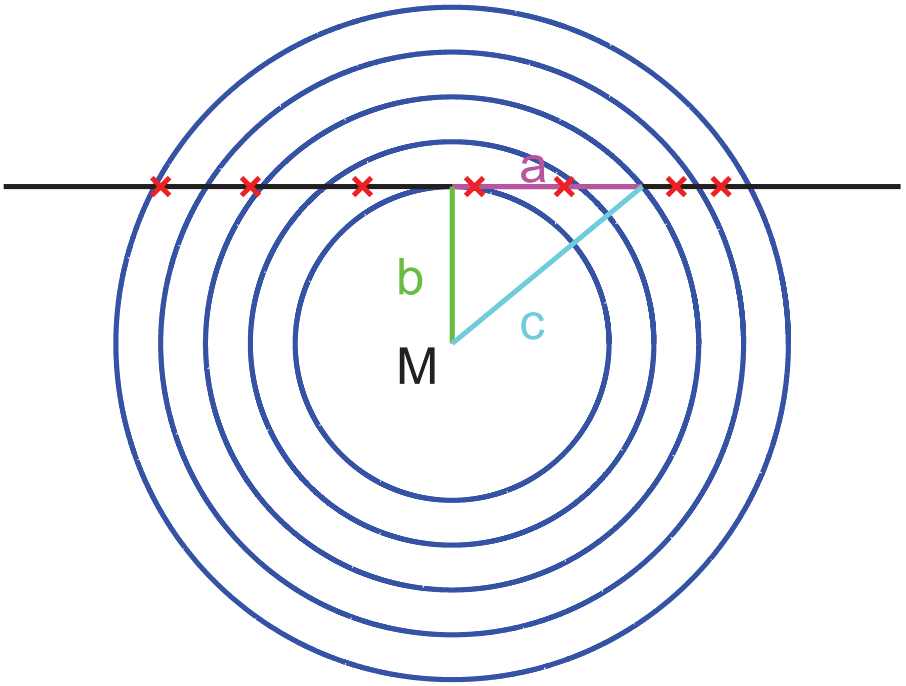

Fig. 6.16: 2D intersection of Earth's atmosphere, with the center of Earth, M, and the altitudes as radii of the circles. Path lengths along the line-of-sight for the vertical grid can be found with right-angled triangle algebra. Changes of the latitude in a vertical grid cell are added as additional sides a (red crosses). The path length in each grid cell is the difference of the sides a of neighboring grid cells. Note, that, depending on the binning of the latitudes, it is possible that all grid cells only have one latitude.

To calculate the changes of latitudes within one altitude interval, additional sides a have to be found, which represent the latitude changes. These additional sides are sorted into the vector of the opposite sides, before deriving the differences of neighboring sides a in a loop, and each time one of the additional sides a is found

in the loop, the latitude interval in the 2D grid changes. The straight line equation of the LOS is set equal to the cone equation of the boundary latitude for each latitude cone, to obtain the additional sides. There may be 3 different kind of solutions for this mathematical problem. No intersection at all, one (tangential) intersection, or two intersections, and these special cases have to be treated individually. A double cone, that starts at the origin (center of Earth) into z direction can be described in cartesian coordinates by:

$$x^2 + y^2 - tan^2(\alpha)z^2 = 0 \tag{6.24}$$

α is the opening angle of the right-angled triangle, that is spinned in a circle to form the cone. The opening angle is 2α. The following parameterization of the straight line equation for the LOS is used:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_{TP} \\ y_{TP} \\ z_{TP} \end{pmatrix} + \lambda \begin{pmatrix} x_{LOS} \\ y_{LOS} \\ z_{LOS} \end{pmatrix} \tag{6.25}$$

TP is the location vector of the tangent point, while x_{LOS}, y_{LOS} and z_{LOS} are the components of the vector along the LOS. The direction vector is normalized (e.g., to 1 km), so λ is the length of the additional opposite sides a to be determinated. Putting (6.25) into (6.24) leads to a quadratic equation for λ as follows:
Cone equation:

$$x^2 + y^2 - tan^2\alpha z^2 = 0 \tag{6.26}$$

Putting in straight line equation:

$$(x_{TP} + \lambda x_{LOS})^2 + (y_{TP} + \lambda y_{LOS})^2 - tan^2\alpha(z_{TP} + \lambda z_{LOS})^2 = 0 \tag{6.27}$$

Dissolving the brackets:

$$\begin{aligned} &x_{TP}^2 + 2\lambda x_{LOS}x_{TP} + \lambda^2 x_{LOS}^2 \\ &+y_{TP}^2 + 2\lambda y_{LOS}y_{TP} + \lambda^2 y_{LOS}^2 \\ &-tan^2\alpha z_{TP}^2 - tan^2\alpha 2\lambda z_{LOS}z_{TP} - tan^2\alpha\lambda^2 z_{LOS}^2 = 0 \end{aligned} \tag{6.28}$$

Sorting for λ^2,λ^1 and λ^0:

$$\begin{aligned} &\lambda^2(x_{LOS}^2 + y_{LOS}^2 - tan^2\alpha z_{LOS}^2) \\ &+2\lambda(x_{LOS}x_{TP} + y_{LOS}y_{TP} - tan^2\alpha z_{LOS}z_{TP}) \\ &+x_{TP}^2 + y_{TP}^2 - tan^2\alpha z_{TP}^2 = 0 \end{aligned} \tag{6.29}$$

Dividing by $(x_{LOS}^2 + y_{LOS}^2 - tan^2\alpha z_{LOS}^2)$:

$$\lambda^2 + 2\lambda\frac{x_{LOS}x_{TP} + y_{LOS}y_{TP} - tan^2\alpha z_{LOS}z_{TP}}{x_{LOS}^2 + y_{LOS}^2 - tan^2\alpha z_{LOS}^2} + \frac{x_{TP}^2 + y_{TP}^2 - tan^2\alpha z_{TP}^2}{x_{LOS}^2 + y_{LOS}^2 - tan^2\alpha z_{LOS}^2} = 0 \tag{6.30}$$

Solving the quadratic equation:

$$x^2 + px + q = 0 \tag{6.31}$$

for the two solutions:

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} \tag{6.32}$$

Using the right parameters:

$$\lambda_{1,2} = -\frac{x_{LOS}x_{TP} + y_{LOS}y_{TP} - tan^2\alpha z_{LOS}z_{TP}}{x_{LOS}^2 + y_{LOS}^2 - tan^2\alpha z_{LOS}^2} \pm \sqrt{(\frac{x_{LOS}x_{TP} + y_{LOS}y_{TP} - tan^2\alpha z_{LOS}z_{TP}}{x_{LOS}^2 + y_{LOS}^2 - tan^2\alpha z_{LOS}^2})^2 - \frac{x_{TP}^2 + y_{TP}^2 - tan^2\alpha z_{TP}^2}{x_{LOS}^2 + y_{LOS}^2 - tan^2\alpha z_{LOS}^2}} \tag{6.33}$$

This solution still looks quite long, but more simplifications cannot be made. However, some useful comments can be made. As it is favorable to split the LOS at the tangent point into two sections to avoid ambiguities, only a positive solution of λ will be discussed here, because a similar discussion for the other side of the LOS can be easily made using $-\vec{x}_{LOS}$ instead of $\vec{x}_{LOS}$. As for one measurement geometry the tangent point and the LOS are given, the only dependence that is still in λ is the latitudinal dependence of $\alpha = 90° -$ latitude. The square of the tangent occurs in (6.33). Therefore, it has to be evaluated on which hemisphere the straight line intersects the double cone. There is a maximum latitude, which is passed by the LOS, and the argument of the square root is zero at this latitude and negative for higher latitudes. The other case, when the square root becomes zero is for equator crossing. In this case α is 90° and $\tan(\alpha)$ is infinite. For the equator crossing Eq. (6.33) is simplified to:

$$\lambda_{eq} = -\frac{z_{TP}}{z_{LOS}} \tag{6.34}$$

In the cases with 2 valid solutions for λ the hemisphere can be determined by comparison with λ_{eq}. For $\lambda < \lambda_{eq}$ the intersection is on the same hemisphere as the tangent point and for $\lambda > \lambda_{eq}$ the intersection is on the other hemisphere as the tangent point. Since this calculation is not done for only one latitude, but for a latitude interval, it has to be checked, whether the upper or the lower limit of the latitude interval has to be used. For the latitude grid with day and night side separated at 82° consecutive latitude grids are passed in a monotonic direction along the LOS.

6.3.4 Calculation of self-absorption contribution f and f'

As pointed out in Sect. 6.2.3, the emission depends on the product of the incoming solar irradiation and the absorption cross section. The same applies for the absorption. However, while for the emission the spectrally integrated values of both quantities are sufficient, proper consideration of the spectral variations is crucial for the treatment of the absorption part. For monochromatic radiation, the exponential Beer–Lambert absorption law $I = I_0 e^{-\sigma \int n ds}$ can be applied with a path independent absorption cross section σ in the exponent. The density n is integrated along the absorption path s. However, because of the relatively low spectral resolution of SCIAMACHY, the spectral radiance measured is integrated over a certain spectral range, which is essentially determined by the width of the instrument function.

Therefore, the Beer–Lambert law is simulated for monochromatic radiation and integrated for different $g = \int n\mathrm{d}s$ (Eq. 6.11). The emission is large for wavelengths with large absorption cross sections. However, a larger absorption cross section also means more absorption. Therefore, the metallic layer becomes optically thick for lower g in the center of the line compared to the wings of the line. This leads to a "hole-burning" effect in the product of cross section σ, irradiance πF and absorption attenuation $e^{-\sigma \int n\mathrm{d}s}$. This is illustrated in Figs. 6.17 and 6.18.

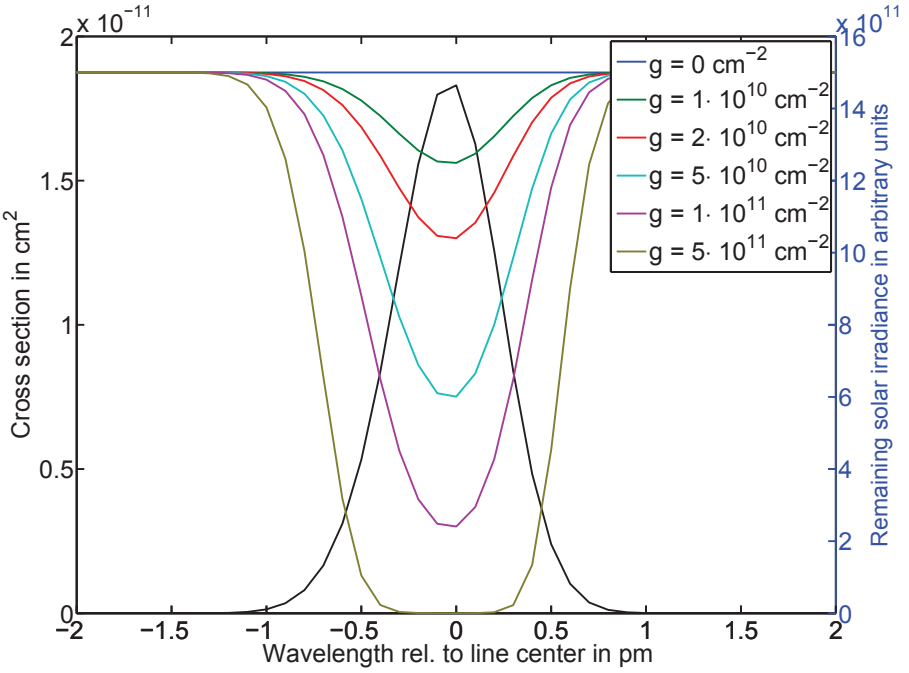

Fig. 6.17: Absorption cross section for the Mg 285.2 nm line (black line and left-hand ordinate), which has to be multiplied with the attenuated solar irradiance (other lines and right-hand ordinate) to obtain the wavelength-specific emissivities shown in Fig. 6.18. The initially incoming radiation is partly absorbed along the line from Sun (LFS) before the absorption and re-emission process into the line-of-sight (LOS). After the re-emission process, radiation is further absorbed along the LOS. The parts of g from the LOS and the LFS may not be treated separately. The spectral irradiance profile, which is initially assumed to be flat, is absorbed where the cross section is high. This leads to a spectral hole burning.

For small $\int n\mathrm{d}s$ one obtains an approximate Beer–Lambert Law with a constant effective cross section. Hunten (1956) showed that, for a Gaussian-shaped cross section profile, the effective cross section is $\frac{1}{\sqrt{2}}$ times the cross section at the maximum of the Gaussian. For larger $\int n\mathrm{d}s$, the effective cross section becomes smaller as the remaining light comes more and more from the edges of the absorption cross section profile. f is determined numerically using Eq. (6.35):

$$f = \frac{\int \sigma(\lambda)\pi F(\lambda) \cdot e^{-\sigma(\lambda)g}\mathrm{d}\lambda}{\int \sigma(\lambda)\pi F(\lambda)\mathrm{d}\lambda}. \tag{6.35}$$

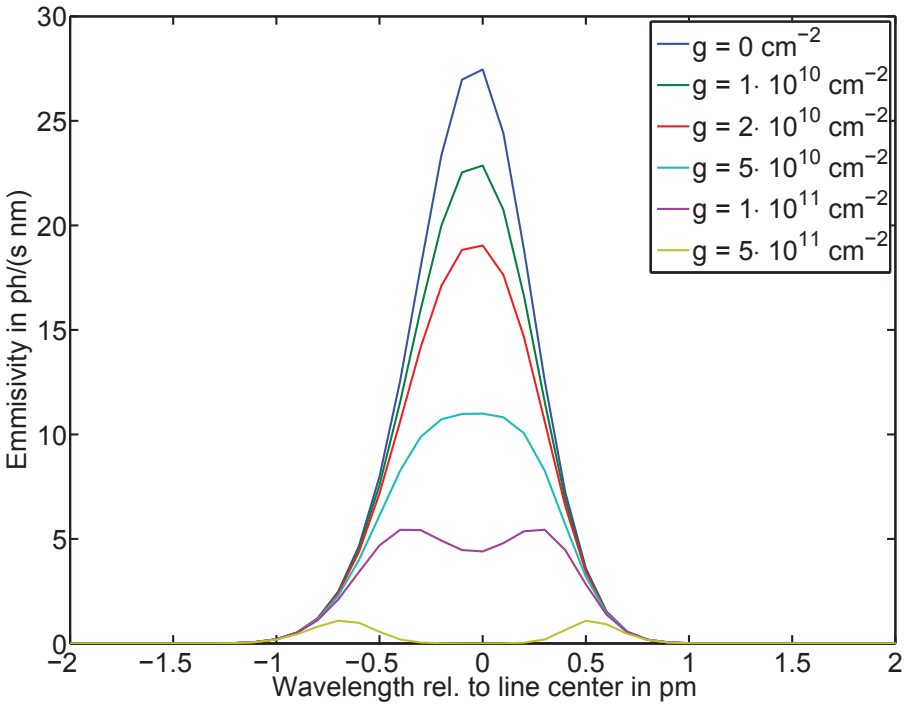

Fig. 6.18: Calculation of the wavelength-specific emissivity for the Mg 285.2 nm line as a function of wavelength difference from the center of line. It is the product of the solar irradiance, attenuated along the line from Sun and the line-of-sight, and the absorption cross section (both shown in Fig. 6.17). To obtain the emissivity of the spectral line, this spectrum has to be integrated over all wavelengths. For strong absorption the light comes mainly from the edges of the line instead of the line center. This reduces the effective cross section for the total integrated profile.

f' is obtained by numerical differentiation of f. For Mg and Mg^+, isotopic shifts reduce the cross section in the center of the line (Fig. 6.19) and therefore lead to reduced self-absorption. Results of the numerical calculation of f are shown in Fig. 6.20.

For a ray starting from the satellite along the LOS through a strongly absorbing layer, the f factors along the LOS decrease rapidly. For the total emission from the LOS, the single emissions of each line segment have to be integrated. Because f decreases more rapidly for higher densities, this leads to a compression, and for a monochromatic line even to a saturation effect for the conversion of true SCD $g = \int n ds$ to the measured SCD ($\int f n ds$). This is illustrated in Fig. 6.21. For typically obtained values of SCDs in the order of $0.5 \times 10^{11}\,\mathrm{cm}^{-2}$ for Mg, this can lead to an issue with measurement noise (see also Fig. 6.26).

6.3.5 Linearization and iteration

The system of linear equations $\mathbf{J}^T\mathbf{J}\vec{x} = \mathbf{J}^T\vec{y}$ has to be solved. This equation system is not linear as $\mathbf{J} = \mathbf{J}(\vec{x})$, because of the self-absorption contributions in f and f'. This problem is overcome by an iterative approach using $\mathbf{J} = \mathbf{J}(\vec{x}_{i-1})$ to calculate

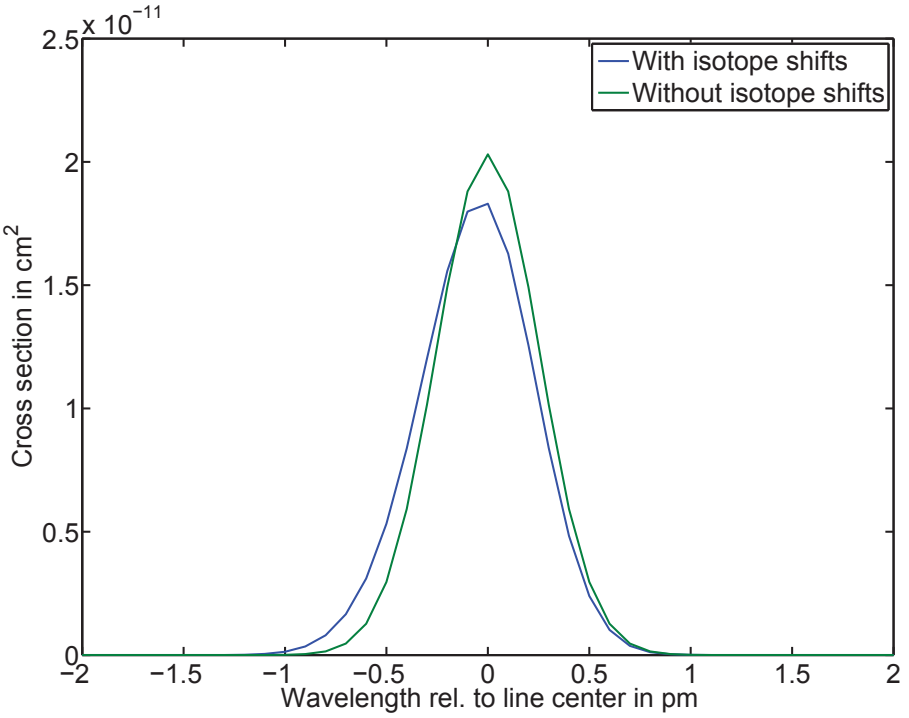

Fig. 6.19: Absorption cross section for the Mg 285.2 nm line with and without considering isotopic shifts. The wavelength integrated cross section is the same for both cases. However, through the isotopic shifts for the heavier isotopes ^{25}Mg and ^{26}Mg, a part of the cross section is shifted from the center of the line to the edge of the line, which effectively reduces the absorption cross section for the integrated line, because the monochromatic Beer-Lambert law is exponential in the cross section.

x_i with starting values $\vec{x}_0$ for the first iteration step. It is assumed, that the metal layer is not opaque, because otherwise the SCD profile would be almost viewing angle independent and, therefore, almost constant below a certain tangent altitude, which is not the case (see, e.g., Fig. 6.9). Under this assumption, starting with $\vec{x}_0 = \vec{0}$ is a good choice, because when choosing $\vec{x}_0 = \vec{0}$ the Jacobian is the same, as if self-absorption was not considered at all, and a profile shape is obtained that is already close to the final solution after the first step. This first step solution can be interpreted as the actual starting condition, which is close enough to the final result so that the iteration converges to the final result.

In order to investigate how many iterations typically are required to achieve convergence, up to 50 iterations were run for typical profiles. Convergence is already achieved quickly. After 20 iterations, which are used to obtain the results, the largest difference for an individual grid cell from step 19 to step 20 is less than 1 %. The changes in most of the other grid cells are at least a factor of 10 to 100 smaller.

6.3.6 Constraints

Minimizing only the sum of the squares of the differences between the forward model $\mathbf{K}\vec{x}$ and the SCD profiles $\vec{y}$ typically leads to strongly oscillating solutions $\vec{x}$. There-

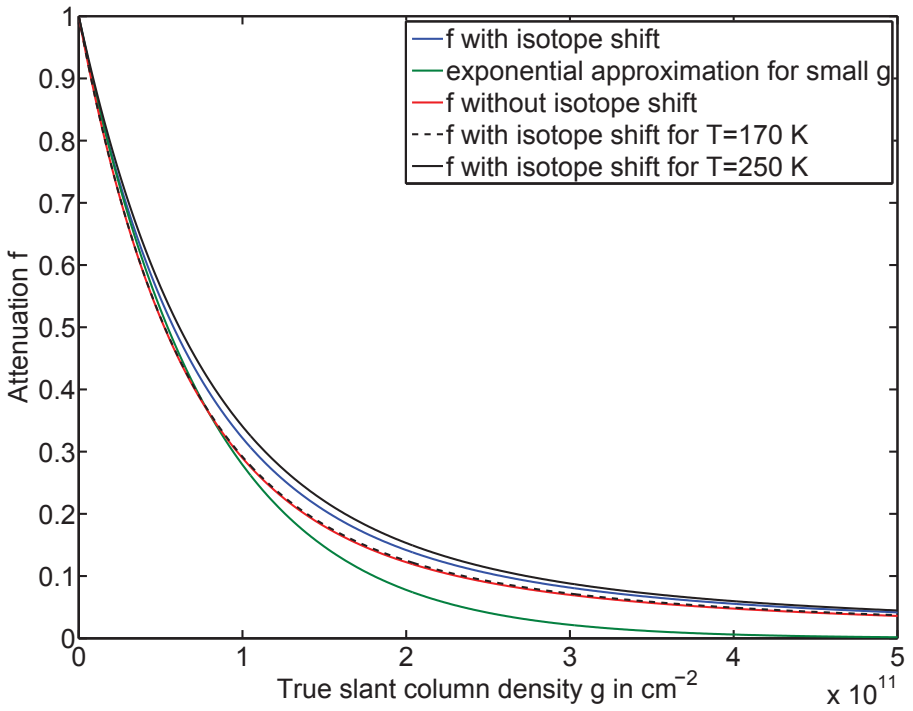

Fig. 6.20: Calculated attenuation f as a function of g for the Mg 285.2 nm line. Taking isotopic shifts into account reduces the absorption effect and f is higher for this case. The black lines show variations in the absorption behavior in the range of typical temperatures in the 90 km to 100 km altitude range, resulting in a different Doppler broadening of the absorption cross section profiles.

fore, constraints are introduced to stabilize the retrieval and to obtain more realistic results by reducing the oscillations. An a priori constraint (this is also called Tikhonov regularization see, e.g., Tikhonov, 1943) and smoothing constraints for neighboring latitude and altitude grid cells are used. The constraints have to be chosen as strong as required to stabilize the solution, but as weak as possible to not influence the solution too strongly. The final equation that has to be solved is Eq. (6.36), which represents a regularized iterative linearized non-linear least square fit with a regularization dominated by smoothing constraints:

$$\left(\mathbf{J}^T\mathbf{S}_y\mathbf{J} + \mathbf{S}_\mathrm{a} + \lambda_h\mathbf{S}_H^T\mathbf{S}_H + \lambda_\phi\mathbf{S}_\phi^T\mathbf{S}_\phi\right)\vec{x} = \mathbf{J}^T\mathbf{S}_y\vec{y} + \underbrace{\mathbf{S}_\mathrm{a}\underbrace{\vec{x}_\mathrm{a}}_{=0}}_{=0} \tag{6.36}$$

with the Jacobian $\mathbf{J}$ of the forward model $\mathbf{K}\vec{x}$. The a priori covariance matrix $\mathbf{S}_\mathrm{a}$ is in fact a scalar ($\lambda_{\text{a-priori}}$) multiplied with an identity matrix. $\mathbf{S}_H$ and $\mathbf{S}_\phi$ are the matrices for altitudinal and latitudinal constraints (large sparse matrices with only 2 diagonals of non-zero values) and λ_h and λ_ϕ are the scalar weighting factors for both constraints. $\vec{x}$ is the vector of number densities. On the right-hand side there is the covariance matrix for the SCDs ($\mathbf{S}_y$), which is assumed to be diagonal: the SCDs $\vec{y}$ and the a priori solution $\vec{x}_\mathrm{a}$. As there should not be any bias on the form

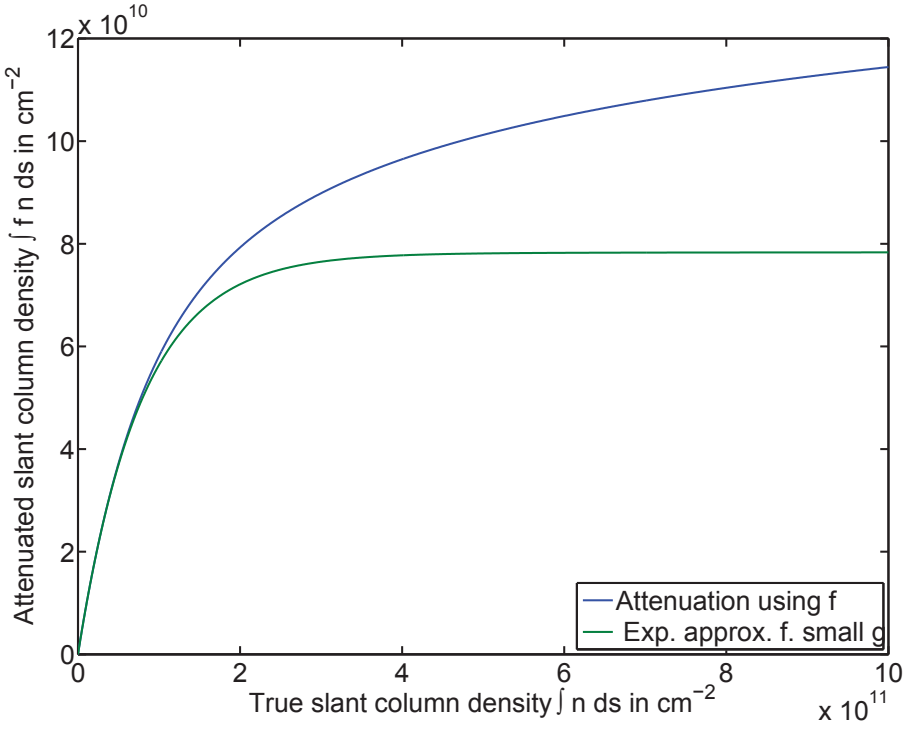

Fig. 6.21: Plot of attenuated slant column densities (SCDs) versus the true SCDs for the Mg 285.2 nm line. The attenuated SCDs are determined using f, as well as the exponential approximation of f for small true SCDs. Typical measured SCDs (y axis) in the peak region for Mg are 6–8 $\times 10^{10}\,\mathrm{cm}^{-2}$. The exponential approximation cannot be used in the retrieval, since it is already saturated below $8 \times 10^{10}\,\mathrm{cm}^{-2}$. For Mg the self-absorption effect is so strong that the measured SCDs are a factor 3 to 4 smaller than the true SCDs in the peak region. For Mg^+(not in the plot) the conversion factor is close to 1.

of the profile, $\vec{x}_\mathrm{a} = \vec{0}$ is used. The effect of choosing different vertical smoothing constraints is demonstrated in Sect. 6.4.2.

6.3.7 Optimizations

The number of parameters of the solution and the measurement are n and m, and n and m are nearly equally large. The slowest operations used in the iteration are matrix-matrix-multiplications, as well as the LU-decomposition for solving the equation system. Both scale with n^3 where n is one dimension of the matrix. For all these operations library routines of ATLAS (see, e.g., Whaley et al., 2001) and Lapack (see, e.g., Anderson et al., 1999) are used. Furthermore, the sparsity of the used matrices can be used. Figure 6.22 shows the non-zero elements of the matrices $\mathbf{J}$ and $\mathbf{J}^T\mathbf{J}$ to point out the sparsity of these matrices, that will be used for the optimization.

For multiplications, only the blocks that lead to non-zero results are evaluated. Furthermore, it is advantageous to have the Jacobian J in a band diagonal form with a low bandwidth. This can be achieved by arranging the representation of the

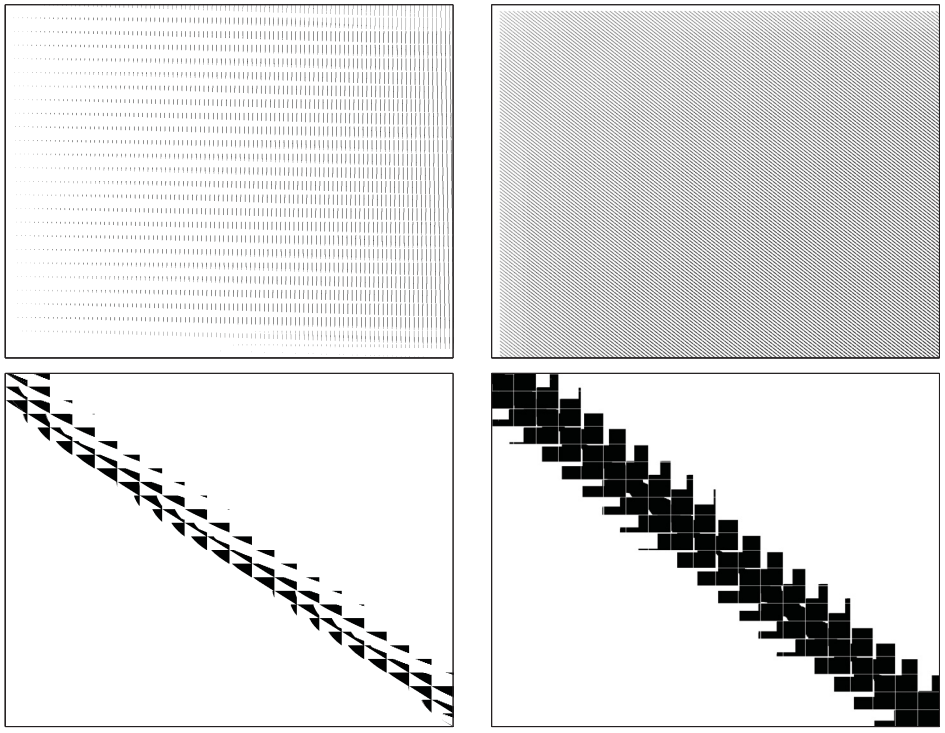

Fig. 6.22: Non-zero elements of $\mathbf{J}$ for a grid that is sorted for altitudes first and then for latitudes (upper left), non-zero elements of $\mathbf{J}^T\mathbf{J}$ for a grid that is sorted for altitudes first and then for latitudes (upper right), non-zero elements of $\mathbf{J}$ for a grid that is sorted for latitudes first and then for altitudes (lower left), and non-zero elements of $\mathbf{J}^T\mathbf{J}$ for a grid that is sorted for latitudes first and then for altitudes (lower right). In both cases $\mathbf{J}$ and $\mathbf{J}^T\mathbf{J}$ are sparse. However, in the lower case, $\mathbf{J}$ has a clear block structure, which is used in later optimization. Similarly $\mathbf{J}^T\mathbf{J}$ has a clear diagonal structure in the lower case, so that a special algorithm for the inversion of band-diagonal matrices is used.

retrieval grid in the J Matrix in such a way, that all measurements of one latitude are grouped. The distance of measurements – in the row of the J Matrix – of same latitude is smaller than the distance of measurements with the same altitude[1]. Using these optimization steps allows to make faster calculations, which allows to do the calculations on finer grids. Furthermore, the speed deficits of the 2D algorithm to a multiple run of 1D algorithms, which arise from the much bigger matrices to be inverted, scaling with n^3, can be nearly canceled out[2], so that the 2D algorithm is as fast as the 1D algorithm on a grid with latitude intervals chosen that wide, that only one latitude interval is passed for each measurement. The bottleneck of the algorithm is the formulation and inversion of Eq. (6.36). Therefore, an analysis of this equation follows:
First the RHS of the equation is discussed:
$\mathbf{S}_a\mathbf{x}_a$: is the multiplication of the diagonal $n \times n$ matrix $\mathbf{S}_a$ with the $n \times 1$ array

[1]Note, that a 2D grid has to be put into a 1D entity, which can be done row-wise or column-wise.
[2]Actually the difference of the 2D and 1D algorithm is only quadratic (2D: n^3, 1D $a(\frac{n}{a})^3 = \frac{n^3}{a^2}$)

$\mathbf{x}_a$. This means, that only the diagonal vector of $\mathbf{S}_a$ has to be multiplied, which leads to n multiplications. However, as $\mathbf{x}_a = 0$ is chosen, this calculation is not even performed in the current program, and must be added, if another a-priori is used.
$\mathbf{J}^T\mathbf{S}_y\mathbf{y}$: $\mathbf{S}_y$ is a diagonal $m \times m$ matrix and is first multiplied with $\mathbf{y}$ requiring m multiplications. $\mathbf{J}^T$ is an $n \times m$ and sparse matrix. However, even if it is dense the number of multiplications is $O(nm)$.
The O(n^3) can be found on the LHS of the equation:
$\lambda_h\mathbf{S}_H^T\mathbf{S}_H$ and $\lambda_\phi\mathbf{S}_\phi^T\mathbf{S}_\phi$ require at least $O(n^3)$ operations, although the used matrices are sparse. Here, it is better to build up the whole matrices directly, which is simple, as both matrices only have few non zero diagonals with a simple pattern. Building up the whole matrices at once is as simple as building the individual factor matrices and performing the multiplication.
The only left term to consider on the LHS is: $\mathbf{J}^T\mathbf{S}_y\mathbf{J}$. It consists of two consecutive matrix multiplications, and it does not matter which matrix multiplication is done first. For the multiplication with $\mathbf{S}_y$ m^2n steps are needed. As $\mathbf{S}_y$ is diagonal only mn multiplications are needed, but the fact that $\mathbf{J}$ is also sparse can be exploited. $\mathbf{J}^T\mathbf{J}$ needs n^2m multiplications. However, as both matrices are sparse a lot of steps can be saved. As every series of consecutive limb measurements [3] covers nearly all altitudes, but only few latitudes, the path matrix has kind of a block form of number of vertical levels (roughly 100) $\times$ number of measurements with nearly the same latitude (roughly 30), whose blocks can be exploited as a new base unit for the matrix, so that the matrix can be scanned for non-zero blocks, rather than non-zero elements. Depending on how many latitude levels are chosen, there are several consecutive blocks in one block row. It just has to be found out for one matrix $\mathbf{J}$ or $\mathbf{J}^T$ how many consecutive non-zero blocks exist in each line, and for the later matrix multiplication only the non-zero blocks of the corresponding lines have to be multiplied. In practical application tests this approach leads to an acceleration of a factor 20 or more.
The resulting matrix, which has to be inverted, is a band diagonal matrix and using an optimized band diagonal algorithm brings a reduction in calculation time by up to a factor 3. Using the optimizations described in this section, the algorithm can be run on finer grids, which reduces discretization errors. However, the resolution of the algorithm does not get better, because it is determined by the instrument's measurement sequence, the measurements geometry, as well as the vertical field of view of the instrument.

6.4 Error estimations and sensitivity to retrieval parameters studies

In this section, sample results are briefly presented to demonstrate, that the algorithm works well. As most of the discussed parameters and effects only have an influence on the vertical profile, only results for vertical profiles in a limited equatorial latitude range are shown.

[3] The consecutive measurement series meant here is also called a state by SCIAMACHY users.

6.4.1 Equatorial vertical retrieval results and error estimations

Figures 6.23–6.25 show the results of vertical profile number density retrievals of Mg at 285.2 nm and Mg^+ at 279.6/280.4 nm for the equatorial region (10° S to 10° N). The red curve in these figures shows the retrieval result using the mean equatorial SCD profile from 2008–2012 as a test profile.

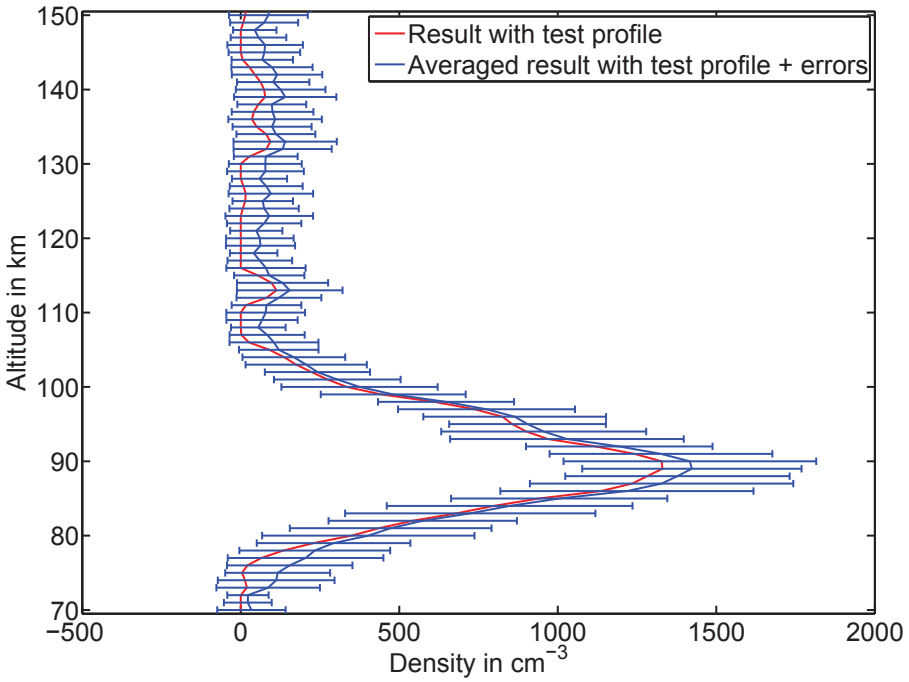

Fig. 6.23: Plot of retrieved Mg (285.2 nm) number density for the 4 year averaged equatorial data set (red). To estimate the error, the retrieval is repeated 1000 times with a typical random Gaussian error for daily averaged data applied to the initial slant column densities. The blue line shows the mean result of the 1000 runs, while the error bars show the standard deviation from the mean values of the 1000 runs. For daily averaged data (which are used here) the differences between both methods are small.

The Mg profile peaks at 90 km and has a FWHM of ≈ 15 km. The Mg^+ profile peaks at 96 km and has a FWHM of ≈ 12 km. Both Mg^+ lines show similar peak values. However, for small peak values the more weakly absorbing 280.4 nm line shows slightly higher peak values, while for regions with high peak values (3000–4000 cm^{-3}), the more strongly absorbing 279.6 nm line shows the higher values. However, the largest differences are smaller than 25 % in the peak region.

To calculate the Mg/Mg^+ number density errors, the mean error of the SCE of the line is taken, corresponding to roughly 1×10^8 ph cm^{-2} s^{-1} sr^{-1} for single measurements. It is half as large as the highest SCE for Mg, and 1/4 (280.4 nm), and 1/8 (279.6 nm) as large as the SCE for Mg^+ in the peak region. To obtain the SCDs, the SCEs have to be multiplied by 4π and divided by γ. The same applies to the errors. For roughly 20 measurements per day (see Sect. 4), the single measurement error is divided by $\sqrt{20} \approx 4.5$ to derive the daily error. Note, that the

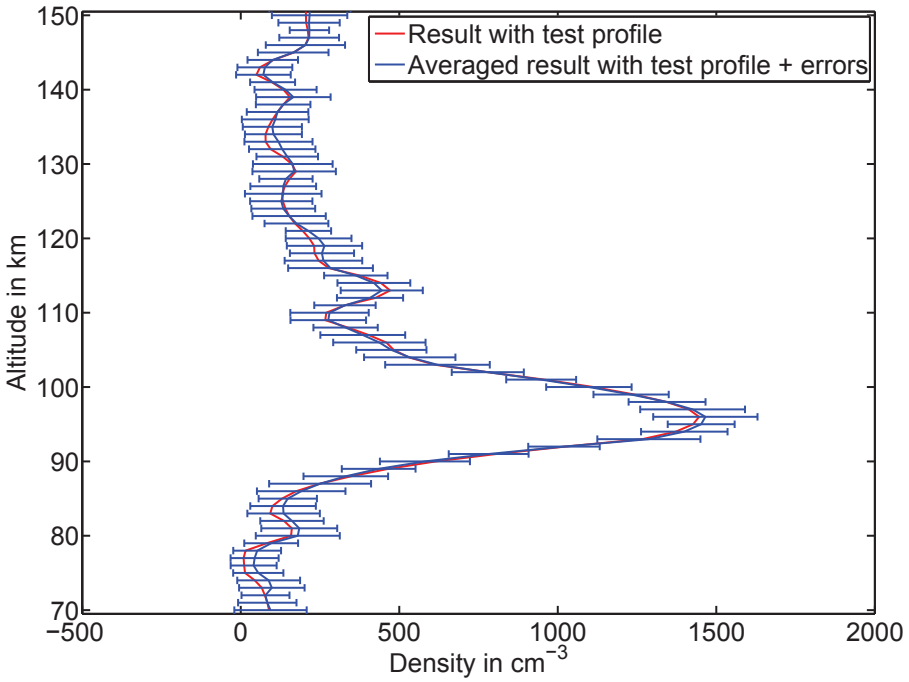

Fig. 6.24: Mg^+ (279.6 nm) number density retrieval result. The same methods as in Fig. 6.23 are applied. The errors for Mg^+ are smaller than for Mg and furthermore non-linear self-absorption effects are smaller for Mg^+. Therefore, both methods lead to the same mean results.

daily error for Mg is roughly as large as the single measurement error for Mg^+ at 279.6 nm.

A Monte Carlo method is used to propagate the radiance errors to the retrieved Mg/Mg^+ density profiles. The Gaussian error is applied to the mean SCD profile, and then the density profile is retrieved. This is repeated often enough (1000 times), that the mean and the standard deviation of the result converge, and the standard deviation is interpreted as the error of the retrieval. The blue line in Figs. 6.23–6.26 shows the mean of these Monte Carlo runs with 1σ error bars. Note, that the error of the mean is the standard deviation divided by a factor $\sqrt{1000} \approx 32$ and, therefore, the mean itself is very accurate.

For Mg the relative errors for daily averages in the peak region are 20–30 %. The mean of the retrievals with additional errors is up to $100\,\mathrm{cm}^{-3}$ higher than the retrieval without errors. This is explained by the non-linearity of the forward model (see Sect. 6.3.4). Figure 6.26 shows the errors for the Mg retrieval using single measurements, and for this case the mean is shifted significantly between both methods. This leads to a systematic difference between averaging the SCDs before the retrieval and averaging densities after the retrieval. Therefore, an averaging must be applied to the SCDs before the retrieval.

On the other hand side the true natural variability has to be taken into account, which is still there, even if perfect measurements without errors are used. Like the variability resulting from errors, the true variability leads to an increase of

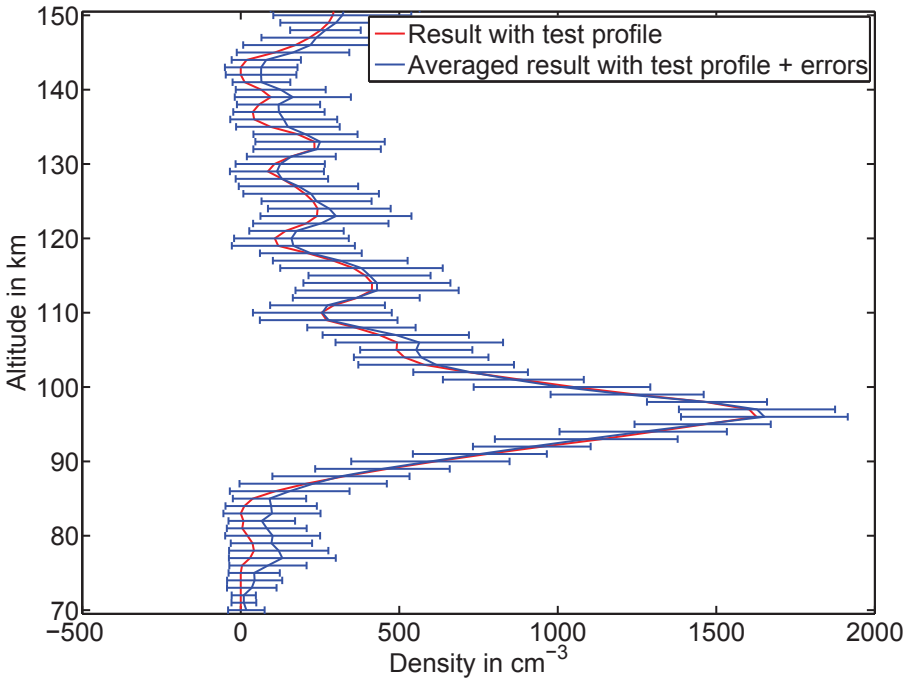

Fig. 6.25: Mg^+ (280.4 nm) number density retrieval result. The same methods as in Fig. 6.23 are applied. The errors for Mg^+ are smaller than for Mg and furthermore non-linear self-absorption effects are smaller for Mg^+. Therefore, both methods lead to the same mean results.

the mean values. Furthermore, the longer the time span of averaged SCEs is, the more the boundary conditions, like scattering angles and solar zenith angles, which are needed for the retrieval, change. Therefore, a compromise between averaging enough data before the retrieval step so that shifts due to errors are excluded, and averaging enough individual results after the retrieval step to account for the natural variability, is made to get the mean density profile at the Equator. As there is one day of measurements available roughly every 14 days, daily averages of the spectra are formed, and densities are retrieved from these daily averages. These densities are further averaged, e.g., to monthly averaged results, to reduce the errors.

The Mg density errors are only small in the peak region, while below 80 km and above 100 km the relative error is bigger than 100 %. However, for the 4 year average (average of all spectra to obtain one SCD profile before retrieving densities), the errors are negligibly small, so that the small peaks at 113 km and between 130–140 km may also be real and, e.g., could originate from sporadic layers. Furthermore, Mg^+ also shows small peaks at the same altitudes.

For Mg^+ the relative errors in the peak region are less than 20 %. The 1σ error is close to 100 % below 85 km, but although the error is also large above 105 km, the density is still larger than 50 cm^{-3} within the error limits and not zero at the top altitude of 150 km.

Although the same constraints are applied for both Mg^+ lines, the results for the 280.4 nm line oscillate more strongly above the peak. And while the peak at

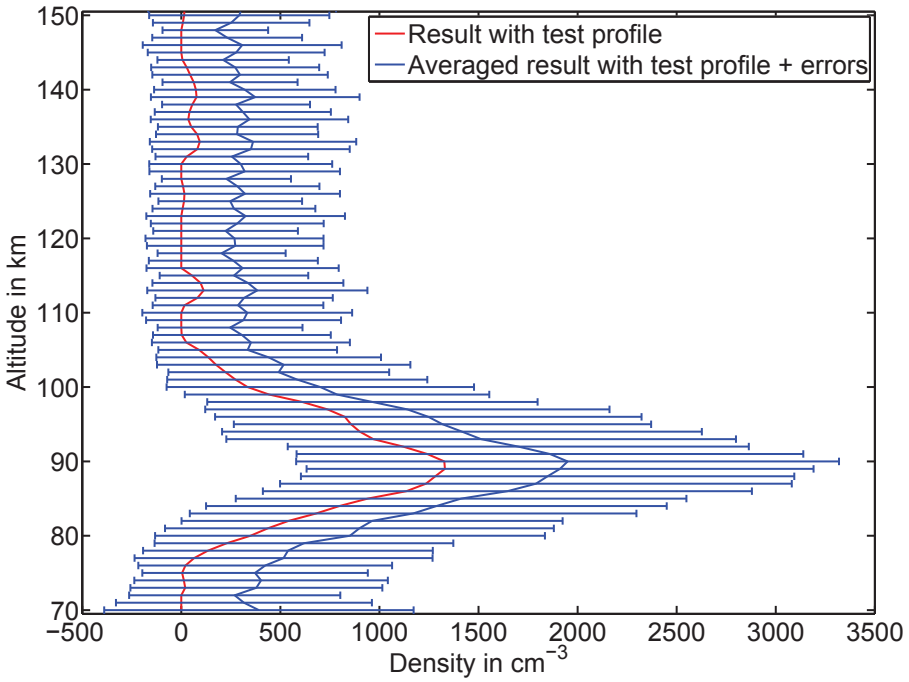

Fig. 6.26: Retrieval of Mg (285.2 nm). The same methods as in Fig. 6.23 are applied. However, here the single measurements error is used. As for high relative errors the error propagation cannot be approximated to be linear and non-linearities in the forward model lead to a shift of the mean to higher values. Therefore, a certain averaging of the data is needed before applying the retrieval algorithm. Note, that the error bars estimate the error of single measurements while the error for the blue line is a factor $\sqrt{1000} \approx 32$ smaller.

113 km occurs for both lines, the higher peaks are at different altitudes. However, both Mg^+ profiles are in good agreement within the error range.

6.4.2 Influences of different constraint strengths

As mentioned in Sect. 6.2.2, different constraints are used. To be effective these constraints should influence but not dominate the retrieval. The a priori and vertical-smoothness-constraint have a similar effect, so both must be tuned together. To simplify the search for an optimal constraint, a fixed ratio of $\lambda_h : \lambda_\phi : \lambda_{\text{a-priori}}$ (see Eq. (6.36), where $\lambda_{\text{a-priori}}$ influences $\mathbf{S}_\text{a}$) of $10 : 2 : 1$ is used, as this was empirically found to yield realistic solutions. To investigate the sensitivity of the retrieval to the constraint parameters, many different combinations with the fixed ratio were tested, ranging from solutions that are obviously too heavily smoothed, to solutions that are obviously too strongly oscillating. Figures 6.27–6.29 show results for the vertical equatorial profile for Mg/Mg^+ for one example day. From this, as well as from latitude–altitude–density plots, and considering other boundary conditions, which cannot be well quantified (e.g., similarity of results for same months, estimation of the difference between retrieving densities from the 4 year averaged data and averaging results from single days, seasonal variations, smoothness of the result

compared with the resolution of at least 3.3 km steps, etc.), the range of acceptable parameters for λ_h for both species is estimated to be between 5×10^{-7} and 5×10^{-9}. For further studies 5×10^{-8} will be used for Mg and 1×10^{-7} will be used for Mg^+ as the optimal solution. Within the range of acceptable parameters the uncertainty of the peak value can be estimated to be $\approx 30\,\%$ for Mg^+ and $\approx 50\,\%$ for Mg.

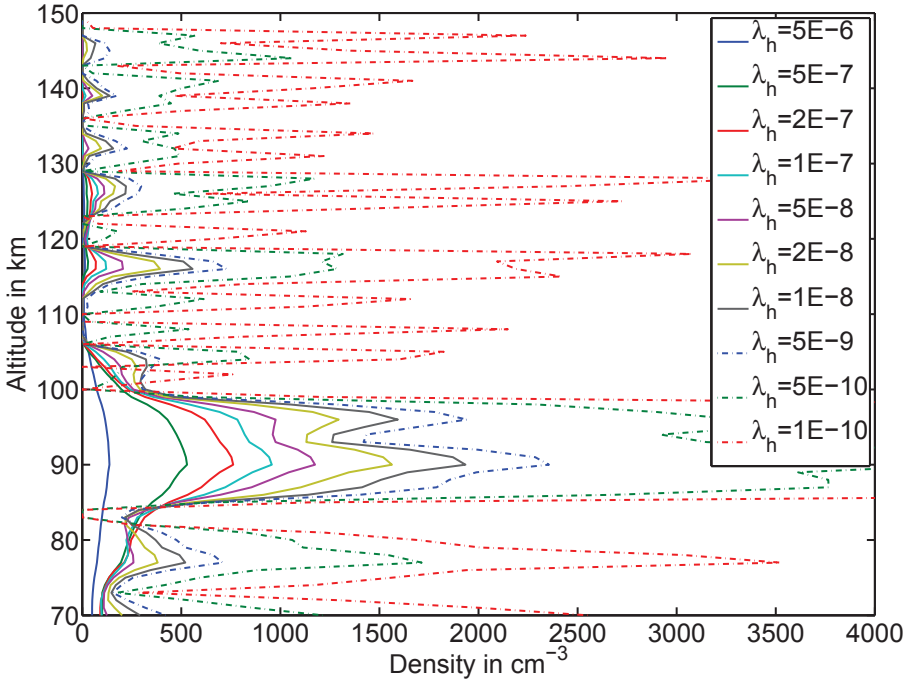

Fig. 6.27: Single day (7 September 2009) vertical profile for Mg (285.2 nm) in the equatorial region from 15° S to 15° N. For too low constraint strength ($\lambda_h < 5 \times 10^{-9}$) the profile is strongly oscillating, while for too high constraint strength ($\lambda_h > 5 \times 10^{-7}$) the profile is too smooth. Acceptable solutions between these two extrema show peak densities from 500 to 2000 cm^{-3}. However, at least the peak altitude is nearly independent of λ_h. The solution with $\lambda_h = 5 \times 10^{-8}$ is favored, considering, e.g., seasonal variations etc.

One of the reasons for the fuzzy transition between the too smooth and too oscillating solutions can be found in the errors. Figure 6.30 shows the sensitivity of the retrieval to the constraint strength using the 4 year averaged data, while Fig. 6.31 shows the results if an error is applied to the same data the same way as discussed in Sect. 6.4.1 (both figures show results for Mg). For simplicity a 1D pure vertical model has been used for both figures, which is very similar to the 2D model.

6.4.3 Influence of different Raman scattering percentage on the profile retrievals

As discussed in Sect. 6.2.2, the Ring effect has significant influence on the Mg profile retrievals for altitudes below 80 km. Figures 6.32 and 6.33 show the retrieved vertical

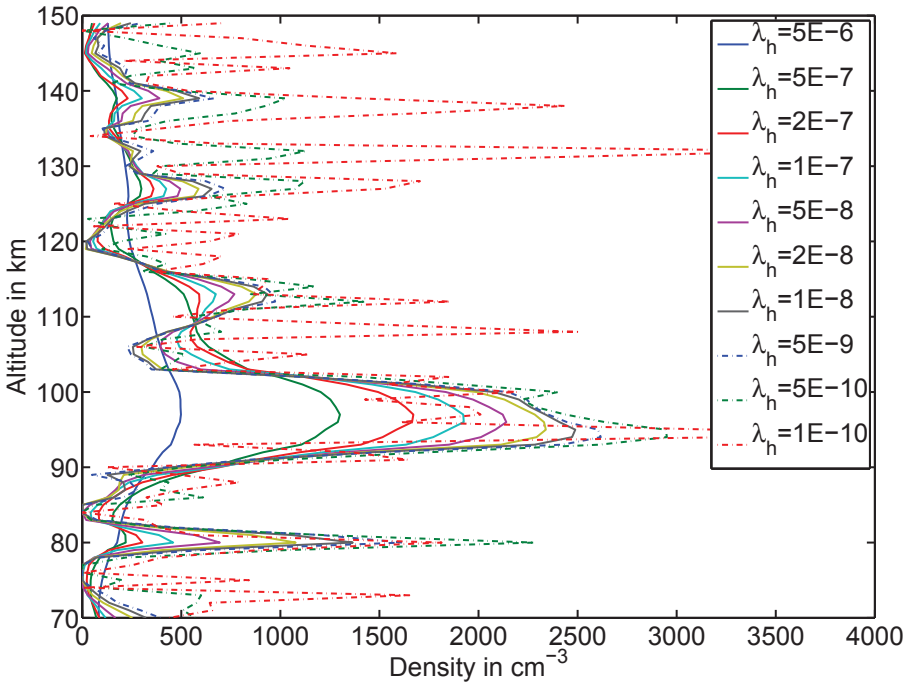

Fig. 6.28: Single day (7 September 2009) result for Mg^+ (279.6 nm) in the equatorial region from 15° S to 15° N. The same approach to find the optimal constraint strength for Mg (see Fig. 6.27) is applied for Mg^+. However, the variation of the peak density with constraint strength, especially for low strength, is much weaker, so errors from a wrong choice of the constraint strength are smaller. For Mg^+ the solution with $\lambda_h = 1 \times 10^{-7}$ will be used for further investigations.

profiles for Mg and Mg^+ with a different contribution of Raman scattering to the total scattering (Raman and Rayleigh).

The Raman scattering percentage is assumed to be $\approx 4\,\%$. The results for Mg still show seasonal and latitudinal variations for the densities at low altitudes, which are not expected to be there (not shown here). However, the densities at low altitudes are zero (within the error limit) for Raman scattering contribution of 3–6 %. The profiles with no or only very weak Ring effect correction still show significant densities at 70 km and only the constraints force the profile near the lower peak edge not to diverge. The densities above 95 km are nearly independent of the Ring correction. The peak value is higher for lower Raman scattering contribution. However, the variations between the most extreme expected contributions of 3–6 % are less than $100\,\mathrm{cm}^{-3}$, so changing the Raman contribution by 1 % changes the peak value by only 2 %. The Mg peak altitude is slightly lower for lower Raman scattering contributions. However, within the range of expected Raman scattering contributions the peak altitude does not change by more than 1 km. Below the peak the differences are bigger, but down to a certain altitude the descent is parallel for all of the considered Raman scattering contributions. The differences are largest below 80 km, as different Ring corrections lead to different residual densities at the lowermost altitudes.

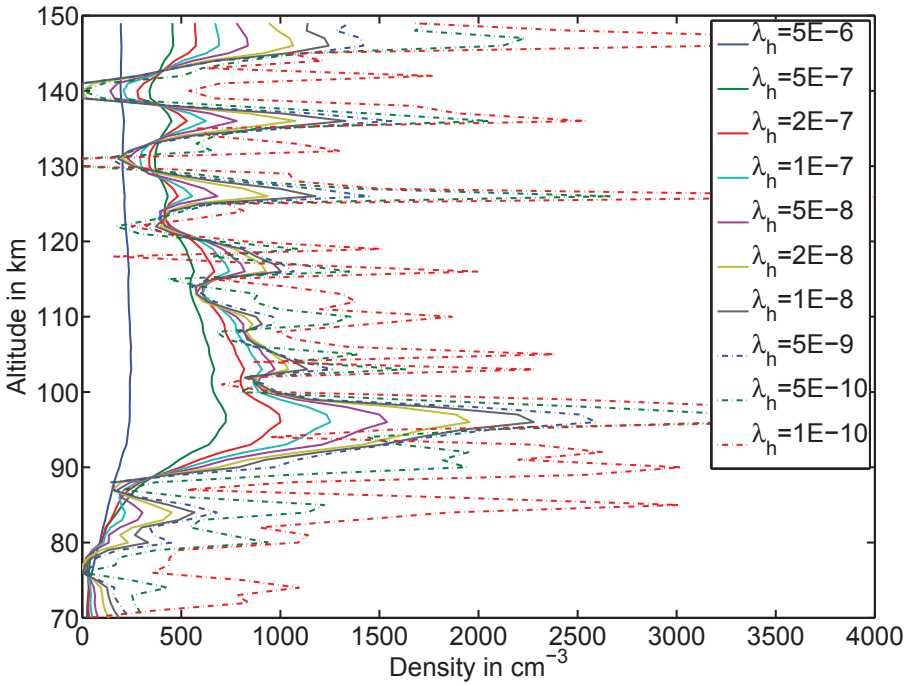

Fig. 6.29: Result for Mg^+ (280.4 nm) for the same conditions as in Fig. 6.28.

For Mg^+ the Ring contribution above 70 km is small, as the double line structure leads only to a filling-in from one side of the absorption line. Furthermore, the Mg^+ lines are not as deep as the Mg line (see Fig. 6.10) and the Mg^+ lines have emission lines in the core. The lower and upper peak edges are similar. The only differences for retrievals with different Ring effect corrections is the peak value, which is higher for a larger Raman scattering contribution, and the densities at the lowermost altitudes that are slightly higher for lower Raman scattering contribution.

The Mg^+ peak is at a higher altitude than the Mg peak, so that the signal in the peak region is mainly dominated by pure emission, as the Rayleigh and Raman scattered background signal is much smaller than the emission signal there. As discussed in Sect. 6.2.2, this leads to slightly larger values in the peak region. And since the corrections above 70 km are negligibly small, it is better not to use the Ring effect correction at all for Mg^+.

6.5 Extension of the retrieval algorithm for Na

Sodium (Na) number densities can be retrieved from the Na D double lines at 590 nm. The vacuum wavelengths of the lines are 589.756 nm for the D_1 line and 589.158 nm for the D_2 line. The two transitions are from the lower most excited states $3^2P_{\frac{1}{2}}$ for D_1 and $3^2P_{\frac{3}{2}}$ for D_2 to the ground state $3^2S_{\frac{1}{2}}$. The relative Einstein coefficient for both line are 1, because only the lowest states are involved, and the transition between both P states is forbidden, because of $\Delta L = 0$.

There are several differences of the Na lines compared to the previously treated Mg and Mg^+ lines, which need to be considered in the radiative transfer model. The

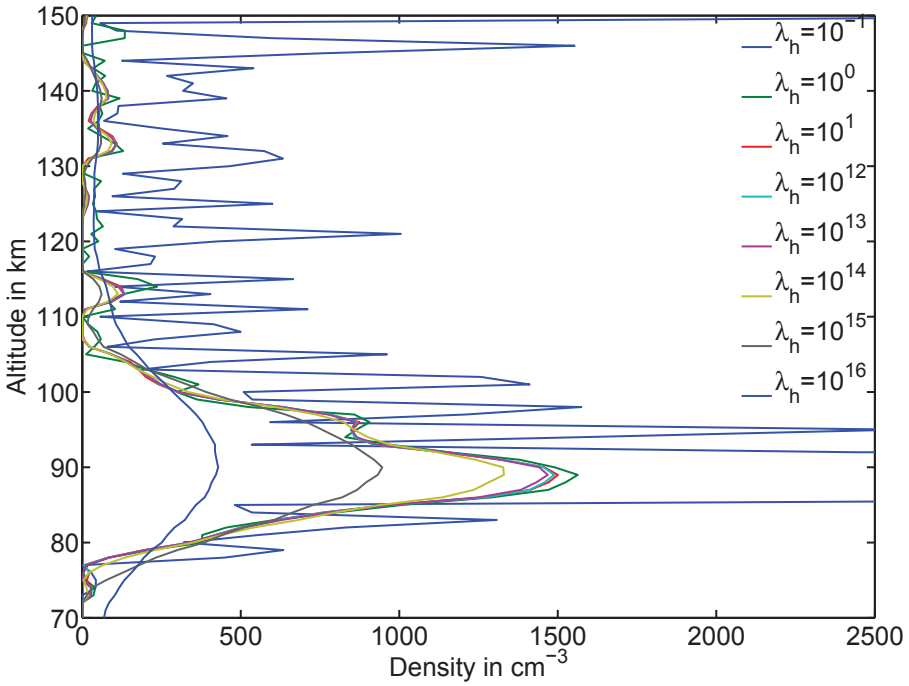

Fig. 6.30: Retrieval results with a 1-D model for the error free 4 year averaged equatorial profile for Mg (comparable with the red line solution in the plots in Sect. 6.4.1). Different vertical constraints are chosen and $\lambda_h = 10 \times \lambda_{\text{a-priori}}$. With increasing constraint strength there is a sharp transition from a totally oscillating solution to a smooth solution. This result is stable within 12 orders of magnitude in the constraints strength until the result is smoothed too strongly. Note, that the numerical value of the constraints is slightly different here than for the 2D model.

stratosphere and troposphere are transparent for electromagnetic radiation in this wavelength range, which introduces the advantage, that the Na D-lines can be observed from ground very well, but also the disadvantage, that lower atmospheric influences have to be considered in the radiative transfer model. The solar Fraunhofer lines for Na are much narrower than the Mg and Mg^+ lines. Therefore, small wavelength shifts have to be considered. The signal to background ratio is higher than for the Mg/Mg^+ lines. Therefore, the Ring effect is negligibly small and its treatment is omitted.

6.5.1 Solar Fraunhofer lines

For the retrieval program, the shift and lineshape information from McNutt and Mack (1963) are used. Following McNutt and Mack (1963), the solar irradiance in the vicinity of the Na D_1 and D_2 lines can be calculated for x (x is defined further below):

$$I(x) = I_0 \cdot e^{(\frac{|x|}{x_e})^A} \tag{6.37}$$

The ratio $\frac{I_0}{I_{\text{baseline}}}$ of the intensity at the line center and the baseline intensity at the edge of the Fraunhofer lines is stated in McNutt and Mack (1963). To scale this to

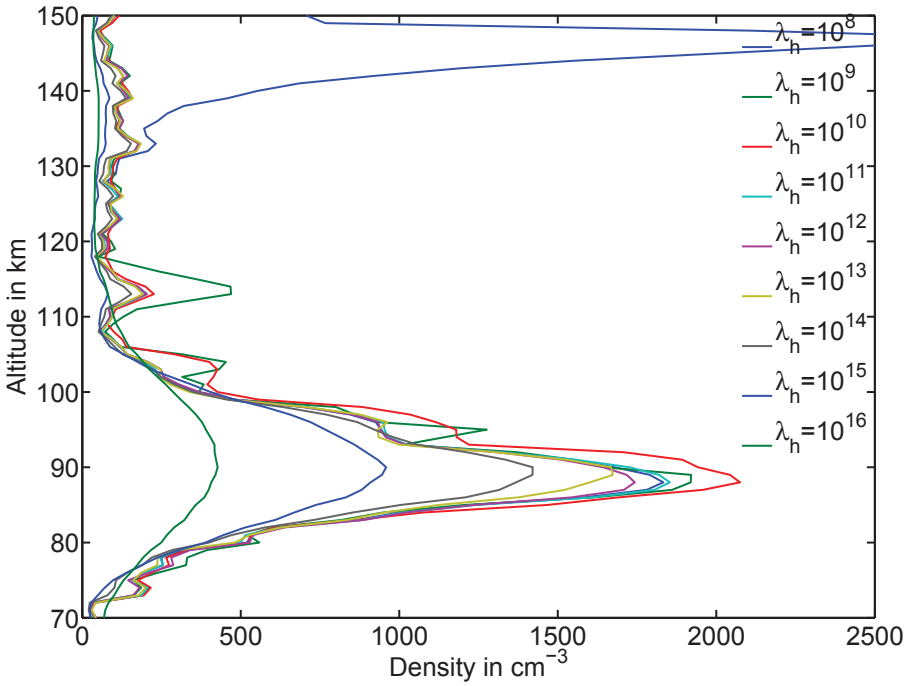

Fig. 6.31: Retrieval results with a 1-D model when an error is applied for Mg with the same method as for the blue line in Fig. 6.23. The constraint strength is varied. Oscillating results start at a much higher constraint strength than for the error-free case shown in Fig. 6.30. Furthermore, the result varies much stronger with changing constraint strength, so that it is difficult to find the constraint strength for the optimal solution. For the case of very strong smoothing, the same profiles as in the error-free case are retrieved for the same constraint strength.

the SCIAMACHY spectrum, a solar irradiance of $5.44 \cdot 10^{14} \frac{\text{photon}}{\text{s cm}^2 \text{ nm}}$ is used for the edge of the line. For the D_2 line the following values are used:

$$\begin{aligned} I_0 &= 0.0444 \cdot 5.44 \cdot 10^{14} \frac{\text{photon}}{\text{s cm}^2 \text{ nm}} \\ A &= 2.16 \\ x_e &= \frac{\boldsymbol{\sigma}}{k_{\text{linecenter}}} = \frac{0.228}{16973} = 13.4 \cdot 10^{-6} \\ x &= \frac{k - (k_{\text{linecenter}} - \text{shifts})}{k_{\text{linecenter}}} = \frac{k - (k_{\text{linecenter}})}{k_{\text{linecenter}}} + \frac{\text{shifts}}{k_{\text{linecenter}}}. \end{aligned}$$

The formula is given for wavenumbers k and the parameters x and x_e are normalized to the wavenumber of the line center. $\boldsymbol{\sigma}$ is the width parameter of the line in wavenumbers (see McNutt and Mack, 1963). A positive value for shifts leads to a red shift as the linecenter is moved toward shorter wavenumbers and higher wavelength. The constant red shift is: $\frac{\text{grav}_{\text{shift}}}{k_{\text{linecenter}}} = 2.7 \cdot 10^{-6}$. Consequently the Doppler shift needs the negative velocity, to point into the right direction $\frac{\text{Doppler}}{k_{\text{linecenter}}} = -\frac{k_{\text{linecenter}} \frac{v}{c}}{k_{\text{linecenter}}} = -\frac{v}{c}$. The most extreme value for the Doppler shift is $\pm 3.2 \cdot 10^{-6}$. For the D_1 line the

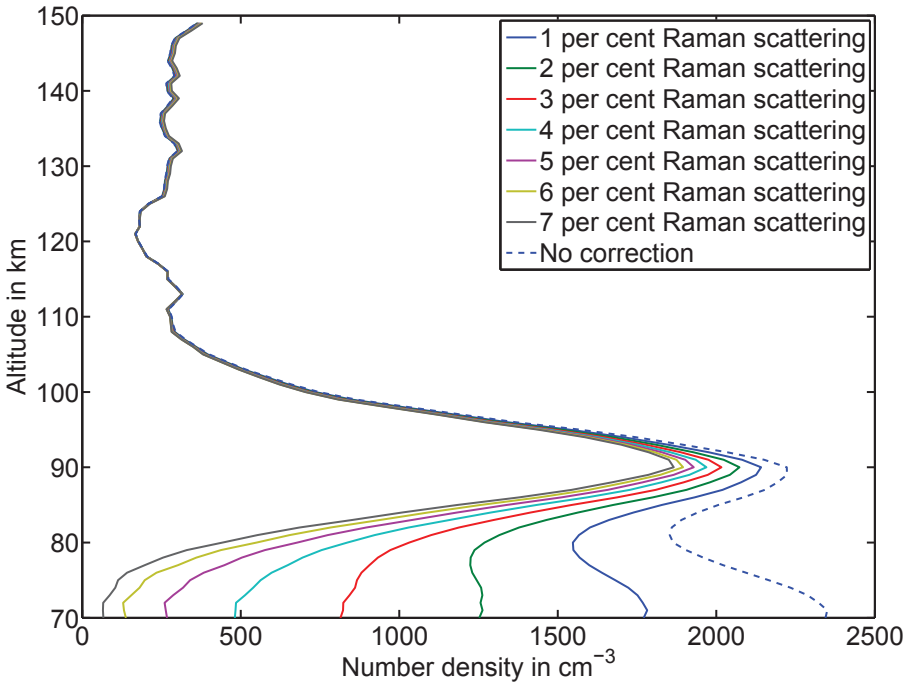

Fig. 6.32: Plot of Mg average profile of all single day results of the 2D retrieval in the equatorial region (15° S to 15° N) for different percentages of Raman scattering in the Ring effect correction. The different corrections mainly affect the peak density and the lower peak edge. The peak altitude is nearly unaffected. Differences for variations of 1 % in the Raman scattering percentage vary the lower peak edge altitude by ≈ 1 km for altitudes above 80 km. For altitudes below 80 km, the differences drastically rise.

following parameters are used:

$$\begin{aligned} I_0 &= 0.0495 \cdot 5.44 \cdot 10^{14} \, \frac{\text{photon}}{\text{s cm}^2 \, \text{nm}} \\ A &= 2.14 \\ x_e &= 12.8 \cdot 10^{-6}. \end{aligned}$$

As there are hyper-fine splittings for both D lines, the line center is the weighted average of the individual line wavelengths and strengths.

6.5.2 Shifts between Sun and Earth's mesosphere

The following shifts are listed in McNutt and Mack (1963): Doppler shifts due to the orbit of the Earth around Sun, as well as the Earth's rotation and Doppler shifts from the rotation of the Sun. Furthermore, there are other shifts, which are always biased to red shifts like gravitational red shift and Lindholm shifts and shifts due to radial convective currents on the Sun. The gravitational red shift comes from the fact that the gravitational field at the Sun's surface is much higher than in the mesosphere. From Einstein's Special Relativity and the principle of equivalence from

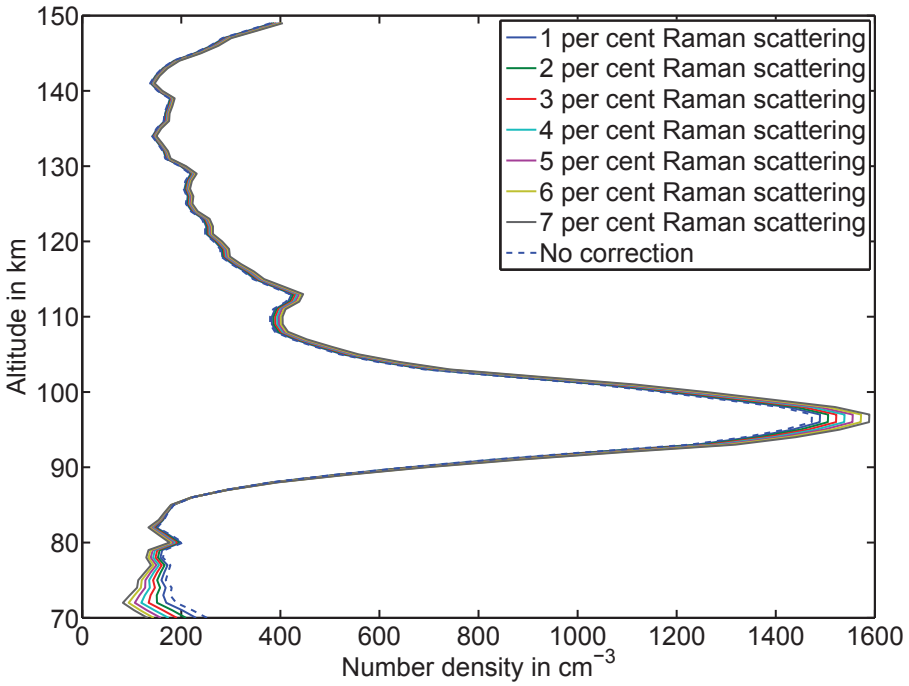

Fig. 6.33: Plot of Mg^+ (279.6 nm) average profile of all single day results of the 2D retrieval in the equatorial region (15° S to 15° N) for different percentages of Raman scattering in the Ring effect correction. For altitudes above 80 km the Ring effect correction is negligible and it is very small above 70 km. In the peak region the emission dominates the background signal. Because of this the peak value rises with higher Raman scattering percentage (as discussed in Sect. 6.2.2).

the theory of General Relativity (see, e.g., Einstein, 1911) one can derive a shift of

$$z = \frac{\lambda_{\text{observed}} - \lambda}{\lambda} = \sqrt{\frac{1}{1 - \frac{2GM_{\text{Sun}}}{r_{\text{Sun}}c^2}}} - 1 = 2.1 \times 10^{-6}, \tag{6.38}$$

with the gravitational constant G, the radius of the Sun r_{Sun}, the mass of the Sun M_{Sun} and the speed of light c. The so called Lindholm shifts (see, e.g., Lindholm, 1942) are shifts caused by high pressure. In the high pressure gas of the Sun, neighboring helium atoms are so close to the Na atoms, that the energy levels are distorted, which leads to a shift of the spectral lines. Furthermore, the high pressure broadens the line stronger than Doppler broadening. The total measured red shift estimated by McNutt and Mack (1963) is $(2.7 \pm 0.4) \times 10^{-6}$. The Doppler shift due to solar rotation can be very high, especially in the limb region of the solar disk (see, e.g., Marmet, 1989). However, it is assumed, that it cancels out over the whole disk and just leads to a broadening of the line.

The Doppler shift between the Earth and the Sun lies in in the pm range. The Earth rotates around its axis and performs nearly a circular orbit around the Sun. The Earth axis is slightly tilted to its orbit movement. Furthermore, as the Earth orbit around the Sun is an ellipse, the distance from Earth to Sun also changes. The

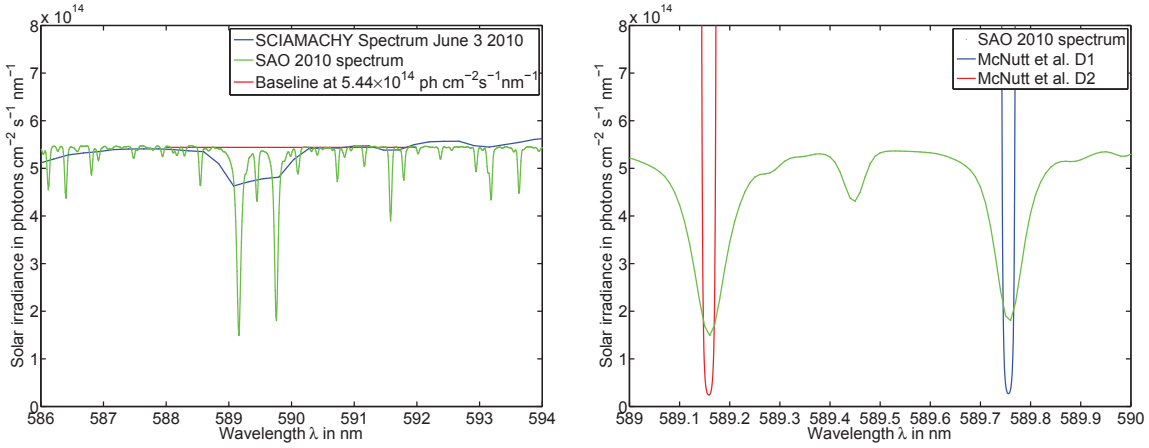

Fig. 6.34: Spectrum of the Na Fraunhofer lines. The left figure shows the SCIAMACHY solar irradiance spectrum, as well as the higher resolved SAO 2010 spectrum from Chance and Kurucz (2010) and the baseline at $5.44 \cdot 10^{14} \frac{\text{photon}}{\text{s cm}^2 \text{ nm}}$. The right figure shows the SAO 2010 spectrum compared to the fully resolved approximation of the D_1 and D_2 line from McNutt and Mack (1963). Note, that this approximation is only valid close to the center of the Fraunhofer lines. The Fully resolved lines are much deeper than the lines in the SAO 2010 spectrum and the SCIAMACHY spectrum.

rotational speed perpendicular to the Earth axis is given by:

$$v_{rot_{\text{Earth}}} = r\Omega \tag{6.39}$$

The rotation time is 24 h (respectively 23 h 56 min, when taking into account, the orbit of Earth around the sun), and the perpendicular radius is latitude dependent. The interesting region is roughly at 90 km altitude, so $R_e = 6370 + 90 = 6460$ km is a good approximation. The rotational velocity at this altitude is:

$$v_{rot_{\text{Earth}}} = \frac{2\pi R_e}{24h} cos(\theta_{\text{Earth}}), \tag{6.40}$$

with θ_{Earth} being the latitude on Earth at the tangent point. As the Earth axis is tilted towards the normal of the ecliptic, the projected component is needed, that is derived with the cosine of the declination. This is the maximum that can appear, when the velocity points towards the Sun. The minimum component appears, when the connection axis between Sun and Earth is perpendicular to the velocity. In this case, the Sun stands in zenith (local at 12 o'clock). The maximum appears, when the velocity points into the Suns direction or opposite (6 and 18 o'clock). At 6 o'clock velocity points to the Sun and at 18 o'clock opposite. So the final equation is:

$$v_{\text{Doppler}} = v_{rot_{\text{Earth}}} cos(\theta_{\text{Sun}}) sin(\phi_{\text{Sun}} - \phi_{\text{TP}}), \tag{6.41}$$

with the latitudinal and longitudinal projection of Sun to Earth θ_{Sun} and ϕ_{Sun}, as well as the longitude of the tangent point ϕ_{TP}. The projected solar coordinates are calculated using the formulas from Blanco-Muriel et al. (2001).

Another velocity component of the same magnitude appears due to the different Sun-Earth-distance, during the year between aphel and perihel. There is a difference of $5 \cdot 10^9$ m between both distances. Assuming a harmonic behavior of the annual cycle the distance between Earth and Sun r can be approximated:

$$r = r_{mid} + \frac{\Delta r}{2} cos(\omega t) \tag{6.42}$$

To calculate the velocity, r has to be derived for time:

$$v = -\frac{\Delta r}{2} \omega sin(\omega t) \tag{6.43}$$

The velocity amplitude of 498 $\frac{\mathrm{m}}{\mathrm{s}}$ for this shift is comparable to the velocity amplitude of 470 $\frac{m}{s}$ for the shift due to rotations.

For the signs of the Doppler shifts, the following considerations are made. For wavenumbers, the Doppler shift is given by:

$$k = k_0(1 + \frac{v}{c}) \tag{6.44}$$

Where k is the observed wavenumber (in the mesosphere) and k_0 is the wavenumber of the signal (from the Sun). If Earth moves towards the Sun, then $\frac{v}{c}$ is positive and the wavenumber and frequency increases. As wavelength $\lambda = \frac{1}{k}$, it will decrease, when k increases (blue shift). Consequently, if the movement is away from the Sun, the light is red shifted to bigger wavelengths. In this definition a positive v means a movement towards the Sun and a blue shift. For the rotation of the Earth the formula is:

$$v_{\mathrm{Doppler}} = v_{rot_{\mathrm{Earth}}} cos(\theta_{\mathrm{Sun}}) sin(\phi_{\mathrm{Sun}} - \phi_{TP}) \tag{6.45}$$

The sign is determined by the sine. Longitudes are assumed to be in degrees East (Sun moves into negative direction in longitudes). If the tangent point is at 0 degrees in Europe, and the Sun is at 90 degrees (China) the velocity points towards the Sun. So $sin(\phi_{\mathrm{Sun}} - \phi_{\mathrm{TP}}) = 1$ and the sign of this velocity is already as it should be. Considering the motion of Earth on its elliptic orbit, Earth is closest to the Sun in winter (perihel around 4.1. and in aphel in July). This means, that from January to July Earth moves away from Sun and from July to January the Earth moves towards the Sun. The first half of the year, there is a red shift, and the second half, there is a blue shift. Consequently (with a small shift of a few days) if $t = 0$ at the beginning of the year $-sin$ gives the right sign.

6.5.3 Hyper-fine structure of the Na D-lines

Na has only one stable isotope $^{23}_{11}$Na and therefore has no isotope effect. The stable isotope has a nuclear angular momentum of $I = \frac{3}{2}$, which leads to a hyper-fine splitting of the energy levels. The splitting for the lower $3^2S_{\frac{1}{2}}$ state is stronger than the splitting of the upper states $3^2P_{\frac{1}{2}}$ and $3^2P_{\frac{3}{2}}$. This can be explained phenomenologically by the lower distance of the valence electron in the S state, which leads to a larger overlap of the nuclear and electron wave functions and therefore a stronger perturbation of the electronic state. Due to the stronger splitting of the S state,

compared to the P states, the D_1 and D_2 lines each split in 2 groups of narrow lying lines. The Doppler width of the Na lines in the mesosphere is approximately 1.2 pm. The 2 groups of narrow lying lines have a separation of about 2 pm and thus have a very small overlap and can be in principle (not in SCIAMACHY resolution) separated. However, the lines within a group are too narrow to be resolved at mesospheric temperatures (respectively temperatures larger than 4 K). As only the separation of the S state can be resolved, the resulting lines are called s-resolved lines. The existence of several degenerated lines, however, is important for correct weighting of the 2 s-resolved lines. This is well explained, e.g., by Chamberlain et al. (1958), McNutt and Mack (1963) and Fricke and von Zahn (1985).

The solar spectra with the shifts discussed in Sect. 6.5.2, and the mesospheric absorption cross section profile of the Na D_2 are shown in Fig. 6.35. The calculation

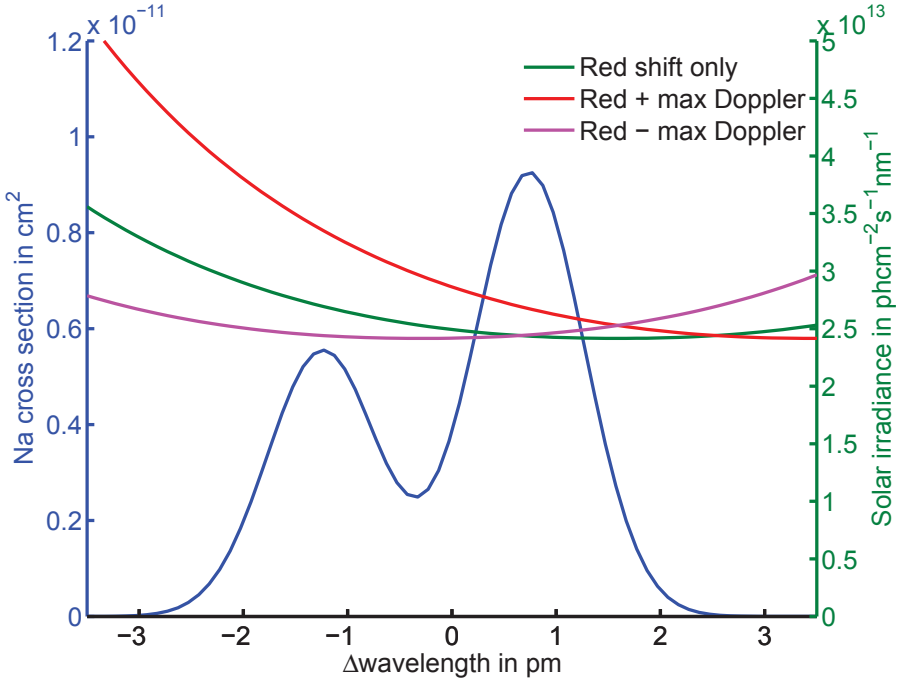

Fig. 6.35: Absorption cross section for the D_2 line at mesospheric temperatures (in blue) and solar irradiance spectrum for different spectral shifts.

of the top of atmosphere emissivity (γ-factor without absorption) is similar to the one for Mg/Mg^+ in Sect. 6.2.3, beside a small change in the phase function (Eq. (6.8)), which will be discussed in the Sect. 6.5.6. The factors E_1 and E_2 for Na are listed in table 6.5. The calculation of the factor f for the self absorption is the same as in Sect. 6.3.4 (Eq. (6.35)), just with the different solar and cross section spectra.

6.5.4 Multiple scattering and albedo factor – first approaches

The only real difference between the retrieval of Na and Mg/Mg^+ is, that the single scattering approximation is only valid for Mg/Mg^+. However, in spite of using a full

multiple scattering radiative transfer model, the single scattering approach is extended by an estimation of the increase of the solar irradiation, which passes a grid cell more than once due to backscattering in the lower atmosphere and surface reflection in the lower atmosphere. In the following, this conversion factor is called albedo factor. Different methods for the estimation of this factor have been reported by Gumbel et al. (2007), and Hedin and Gumbel (2011). In Hedin and Gumbel (2011) the background signal for the limb scan at 40 km tangent altitude is compared to a single scattering radiative transfer model for Rayleigh scattering. However, the lowest tangent altitude of the SCIAMACHY limb MLT measurements at 53 km is too high and straylight effects would lead to an overestimation of the albedo factor, which is illustrated in Fig. 6.36. In contrast to the OSIRIS measurements used

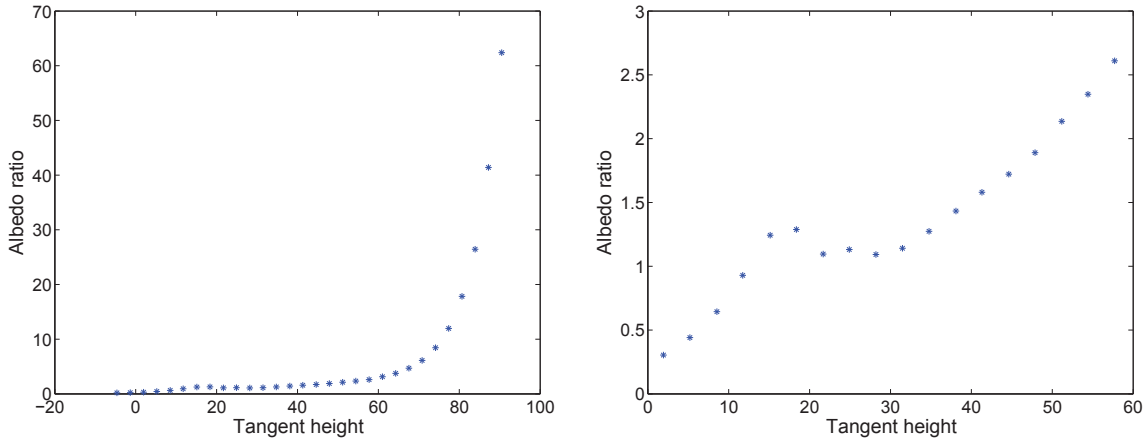

Fig. 6.36: Measured background signal of a nominal SCIAMACHY limb measurement in the vicinity of the Na D lines divided by the simulated Rayleigh single scattered background spectrum. For low altitudes this ratio is influenced by clouds. There is a plateau between 20 km and 35 km altitude, which could be used as the albedo factor. For higher altitudes the ratio rises strongly, as straylight from the lower atmosphere dominates the background signal.

in Hedin and Gumbel (2011), SCIAMACHY has the advantage, that the D_1 and D_2 lines are resolved and densities can be separately retrieved from both lines. As the more strongly emitting and reabsorbing D_2 line varies much more strongly with changes in the albedo factor than the D_1 line, the albedo factor can be estimated as the factor where both Na density profiles retrieved from the D_1 and the D_2 match best. As an example how this approach works, Fig. 6.37 shows the SCD profiles of D_1 and D_2, where the D_2 line has the lower SCD, because of the stronger self absorption, and Fig. 6.38 shows the retrieved density profiles. For too small albedo factors the SCDs are overestimated and the D_2 densities are larger than the D_1 densities. The opposite happens for too large albedo factors. For the optimal albedo factor both profiles match. As discussed in Sect. 6.4.2 the constraints applied in the retrieval may affect the retrieved profiles quite strongly and small changes in the constraint strength may lead to relatively high changes in the retrieved profile, when the statistical errors of the slant column densities are high, so that it is complicated

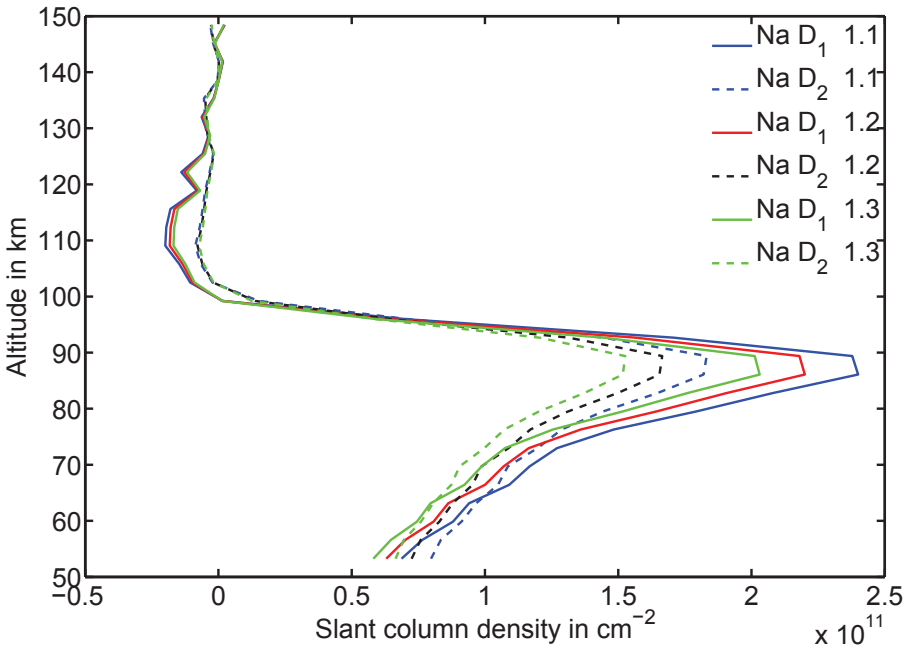

Fig. 6.37: Vertical SCD profile for both Na D lines (October 6, 2009, 35°N) for 3 different albedo factors ($1.1 - 1.3$). The SCDs of the D_2 line are smaller, because of stronger self absorption.

to find the optimal solution. Fortunately, the statistical errors for the visible region at around 600 nm are much smaller than at around 300 nm and Fig. 6.39 shows that the Na retrieval is insensitive enough on the constraint strength, that the D_1, D_2 matching method for the albedo factor should work. Due to the multiply scattered contribution of the radiation, not just the effective incoming solar irradiation is increased, but also the phase function of the incoming radiation is changed. Assuming the multiply scattered contribution of the radiation to be fully isotropic, the phase function changes to $P_{\text{new}} = \frac{P_{\text{old}} + (\text{albedo factor} - 1)}{\text{albedo factor}}$.

In practice, however, this method failed quite often (e.g., for cases, where the differences of the D_1 and the D_2 SCDs were too high to yield reasonable albedo factors), which implies, that probably even harder criteria for the smallness of statistical and systematical error are necessary for this approach to work robustly. Therefore, a different method was used. Nevertheless, the match of the density retrieved from both D lines is a good indicator of how well the retrieval algorithm works.

Simulations of the multiple scattering to single scattering ratio with SCIATRAN (see, e.g., Rozanov et al., 2014) with different ground albedos and for different scattering angles show albedo factors of 1.5 for ground albedos between 0.3 and 0.4, which is a realistic value for the average planetary albedo (including clouds to be part of the surface, which is a reasonable approximation at 90 km altitude for clouds below 15 km). The simulation results are shown in Fig. 6.40 and show a dependence of the albedo factor on the scattering angle. For the retrieval, a linear approximation of the dependence to the sine of the scattering angle is used, and the albedo factor is calculated as $1 + 0.5 \sin(\text{scat. ang.})$. At high latitudes, the ground albedo, which is

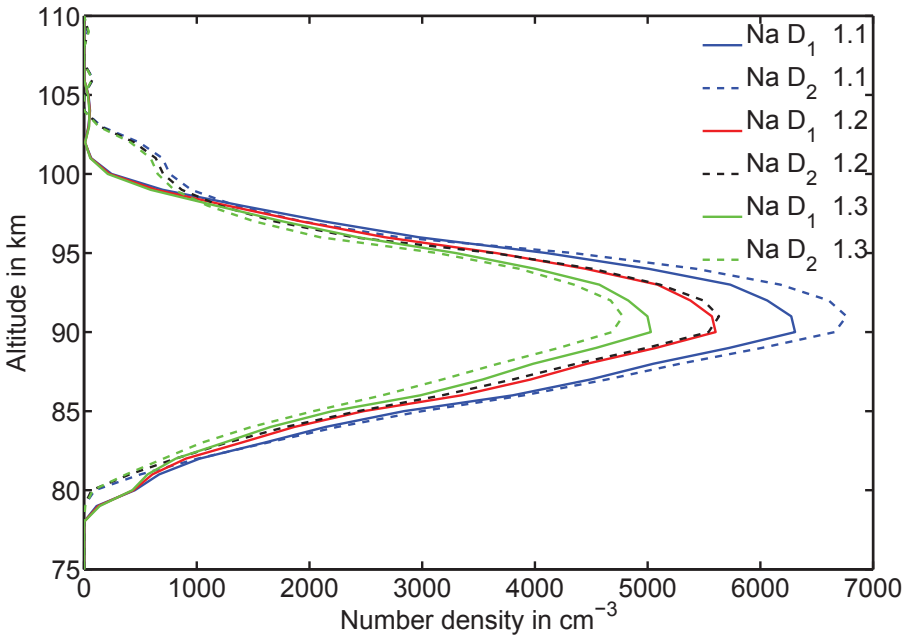

Fig. 6.38: Vertical Na number density profile retrieved from the SCD profiles shown in Fig. 6.37. For too large albedo factors the densities are smaller and the D_1 line shows higher densities. For too small albedo factor the D_2 line shows higher densities. For the optimum albedo factor (here 1.2) both Na lines yield the same densities.

similar to the ground albedo retrieved from MERIS data at 560 nm or 620 nm (see, e.g., Popp et al., 2011), can be higher than 0.8 in this wavelength region. However, the sine of the scattering angle is small in these regions, and there are only small differences for different ground albedos for these conditions (see Fig. 6.40).

For the estimation of how well this method works, Fig. 6.41 shows the absolute and relative differences in the VCDs retrieved from both lines. Beside for the southernmost latitudes in southern hemispheric winter, the match of the results from both lines is quite good. As will be shown in Sect. 8, the densities are high in this region. For high densities the self absorption correction is strong, and small inaccuracies in the radiative transfer model (e.g., the just parameterized multiple scattering part) may lead to larger discrepancies. When using the average of both D lines as the final result, the largest discrepancies to the result of one of the individual lines are smaller than 20%. The statistical error for both lines, shown in Fig. 6.42, which is determined with the same method as for Mg and Mg^+, discussed in Sect. 6.4.1, is $\approx$ 10% in the maximum number density region, and smaller than the mean for altitudes below 80 km and above 100 km.

6.5.5 Multiple scattering and albedo factor – finally used approach

The simple geometric approach of calculating the albedo factor as 1+0.5 sin(scat. ang.) was based on the assumption of a globally constant ground albedo of 0.3 to 0.4. As this assumption appeared to be too strong, another method was developed to calculate the albedo factors from measurement data, which in the end only led to rather

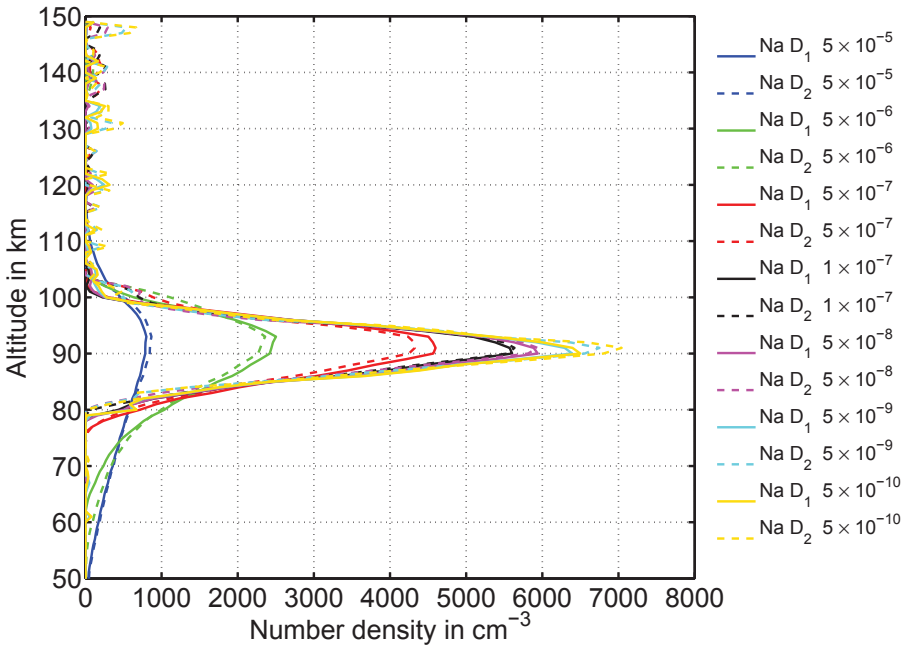

Fig. 6.39: Vertical Na density profile for the D_1 and D_2 for different constraints strengths (Same conditions as in Fig. 6.37, albedo factor 1.2, see legend for constraint strength). For moderate constraint strength ($1\times10^{-7} - 5\times10^{-10}$) the retrieved peak density is nearly independent of the choice of the constraint strength for ≈2 orders of magnitude. However, a factor 5 to 10 in the constraint strength has a similar effect to a change of the albedo factor of 0.1 in Fig. 6.38. The 3 highest constraints strength ($5\times10^{-7} - 5\times10^{-5}$) show too strong smoothing while the lowest constraint strength ($5 \times 10^{-9} - 5 \times 10^{-10}$) shows oscillations at high altitudes. Note, that a stronger smoothing leads to the need of a lower albedo factor to match the D_1 and D_2 density, so a systematic error in one property is rather reduced, than increased, by the tuning of the other one, which results in some robustness in the method.

small changes of up to around 20%, but leads to more exact results than the simple geometric approach. As mentioned in the previous section, the SCIAMACHY limb MLT measurements range from 53 to 150 km tangent altitude, so that the approach used by Hedin and Gumbel (2011), who used the ratio of the limb measurement signal to the simulated Rayleigh scattered measurement signal in the vicinity of the Na lines at 40 km as the albedo factor, cannot directly be used. However, assuming that the vertical profile of this ratio has roughly the same shape for the same latitude and longitude on different days, fitting the MLT measurements profile to this vertical shape profile from the surrounding days enables the determination of the albedo factor for the MLT measurements.

Instead of using a relatively simple radiative transfer model as used for Fig. 6.36, SCIATRAN was used to perform the single Rayleigh scattering calculations. As the dark current signal at 350 km appeared to be too strongly contaminated by the straylight to simply subtract it from the limb signals at the 30 tangent altitudes of a given MLT limb scan, the signal to subtract (b, see below) was obtained by another

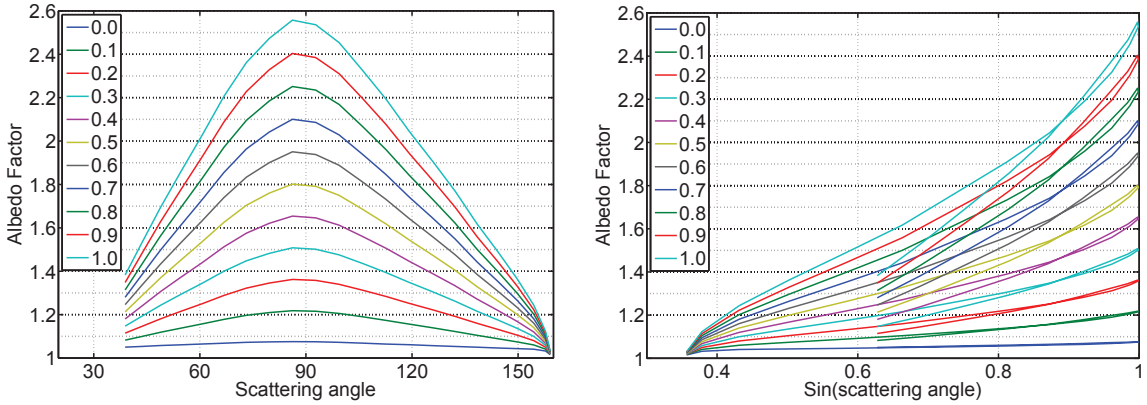

Fig. 6.40: Ratio of the multiple scattering signal (including the single scattering signal) divided by the single scattering signal for different ground albedos (see legend) and realistic SCIAMACHY observation geometries. The simulations were done with SCIATRAN. The albedo factor is highest at a scattering angle of 90°. For strong forward or backward scattering the albedo factor is smaller and differences for different ground albedos are small. The forward scattering shows the steeper descend. The single scattering angle is typically in the range of 30° (high northern latitudes) to 150° (high southern latitudes) (The data for the graphics were provided by Patricia Liebing).

method. We assume that a part of the radiance at each tangent altitude h and wavelength λ has a multiplicative part $aI_{ss}(h,\lambda)$, that is proportional to the simulated single scattering signal $I_{ss}(h,\lambda)$. We further call the proportionality factor a the multiplicative component. Furthermore, there is the actual wavelength independent dark current signal b(h), which we call the additive component. The total incoming radiation $I_{inc}(h,\lambda)$ is the sum of both parts $I_{inc}(h,\lambda) = a(h)I_{ss}(h,\lambda) + b(h)$. The components a and b have to be fitted in a wavelength region with a non spectral constant solar spectrum, so that the roles of a and b cannot be interchanged, i.e., the problem is not ill posed. Furthermore, there must not be any Earth atmospheric features (emission, extinction etc.) in this region. The nearby wavelength interval between 650 nm and 660 nm includes the H α Fraunhofer line at 656 nm and is well suited for this fit. We assume the minimum of the multiplicative component a above 20 km tangent altitude to be very close to the searched albedo factor. Figure 6.43 shows the fit of a and b as well as the fitted vertical profile for an example nominal SCIAMACHY limb measurement.

The albedo factor can be well determined with this method for the nominal SCIAMACHY limb measurements, which scan from ground to 90 km altitude. The MLT measurements, however, scan from 53 to 150 km, so that the altitude region, where the albedo factor can be retrieved from, is not included. However, the latitudinally and longitudinally co-located nominal limb measurements from the neighboring days (SCIAMACHY is in a sun-synchronous orbit, and every latitude-longitude-region is only scanned once per day) show very similar profile shapes, which can be fitted to the MLT-data at altitudes, where both overlap. Figure 6.44

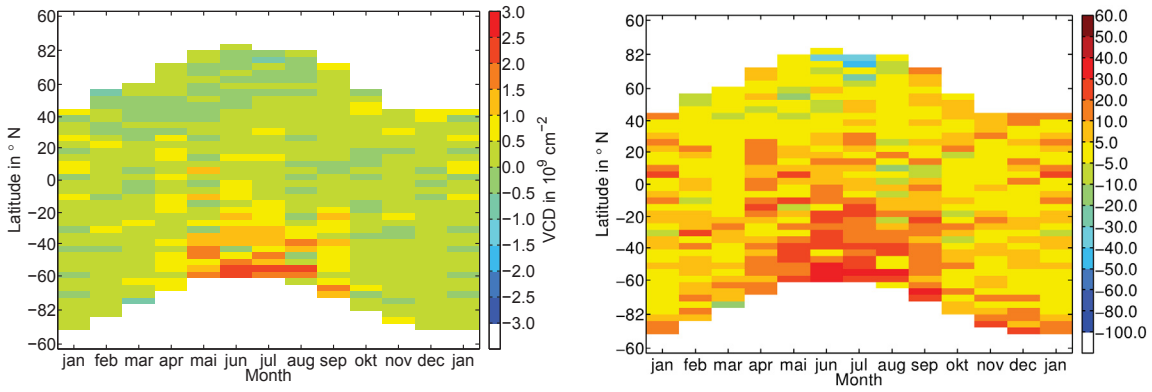

Fig. 6.41: Absolute (left, D_1-D_2) and relative (right, $2\frac{D_1-D_2}{D_1+D_2}$) difference of the Na VCDs retrieved from the D_1 and D_2 lines. For the method that is finally not used.

shows the fit that was finally used to retrieve the albedo factors for the MLT measurements. The median of the multiplicative component in the latitudinally and longitudinally co-located region for the days in the same time period ($\pm$ 200 orbits) of nominal limb measurements is formed for all altitudes. The albedo factor A for this median profile is determined. Between 50 km and 70 km the logarithms of the nominal and the MLT measurements are fitted as factor B (ln MLT = B ln nominal). Fitting the factor B this way puts more weight on the match at the lower altitudes, where it is assumed, that the match should be better, as the spectral straylight error is smaller there. The albedo factor for the MLT measurement finally is the product AB.

6.5.6 Polarization correction for the metal emission lines

The standard polarization correction for SCIAMACHY is described in Slijkhuis (2008). This correction is done assuming Rayleigh scattered radiation. However, the metal emission lines need a different polarization correction, because their degree of polarisation P is smaller or equal to the P that is used for Rayleigh scattering. Furthermore, this correction is different for each individual metal emission line, which can be exploited to find correction parameters for wavelength regions close to the metal emission lines. As the metal's emission lines correction is still similar to the correction for Rayleigh scattering (in fact the Mg line even needs the same correction as Rayleigh-scattered radiation), the relevant formulas from Slijkhuis (2008) are recapitulated with the same nomenclature. Note, that this leads to the double use of the symbols Q, U and V, which are also used as $\frac{Q}{I}$, $\frac{U}{I}$ and $\frac{V}{I}$, which are explained in the following paragraph.

Polarized light, considering a reference plane, is described by the so called Stokes-vector (I, Q, U, V). We denote P_ϕ as the intensity component linearly polarized in an angle P_ϕ to the incident plane with P_{0° being the horizontal and P_{90° the vertical component. Q and U are defined as: $Q = P_{0^\circ} - P_{90^\circ}$ and $U = P_{45^\circ} - P_{135^\circ}$. V is the difference between left and right circular polarized intensity. I is the total intensity, which can be formed from the sum of the perpendicular components, i.e.,

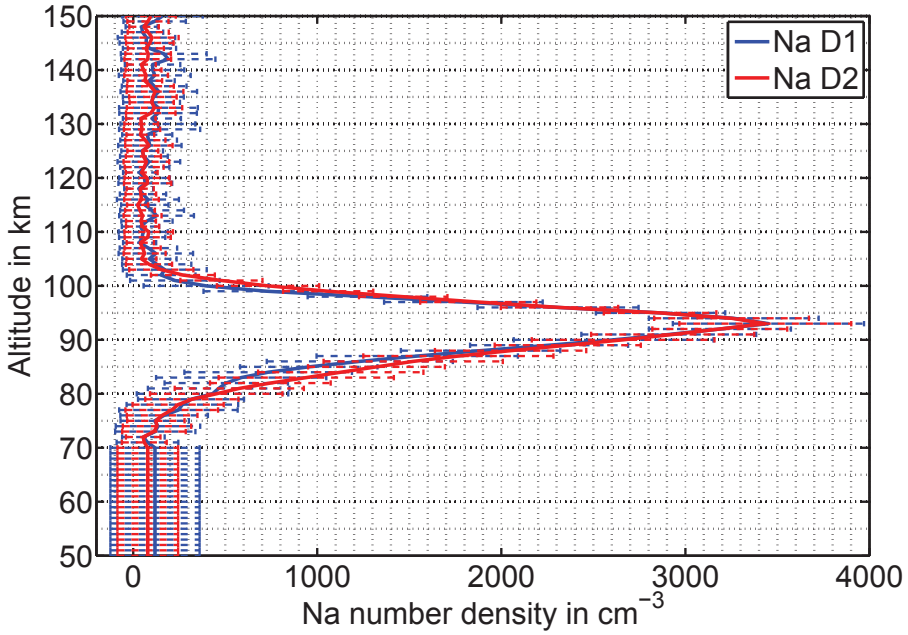

Fig. 6.42: Mean and standard deviation (error bars) of the Na D_1 and D_2 line retrieval for the equatorial SCIAMACHY measurements on March 20, 2010.

$I = P_{0^\circ} + P_{90^\circ} = P_{45^\circ} + P_{135^\circ} = P_{\text{left}} + P_{\text{right}}$. The normalized Stokes-vector is the Stokes-vector divided by I, $(1, Q/I, U/I, V/I)$, which in the following will be substituted by $(1, Q, U, V)$. The incoming solar radiation is assumed to be unpolarized. If the incoming light is not already circularly polarized, this does not change due to resonance fluorescence (see, e.g., Chandrasekhar, 1960), so that the treatment of the V component is furthermore omitted.

For the remaining components Q and U, the reference frame independent degree of polarization P is defined as:

$$P = \sqrt{Q^2 + U^2} \tag{6.46}$$

and the polarization angle χ is:

$$\chi = \frac{1}{2} atan\left(\frac{U}{Q}\right) \tag{6.47}$$

The factor $\frac{1}{2}$ is due to the fact, that the polarization plane in limited to an angle of 180° and not 360°. Knowing this, it is sufficient to calculate P in the scattering plane and to know how χ is transformed from the scattering plane to the instrumental entry plane of SCIAMACHY. This information can be found in Slijkhuis (2008), Eq. (86):

$$\chi = acos\left(\pm\frac{sin\theta_0 sin(\phi_0 - \phi)}{\sqrt{1 - cos^2(\Theta)}}\right) \tag{6.48}$$

With + for $\theta_0 < 90^\circ$, θ_0 being the solar zenith angle, $\phi_0 - \phi$ being the relative azimuth angle and Θ being the scattering angle, which are all available in the SCIAMACHY

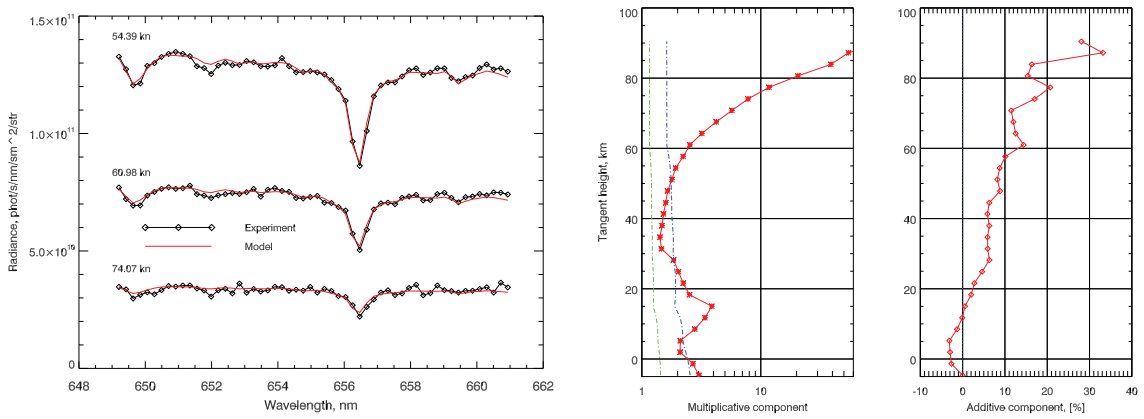

Fig. 6.43: left: Fit of the multiplicative and additive component in the vicinity of the H α Fraunhofer line. right: Result for the multiplicative and additive components (very similar to Fig. 6.36; the graphics were provided by Vladimir Rozanov).

data product. The sign in Eq. (6.48) affects the sign of U, but not of Q. Knowing P and χ, the components Q and U for the Stokes vector measured by the instrument can be calculated as follows:

$$Q = Pcos(2\chi) \tag{6.49}$$

$$U = Psin(2\chi) \tag{6.50}$$

The polarization of the resonance fluorescence process of metals is well described by Hamilton (1947) and can also be found in Chandrasekhar (1960) with a slightly different nomenclature in the formulas used. The incoming solar radiation is assumed to be unpolarized, which means, that the parallel and perpendicular intensity components are equal in any observation plane. Following Chandrasekhar (1960) (Eq. (259), p. 51), the incoming radiation components parallel and perpendicular to the scattering plane are transformed by resonance fluorescence as follows:

$$\begin{pmatrix} I_{\parallel out} \\ I_{\perp out} \end{pmatrix} = \left(\frac{3}{2} E_1 \begin{pmatrix} cos^2\theta & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} E_2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right) \begin{pmatrix} I_{\parallel in} \\ I_{\perp in} \end{pmatrix} \tag{6.51}$$

with the degree of Polarization $P = \sqrt{Q^2} = \sqrt{\left(\frac{I_{\parallel \mathrm{out}} - I_{\perp \mathrm{out}}}{I_{\parallel \mathrm{out}} + I_{\perp \mathrm{out}}} \right)^2}$ and the polarization direction is perpendicular to the scattering plane (The reference frame is chosen so, that $U = 0$). The radiative transfer through the instrument can be written in the form:

$$\begin{pmatrix} I_d \\ Q_d \\ U_d \end{pmatrix} = \begin{pmatrix} M_1 & M_2 & M_3 \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} I_0 \\ Q_0 \\ U_0 \end{pmatrix}, \tag{6.52}$$

with M_1, M_2 and M_3 are the elements of the first row of the transformation matrix. Only I_d (the subscript "d" stands for detected) is really measured by SCIAMACHY, so that the other matrix elements do not need to be considered. The actual intensity I_0 is obtained by the inversion of the equation

$$I_d = M_1 I_0 \left(1 + \frac{M_2}{M_1} \frac{Q_0}{I_0} + \frac{M_3}{M_1} \frac{U_0}{I_0} \right), \tag{6.53}$$

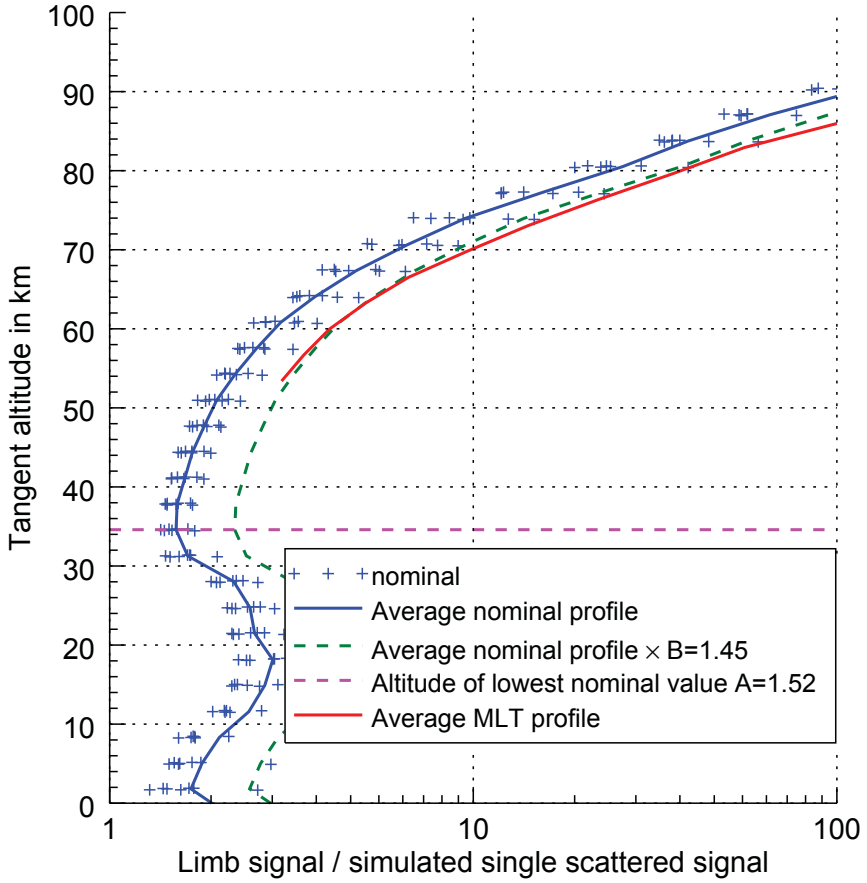

Fig. 6.44: Albedo factor fit for an example MLT-measurement. The albedo factor is the product AB ($AB = 2.20$). Note, that the MLT profile used here comes from a part of the orbit, which has additional solar straylight, and therefore the factor B is unusually large. These measurements are actually sorted out, however this shift better separates the lines in the figure, so that it is easier to read than in the normal case.

and is given by:

$$I_0 = \frac{I_d}{M_1}\left(1 + \frac{M_2}{M_1}\frac{Q_0}{I_0} + \frac{M_3}{M_1}\frac{U_0}{I_0}\right)^{-1}. \tag{6.54}$$

The component M_1 is determined at a different point in the calibration, (and is 1.0 in the key data), so the actual formula for the calculation, which differs a bit from Eq. (6.54), is:

$$I_{cal} = I_0 = \frac{I_d}{1 + \mu_2 Q + \mu_3 U} \tag{6.55}$$

The wavelength dependent factors μ_2 and μ_3 in Eq. (6.55) were measured for SCIAMACHY before launch. However, after launch, it was found out that those pre-flight factors are not valid anymore, and several methods have been used to evaluate those factors in space (see, e.g., Krijger et al., 2014; Liebing et al., 2013). It is possible to derive polarization sensitivities in the wavelength range below 300 nm

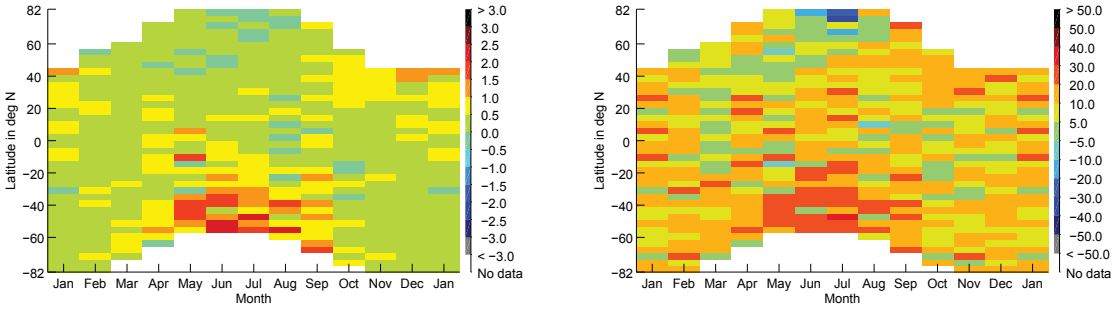

Fig. 6.45: Absolute (left, D_1-D_2) and relative (right, $2\frac{D_1-D_2}{D_1+D_2}$) difference of the Na VCDs retrieved from the D_1 and D_2 lines, with the finally used method. The differences are very similar to the ones in Fig. 6.41.

from in-flight data by comparing measured reflectances with the ones calculated by a vector radiative transfer model (SCIATRAN 3.1, see, e.g., Rozanov et al., 2014), and then fitting the differences to the expected polarization. The results of this fit ($\mu_2 = 0.16$ and $\mu_3 = -0.23$ for Mg at 285.2 nm, and $\mu_2 = 0.15$ and $\mu_3 = -0.21$ for Mg^+ at 279.6 nm) have been applied for Mg and Mg^+ together with Eq. (6.51) and a transformation to the SCIAMACHY polarization reference frame.

The remainder of this section is used to discuss an alternative calculation of the polarisation correction in the vicinity of the Mg/Mg^+ lines. This correction method is not used for the retrieval of Mg/Mg^+. However, as polarisation correction is one of the larger issues in the calibration of satellite-based spectrometers, every independent new approach to test the calibration or to perform the calibration is useful. One of the two Mg^+ lines, the D_1 line at 280.4 nm, does not emit polarized resonance fluorescence, while the other one does. By comparing the densities retrieved from both Mg^+ lines, the quality of the polarization correction can be reviewed. For a good correction the densities retrieved from both Mg^+ lines are more similar than without the correction. This consistency check is fulfilled for the correction actually applied.

Instead of using the known key data, another method for the calculation of the correction factor can be directly found by comparing the two Mg^+ lines. Assuming that the polarization effect is the only effect, that still needs to be corrected for a perfect calibration at the wavelength of the the Mg^+ lines, the total correction factor can be retrieved from the density matching of both Mg^+ lines. However, it is harder to retrieve μ_2 and μ_3 from the total correction factor, as the problem is ill posed, because $U \approx 2Q$ for most scattering angles. Assuming that the wavelength dependence of μ_2 and μ_3 is small, this correction factor could also be used to correct other nearby lines, like the Mg line. Substituting Q and U in Eq. (6.55) by Eq. (6.49) and Eq. (6.50) leads to:

$$I_{cal} = \frac{I_d}{1 + P(\underbrace{\mu_2 cos(2\chi) + \mu_3 sin(2\chi)}_{\Phi})}. \tag{6.56}$$

I_d is the measured intensity to be corrected, and P can be calculated using Eq. (6.51)

and Eq. (6.46). For the Mg^+ D_1 line at 280.4 nm, P is $P = 0$, and a polarization correction is not necessary. As the angles in Φ hardly change during one limb scan, a common Φ is assumed for the whole limb scan. Assuming this, the correction factor for the Mg^+ D_2 line is found by applying different correction factors to all consecutive limb measurements and doing a 1-D vertical retrieval for Mg^+ for the D_2 and the D_1 lines. The correction factor then is the factor, that matches the vertical column density retrieved from both lines best. For finding the optimal correction factor the bisection method is used. Using the vertical column density instead of the density at each altitude is more stable, as it suppresses noise effects in the data, especially in regions with small density, e.g., above and below the peak altitude. For the Mg line P is $P_{Mg} = 2P_{D_2}$. Therefore, knowing the correction factor $\mathrm{F}_{D_2} = \frac{1}{1+P_{D_2}\Phi}$ for the D_2 line, the correction factor for the Mg line is calculated as $\mathrm{F}_{Mg} = \frac{1}{1+2P_{D_2}\Phi}$. Unfortunately, the correction factors for Mg, determined this way, often did not lead to a convergence of the Mg retrieval, so that this kind of correction was forfeit in favour of the above mentioned correction method (which uses μ_2 and μ_3 retrieved by other methods). However, for a better signal to noise ratio in this wavelength region, this method has the potential to be a good alternative/addition to the other correction method.

Part III
Results

7 Results of the Mg and Mg^+ number density retrieval

In this section, monthly mean results for the 4-year MLT limb data set for Mg and Mg^+ are presented. The results are compared to the available observations of Mg and Mg^+ in the literature as well as with theoretical results from the WACCM-Mg model. WACCM is the abbreviation of the Whole Atmosphere Climate Community Model, and the extension Mg stands for the inclusion of the Mg atom and ion chemistry pattern into this model. The results have been published in Langowski et al. (2015b). The text of Langowski et al. (2015b) was mainly written by myself and will be re-used in this section, as this is the best formulation of the text and any reformulation may only reduce the comprehensibility of the text without adding new details.

7.1 Averaging of the spectral data

The signal to noise ratio of the SCIAMACHY observations in the wavelength region of the Mg line at 285 nm and the Mg^+ lines at 280 nm is on the order of 1 for single MLT-measurements above 90 km tangent altitude. As pointed out in Sect. 6.4.1 and illustrated in Fig. 6.26, errors in this magnitude range lead to a systematic bias towards larger densities due to the non linear forward model and non linear error propagation. Therefore, it is necessary to average the spectral data before the retrieval process. However, averaging before the retrieval does not only reduce the noise of the signal, but also the influence of true natural variability, that is non linear propagated. Therefore, a compromise in the averaging has to be found. The data averaging is done in two steps. In the first step, daily zonal mean spectra are formed from the up to 15 single orbits of one day of SCIAMACHY MLT observations and density profiles are retrieved. In the second step, the daily zonal averages for each month of the 4 year period of SCIAMACHY MLT observations are averaged to monthly zonal means, which cumulates to 4 to 8 single day averages forming the multi-annual monthly averages.

Before the data is averaged, it is checked for unsuitable spectra, which are excluded. The excluded measurements are either night measurements with a too low signal or measurements contaminated by extraordinary spikes in the spectra, which are mostly caused by highly energetic particles hitting the detector material. This happens typically in the Southern Atlantic Anomaly (SAA) region. Highly energetic particles typically hit several detector pixels at once, so that these events can be easily identified, e.g., by measuring the average residuum signal after the fit of the spectral lines. Fig. 7.1 shows a plot of the residuum for an orbit affected by the SAA (left) and not affected by the SAA (right). Measurements, that should be excluded, are sorted out through a threshold criterion. There are a few measurements outside the SAA, which also need to be excluded through this criterion. At the start of the dayside measurements of the orbit, the Sun is partly in the total clear field of view (TCFOV) of the instrument, which clearly can be identified by solar spectral features in the dark signal measurement at 350 km tangent altitude. This produces errors in the subtraction of the dark signal especially for the transition region between

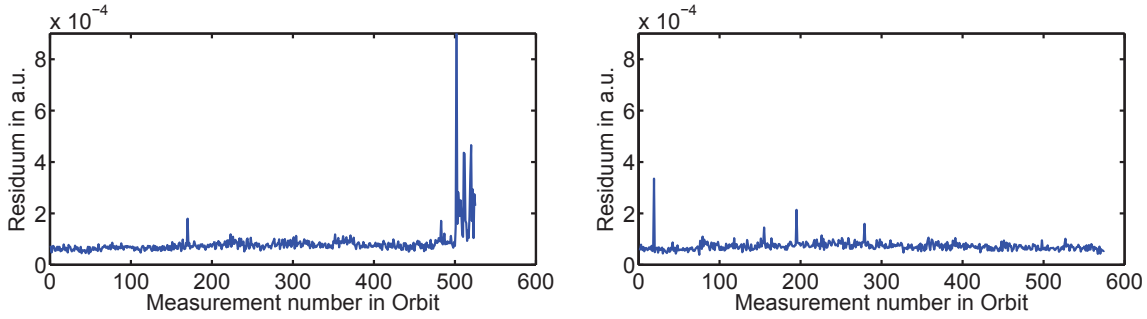

Fig. 7.1: Mean residuum for the Mg line fit window for all measurements of an orbit affected (left) and not affected (right) by the SAA. The SAA can be detected for the last measurements of the orbit at mid to high southern latitudes.

sunlight contaminated and uncontaminated spectra. Dark measurement signals for the first dayside measurements of an orbit as well as slant column density signals are shown in Fig. 7.2. The solar stray light contaminated states and the first not stray light contaminated state are separated out.

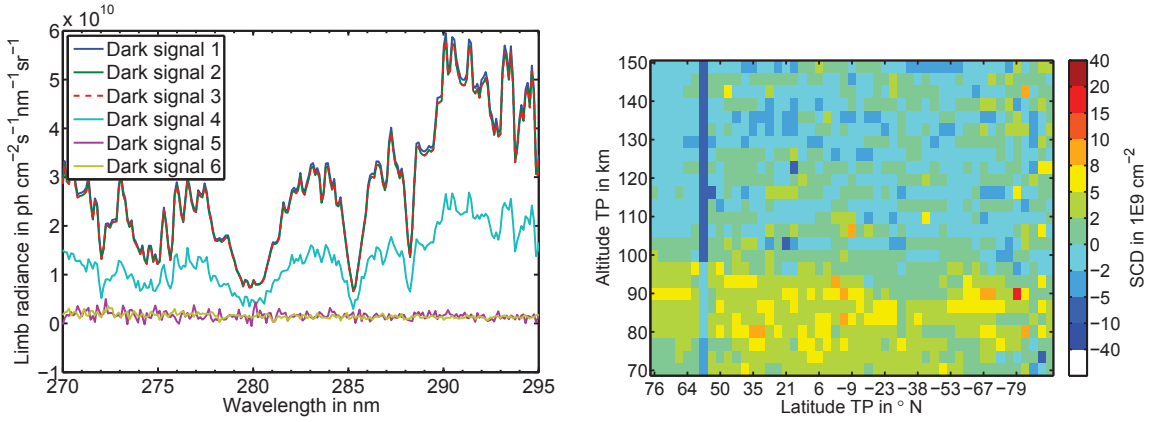

Fig. 7.2: The left figure shows the dark signal for the first 6 dayside states of an orbit. The first 4 states clearly show a solar stray light signal. The right figure shows slant column densities for Mg for an orbit. While the strong stray light contaminated spectra show reasonable slant column densities, there is a clear negative vertical stripe between the strong stray light contaminated and the not stray light contaminated measurements at 60° N.

The averaging process starts with the definition of a reference orbit. Figure 7.3 shows an example of the limb measurements pattern for 4 consecutive orbits (43177-43180). While the longitudes of the states in each orbit change, the latitudinal pattern is alternating for even and odd orbits. The reason for this alternation is most probably, that the limb and nadir states alternate within an orbit. As a result, there would be latitudinal gaps in coverage, if always the same pattern was used. The odd and even orbit with the most limb states are taken and merged to one reference orbit. As the first states in odd and even orbits have very similar latitudes

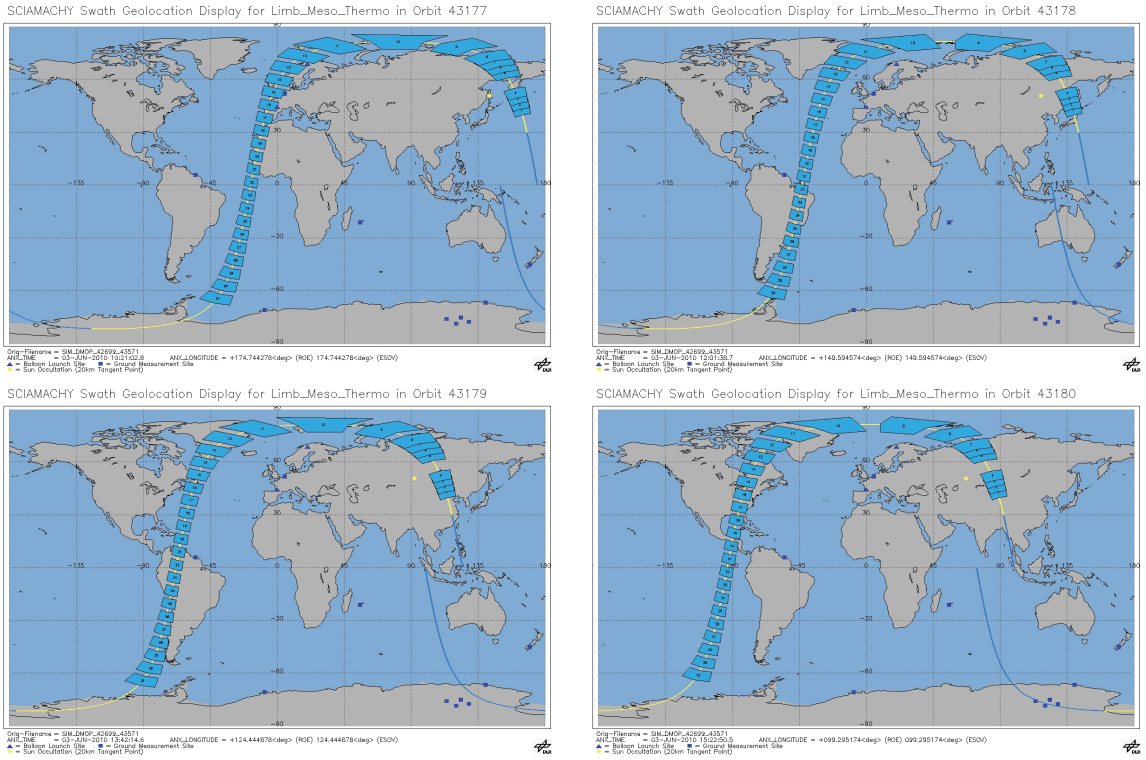

Fig. 7.3: Pattern of SCIAMACHY MLT limb states for the orbits 43177 (upper left), 43178 (upper right), 43179 (bottom left) and 43180 (bottom right). The odd orbits have 31 limb states, while the even orbits have 30 limb states. The latitude of the states in the orbits is similar for consecutive odd orbits and consecutive even orbits. (The graphics are taken from `http://www.atmos.caf.dlr.de/projects/scops/geolocation/limb/L_swath_geo_files_no_ftp.html`)

and local times, only one of those nearly overlapping states is used in the reference orbit. Once the reference orbit with the reference geolocations is found, limb states from all orbits are matched to the reference geolocations and are averaged. This produces a reference orbit with nearly double as much geolocations, as a single odd or even orbit. Instead of a reference orbit with up to 15 single measurement at a geolocation, this orbit has only up to 7 or 8 single measurements at a geolocation, as the separation between odd and even orbits is still there. To improve the number of single measurements in the average, an additional latitudinal smoothing of the spectral data is performed.

7.2 Seasonal variations of Mg

Results of monthly averaged data of Mg, retrieved from the 285.2 nm line, are presented in this section. The averaging process was described in Sect. 7.1. The 2D altitude-latitude grid has 40 latitude intervals between 82° S and N on the descending node side and also 40 intervals for the ascending node side. Note, that the intervals on the ascending node side are actually only sparsely used, because most measure-

ments in this region are excluded, mostly because they are night measurements. If the stray light contaminated measurements (see Fig. 7.2) were not excluded, more grid elements on the ascending node side would show valid results. The monthly averaged results for the MLT data set are shown in Fig. 7.4. Mg shows a peak

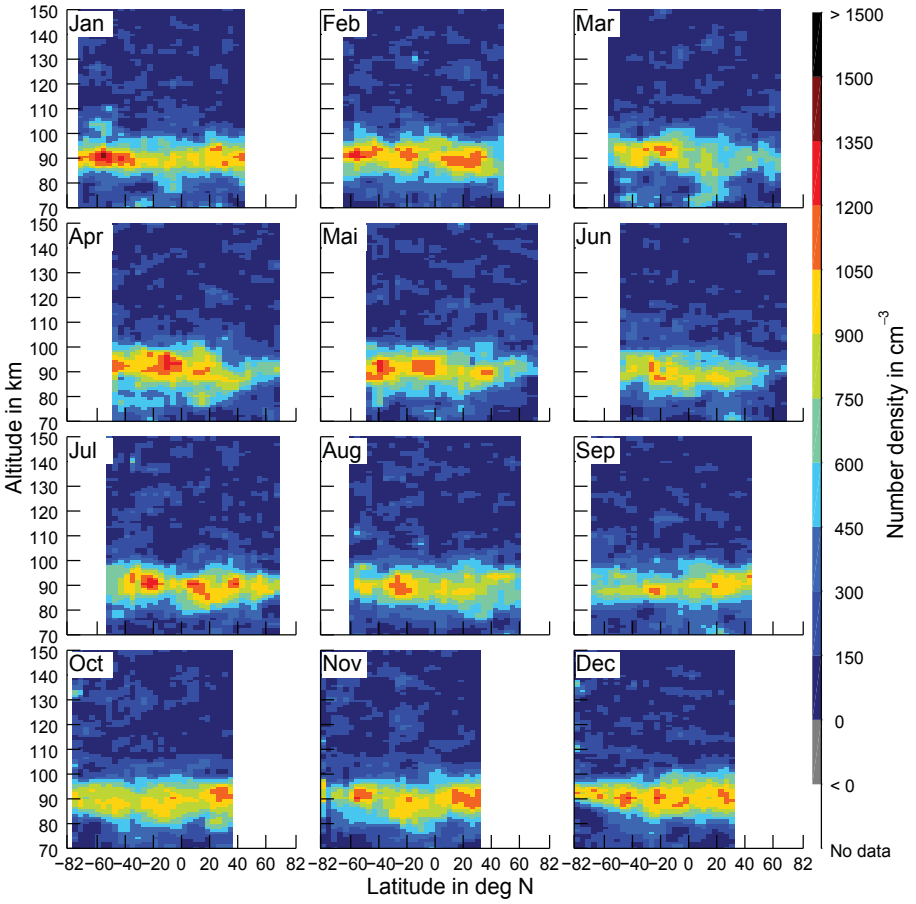

Fig. 7.4: Latitudinal variation of monthly (averaged over all January, February, etc. results from 2008–2012) and zonally averaged Mg density profiles retrieved from the 285.2 nm line. The Mg layer peaks at around 90 km altitude and has a FWHM of about 15 km.

at around 90 km with peak densities between 750 cm^{-3} and 1500 cm^{-3}. There is a strong month-to-month variability in the latitudinal distribution. The averaged yearly results are shown in Fig. 7.5. In the yearly average, the densities in the low latitudes are higher than at the high latitudes in the peak region. At the highest covered southern latitude there are increased densities above 100 km, which, however, are most likely caused by the fact that this latitude is less covered and therefore has a larger statistical error.

Figure 7.6 shows the seasonal variation of the vertical profile for low, mid and high latitudes, and Fig. 7.7 shows the vertical column densities (VCDs). The month-to-month variations are large, and thus the error on the estimate of the amplitude of seasonal variations for mid latitudes is relatively large. The seasonal

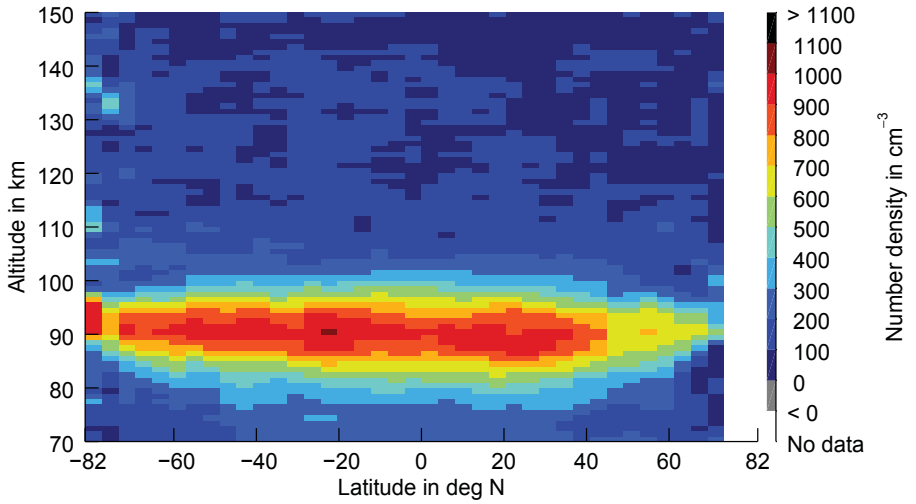

Fig. 7.5: Latitudinal variation of Mg profiles averaged over all available data for the years 2008–2012. Note that only between 45° N and 60° S do measurements from all 12 months contribute to the averages, while for high latitudes there is only summer coverage and/or near-terminator measurements.

variation of the Mg peak altitude is estimated to be less than 5 km. The VCD varies between $0.5 \cdot 10^9$ cm^{-2} and $3.5 \cdot 10^9$ cm^{-2}. The VCD between 40°N and 40°S is higher than for higher latitudes, which are only covered in the hemispheric summer. The seasonal variation is small compared to the mean of the vertical column densities. This relatively small variation is required to retrieve the correct mean value (because of non-linearities in the retrieval) as comprehensively discussed by Langowski et al. (2014a) and in Sect. 6.4.2.

7.3 Seasonal variations of Mg^+

There are two spectral lines for Mg^+, one at 279.6 nm and the second one at 280.4 nm. Mg^+ densities are independently retrieved from both lines. The differences in the peak region are lower than 25%. The 280.4 nm line has the worse signal to noise ratio. However, this line has the advantage of not being polarized, which removes this systematical error source. Therefore, mainly results for the 280.4 nm line are shown here. The densities on an altitude and latitude grid for the monthly averaged results are shown in Fig. 7.8.

The Mg^+ densities peak at an altitude of 95-105 km with peak densities of 500-4000 cm^{-3}. The peak densities show a seasonal variation with a summer maximum between 25°-45° in both hemispheres. In summer, the highest peak altitude is roughly at 45° (N and S). It occurs around 105 km, which is up to 10 km higher than the lowest peak altitude at the equator and higher than that at the poles. There is a minimum at the equator and also a second maximum in peak altitude in the mid latitudes on the winter hemisphere, that, however, is more variable than the one in the summer hemisphere.

The latitudinal variation of Mg^+ is also seen in the average over all available

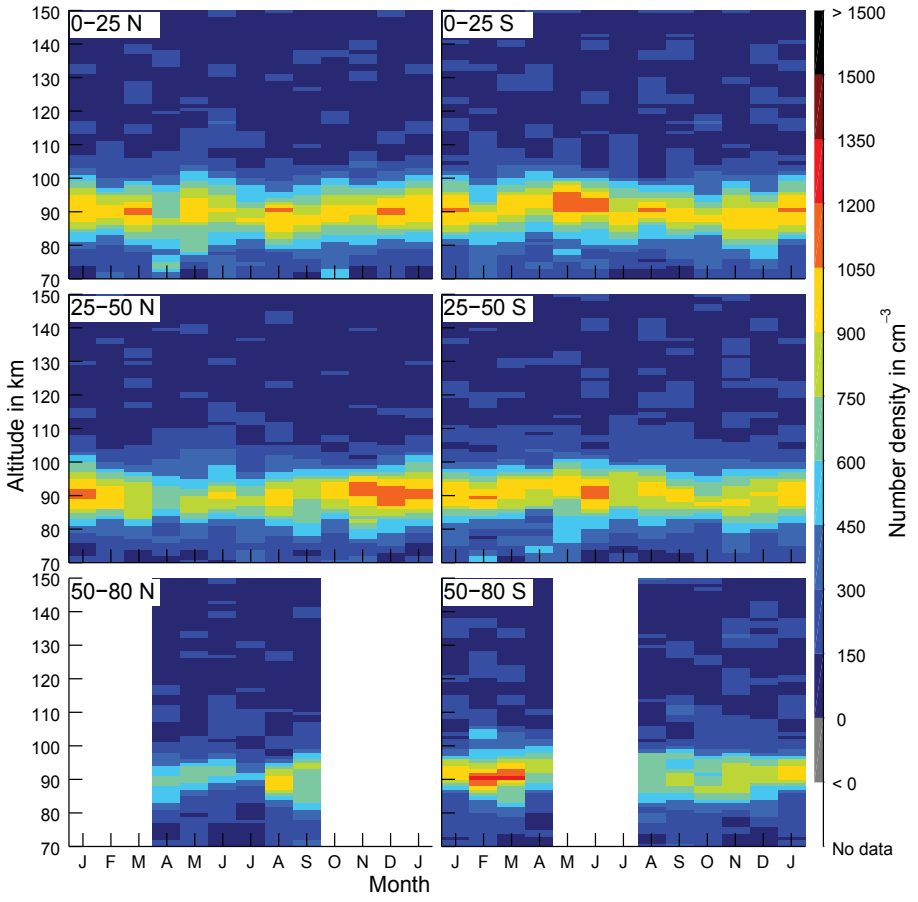

Fig. 7.6: Seasonal variation of the vertical Mg profile for different latitudinal zones (left: Northern Hemisphere, right: Southern Hemisphere, top: low latitudes, mid: mid latitudes, bottom: high latitudes). The peak altitude is at 90 km for all latitudes with variations of ±5 km. The month-to-month variations in peak altitude and density are bigger than any seasonal variations.

measurements in 2008-2012, which is shown in Fig. 7.9. The latitudinal dependence of the peak altitude seems to be symmetric towards the equator. There is an asymmetry in the peak density, which is higher in the Northern Hemisphere. However, the profile in the Southern Hemisphere is broader, so that the differences in the VCDs are smaller. This difference between the hemispheres may be explained by Lorentz force induced transport processes. Since the geographic and geomagnetic latitudes map better at the Northern Hemisphere, the zonal geographic averaging can lead to a wider profile in the south. The better mapping of geographic and geomagnetic latitudes arises from the position of the magnetic poles. The North Magnetic Pole is at around 82° N which is much closer to the geographic pole than the South Magnetic Pole at around 65° S. Furthermore, the South Atlantic Anomaly, which is far away from the South Magnetic Pole (at around Australia) is excluded from the measurements.

When looking at the densities above the maximum density region for mid and

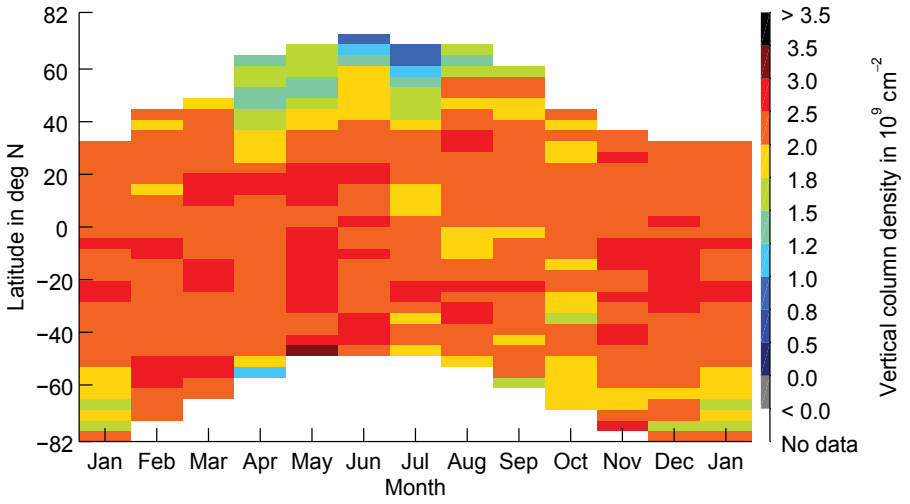

Fig. 7.7: Seasonal variation of the Mg vertical column density (VCD) between 70 km and 150 km for different latitudinal regions. The VCD varies between $0.5 \times 10^9\,\mathrm{cm}^{-2}$ and $3.5 \times 10^9\,\mathrm{cm}^{-2}$. No clear seasonal cycle is observed. Variations between consecutive days with measurements (at least 14 days difference) are even higher than the inter-monthly variations.

low latitudes, the density at same altitude is lower in the equatorial region than at 40° below 110 km, while this behavior is reversed above around 110 km. Furthermore, the density at the uppermost altitude level is enhanced in the equatorial region, pointing to the abundance of Mg^+ above this altitude. This may be explained by enhanced vertical upward transport in the equatorial region.

Vertical transport phenomena of charged particles in the equatorial region are well known and are well connected with the so-called equatorial electrojet (see, e.g., Forbes, 1981). The equatorial electron distribution with transport from the equator to higher latitudes has been observed in ionospheric soundings. It is known as the equatorial anomaly or Appleton anomaly (Kendall and Windle, 1965). Typically, there are fewer electrons at the equator than at 20° on either side of the equator due to this transport. In-situ satellite measurements of the metals above 120 km have been carried out using the Atmospheric Explorer satellite. A discussion of these data can be found in Grebowsky and Aikin (2002).

Drifts of charged particles in electromagnetic fields often can already be explained with the movement of single particles in those fields. As these particles themselves produce those fields, the full discussion is a bit more complex (see the above mentioned equatorial electrojet). The motion of a single charged particle with charge q can be separated into the gyration in the magnetic field $\vec{B}$, and the motion of the guiding centre. For a force $\vec{F}$, which is perpendicular to the magnetic field, the guiding centre moves perpendicular to $\vec{B}$ and $\vec{F}$ with the drift velocity $\vec{v} = \frac{1}{q}\frac{\vec{F}\times\vec{B}}{B^2}$. Using the electric force $\vec{F} = q\vec{E}$ results in the drift velocity $\vec{v} = \frac{\vec{E}\times\vec{B}}{B^2}$. This drift is called $\vec{E} \times \vec{B}$-Drift (see, e.g., the textbook of Chen, 1984).

The vertical upward transport of Mg^+ ions near the equator has been discussed,

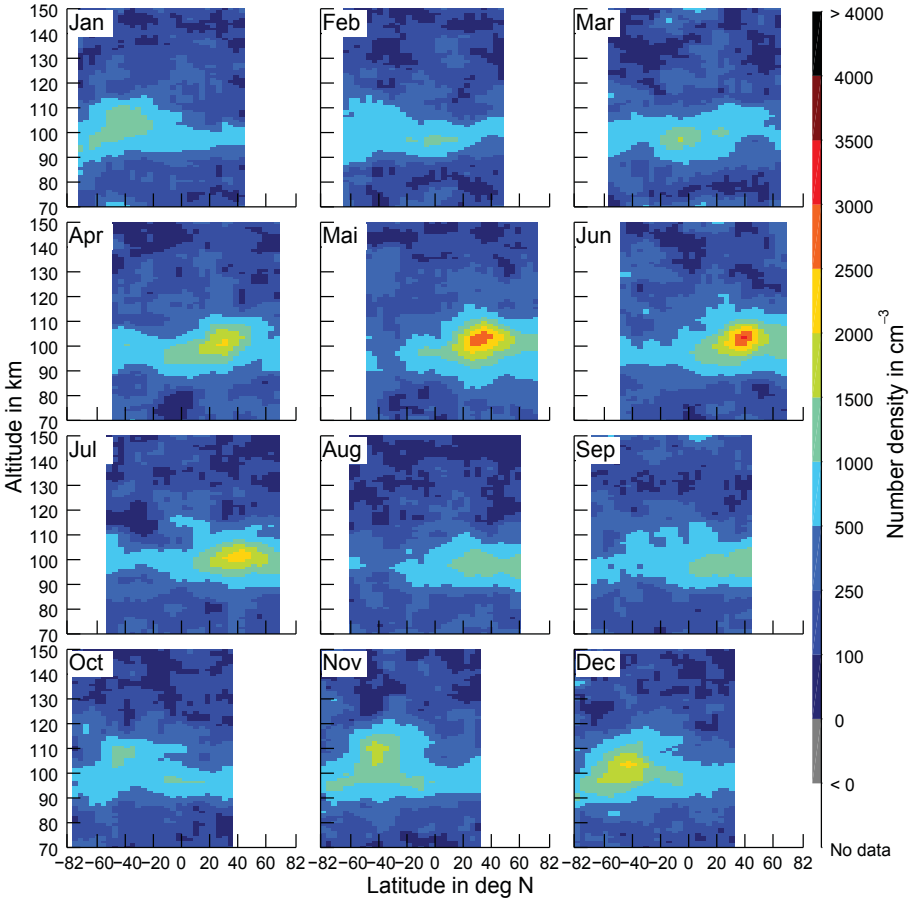

Fig. 7.8: Latitudinal variation of monthly and zonally averaged Mg^+ profiles (for the years 2008–2012) retrieved from the Mg^+ line at 280.4 nm. Mg^+ shows a seasonal cycle with a summer maximum, which is especially pronounced in the region between 25° and 45° latitude in the summer hemisphere. Furthermore, in this region the peak altitude is about 105 km, which is up to 10 km higher than at the equator or at the poles.

e.g., in Hanson and Sterling (1972). The Sun is passing the equatorial region from east to west. The ionisation in the thermosphere is strong for low solar zenith angles (SZAs), which leads to denser plasmas there. This leads to a strong electric field from east to west, which results in a strong eastwards current, the so-called equatorial electrojet (see, e.g., Forbes, 1981). The magnetic field along the meridians and the electric field along the equator lead to an $E \times B$-Drift perpendicular to both in the radial direction, which lifts the ions and electrons to higher altitudes and even above the F-layer. However, this effect can only explain an upward transport within ± 3° latitude. Fesen et al. (1983) showed, that additionally neutral meridional winds have to be taken into account, in order to explain vertical transport also at higher latitudes up to 30°, and this was experimentally shown, e.g., in Fesen and Hays (1982a), Fesen and Hays (1982b) and Gérard and Monfils (1978).

As for Mg, the annual means show higher densities above 110 km at the north-

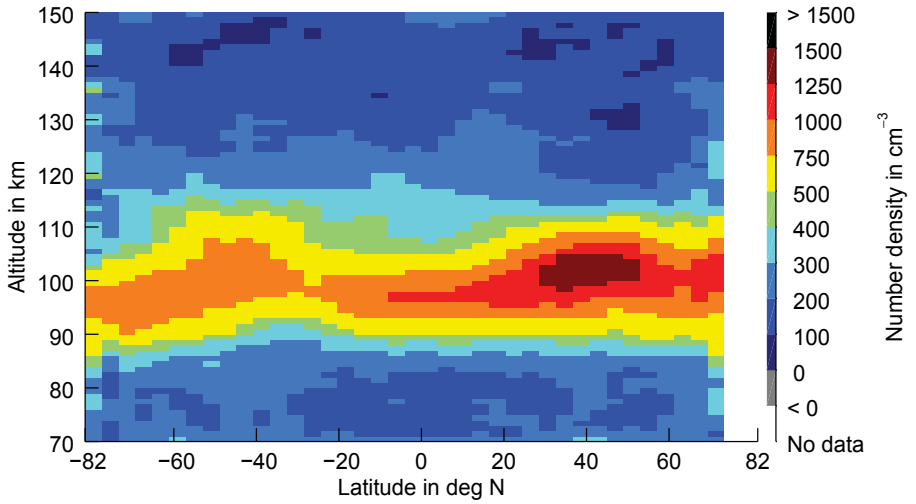

Fig. 7.9: Latitudinal variation of Mg^+ profiles averaged over all available data for 2008–2012 for the Mg^+ line at 280.4 nm.

ernmost and southernmost latitudes due to reduced coverage and, therefore, higher statistical errors. However, at high northern latitudes, which are cut off in the shown results, high densities can be retrieved when also using the stray light contaminated measurements. This is shown in Fig. 7.10. The results as well as the input raw data show low densities below 90 km and at different altitudes in between the high signal region, which makes a differentiation between a stray light effect and the true metal emission complicated. Should this feature be real, it could be explained by the

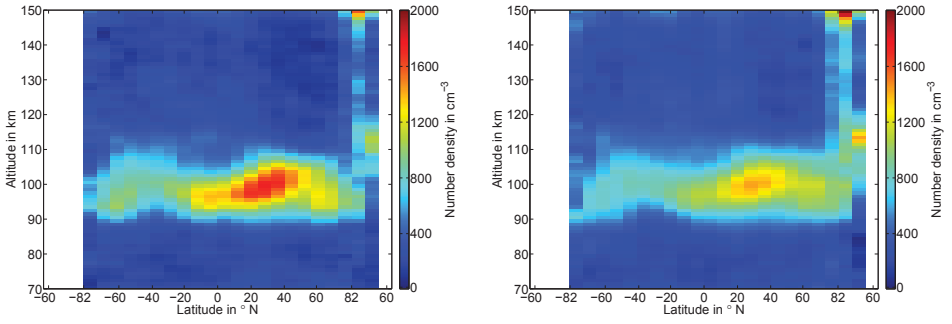

Fig. 7.10: Mg^+ profiles averaged over all available data for 2008-2012 for the Mg^+ line at 279.6 nm (left) and the Mg^+ line at 280.4 nm (right). The stray light contaminated measurements at the start of each orbit are also used here extending the coverage to higher northern latitudes and even the ascending node side. High Mg^+ densities are observed at the north pole. Note that the latitudinal resolution is a factor 2 lower than in Fig. 7.9.

cleft ion fountain found by Lockwood et al. (1985), which describes the transport of charged particles along the magnetic field lines, which close at the pole. This transport may lift charged particles up to several Earth radii until they become

neutralized and sink down if they are heavy enough and are not quickly ionized again.

Figure 7.11 shows the seasonal variation of the vertical Mg^+ profile for low, mid and high latitudes, and Fig. 7.12 shows the corresponding VCDs. In each of the latitude regions, a maximum in peak altitude is found for the summer time in the corresponding hemisphere. The VCD varies between $1 \cdot 10^9\,\mathrm{cm}^{-2}$ and $6 \cdot 10^9\,\mathrm{cm}^{-2}$.

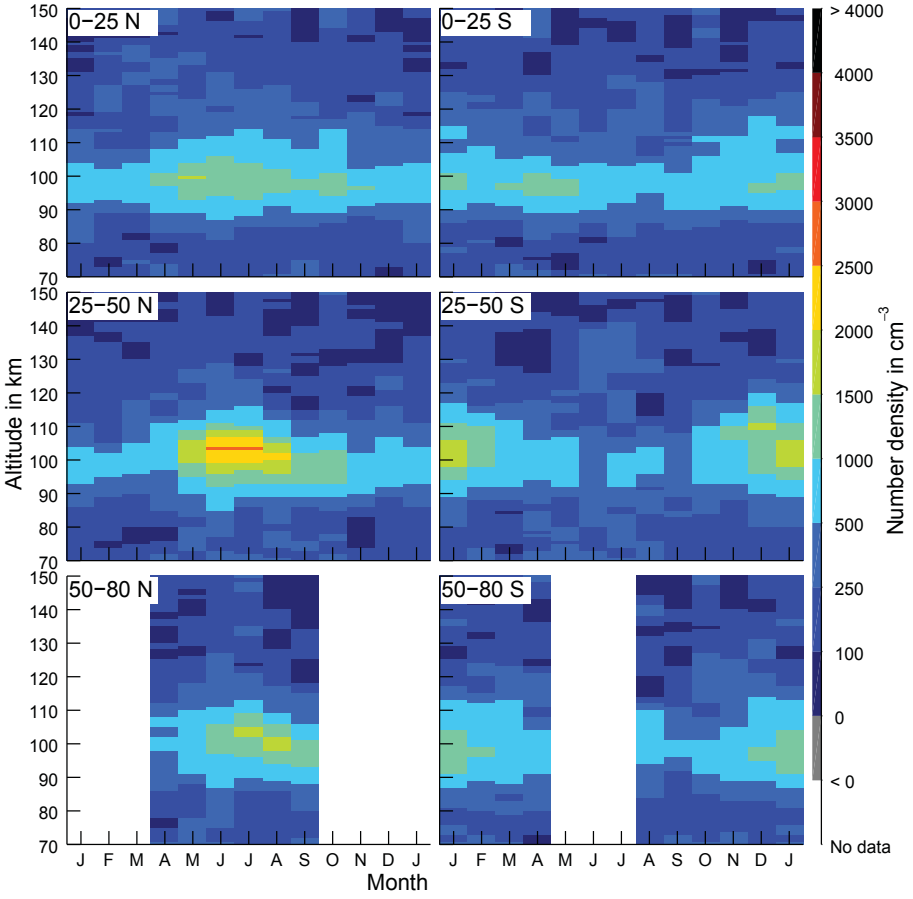

Fig. 7.11: Seasonal variation of the vertical profile of Mg^+ (280.4 nm line) for different latitudinal zones (left: Northern Hemisphere, right: Southern Hemisphere, top: low latitudes, middle: mid latitudes, bottom: high latitudes), averaged over all available observations in 2008–2012. For all these latitudinal zones a seasonal cycle in the peak altitude with a summer maximum and a winter minimum is observed. The seasonal variations are similar or larger than the month-to-month variations. The maximum peak altitude is 5–10 km higher than the minimum peak altitude. For mid latitudes, the peak altitude in the winter hemisphere is still higher than the peak altitude at the equator (see also Fig. 7.8).

A seasonal cycle with a summer maximum is observed and the strongest variations are observed between 20° and 40°. Figure 7.13 shows the ratio of Mg^+ and Mg VCDs. For low and mid latitudes, as well as for the high latitudes in the southern hemispheric summer the ratio is in between 0.5 and 5. The seasonal variability of

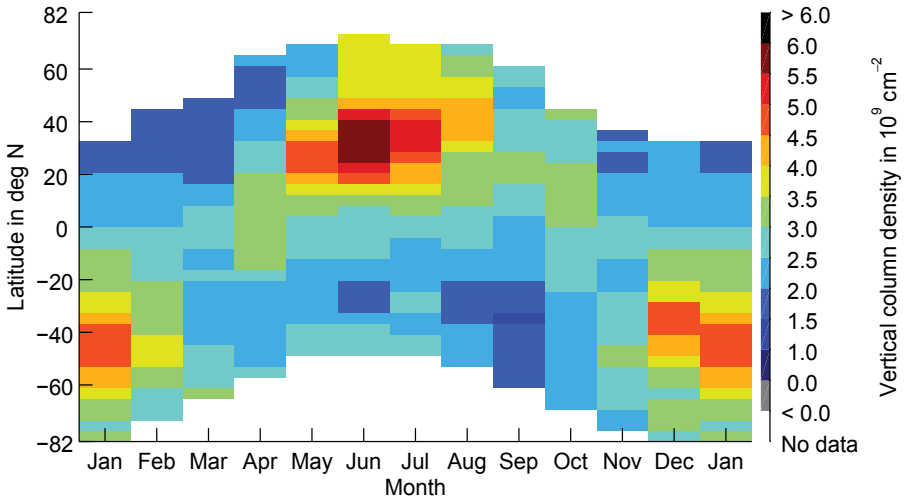

Fig. 7.12: Seasonal variation of the Mg^+ vertical column density (VCD) between 70 km and 150 km and for different latitudinal regions, retrieved from the 280.4 nm line. A clear seasonal cycle with a summer maximum is observed. The summer maximum in the Northern Hemisphere has higher values compared to the one in the Southern Hemisphere. The highest variability can be found between 25° and 50° in both hemispheres.

the ratio follows the one of Mg^+, as Mg does not show a strong seasonal variability.

7.4 Comparison to other measurements

Mg and Mg^+ cannot be observed from the ground, because the wavelengths of the lines are below 300 nm and the emission signal is absorbed strongly by ozone in the stratosphere. As a result, only few measurements are available and rely on observations from sounding rockets and satellites. The first rocket-borne ion mass spectrometer measurement of metal ions was reported by Johnson and Meadows (1955) (May 1954, White Sands) and enhanced ion signals between 93 and 124 km were found. According to Grebowsky and Aikin (2002) approximately 50 flights of rocket-borne mass spectrometers have been made until 2002, probing the region between 80 and 130 km altitude. Results of these flights for Mg^+ can, e.g., be found in Istomin (1963), Narcisi and Bailey (1965), Narcisi (1971), Aikin and Goldberg (1973), Philbrick et al. (1973), Zbinden et al. (1975), Herrmann et al. (1978), Kopp and Herrmann (1984), Kopp et al. (1985b,a), Kopp (1997) and Roddy et al. (2004). Date, local time, latitude and reference publications of these and other rocket flights can also be found in table 1 of Grebowsky et al. (1998). Figure 7.14 shows reproduced plots of vertical number density profiles of Mg^+. The original sources of the data for these plots are listed in table 7.6.

A disadvantage for the comparison with the new data set presented in this study is, that most of these measurements were performed during special events having E_s layers (sporadic electron layers), aurora, meteor showers, stratospheric

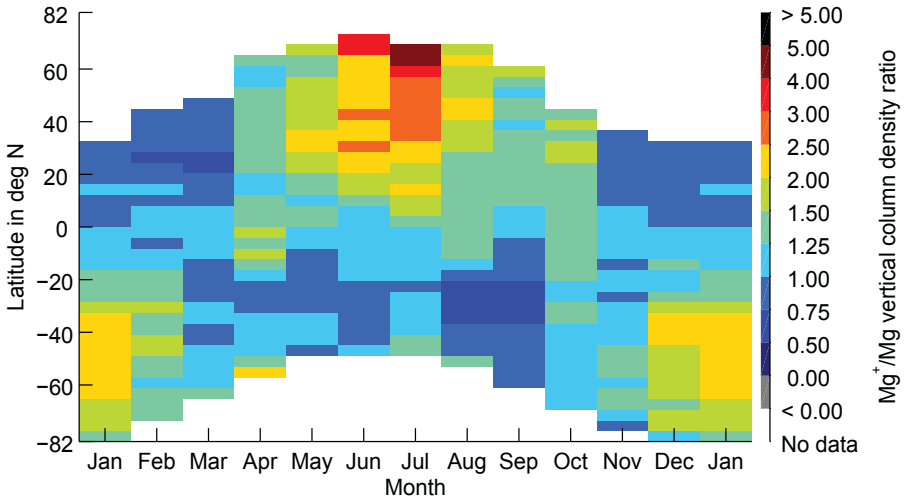

Fig. 7.13: Seasonal variation of the Mg^+ to Mg ratio for different latitudes. The ratio varies between 0.5 and 5 and shows a summer maximum, which is in good agreement with Correira et al. (2008).

Table 7.6: Original sources and geolocations of plots in Fig. 7.14.

No.	Lat	Long	Date	Original reference	Taken from
1	mid N	U.S.S.R.(Eur)	15 June 1960	Istomin (1963)	Istomin (1963)
2	30 N	86 W	31 October 1963	Narcisi and Bailey (1965)	Narcisi and Bailey (1965)
3	30 N	86 W	12 April 1967	Narcisi (1971)	Grebowsky and Aikin (2002)
4	30 N	86 W	12 April 1967	Narcisi (1971)	Grebowsky and Aikin (2002)
5	8 N	77 E	19 March 1970	Aikin and Goldberg (1973)	Aikin and Goldberg (1973)
6	8 N	77 E	19 March 1970	Aikin and Goldberg (1973)	Aikin and Goldberg (1973)
7	8 N	77 E	9 March 1970	Aikin and Goldberg (1973)	Aikin and Goldberg (1973)
8	8 N	77 E	10 March 1970	Aikin and Goldberg (1973)	Aikin and Goldberg (1973)
9	30 N	86 W	20 November 1970	Philbrick et al. (1973)	Grebowsky and Aikin (2002)
10	40 N	9 E	14 December 1971	Zbinden et al. (1975)	Zbinden et al. (1975)
11	38 N	75 W	12 August 1976	Herrmann et al. (1978)	Kopp (1997)
12	51 N	93 W	24 February 1979	Kopp and Herrmann (1984)	Kopp (1997)
13	67 N	20 E	30 November 1980	Kopp et al. (1985b)	Kopp (1997)
14	67 N	20 E	13 August 1978	Kopp et al. (1985a)	Kopp (1997)
15	38 N	75 W	1 July 2003	Roddy et al. (2004)	Roddy et al. (2004)
16	±20 N	all	15 October–29 November 1999	Minschwaner et al. (2007)	Minschwaner et al. (2007)

warmings or NLCs present. Sporadic Mg^+ layers often occurred between 105–110 km and/or at around 120 km altitude. The Mg^+ peak altitude in most of these measurements can be found between 90 and 95 km altitude and the full width at half maximum (FWHM) of the layers is on the order of 5–10 km, but sometimes the FWHM is only 1 km.

When comparing in situ mass spectrometer measurements with satellite remote sensing results, it has to be noted, that the in situ measurements are localized and limited to the direct vicinity of the rocket. In contrast, remote sensing techniques typically cover a large volume with horizontal distances along and perpendicular to the viewing direction of several hundreds of kilometres. This results in smoother appearing layers with a larger FWHM for the remote sensing method. Still, the width of the Mg^+ layers as well as the peak density presented in this study are in good agreement with the in situ rocket measurements. However, the strong latitudinal

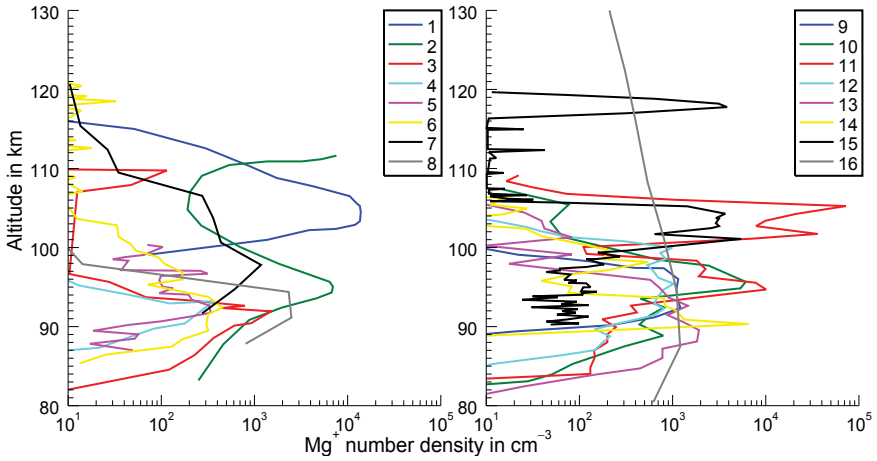

Fig. 7.14: Measurements of the vertical Mg^+ number density profile mentioned in the text and listed in table 7.6. Note that the data points are redrawn from the original figures, and not all details of the original figures may be captured.

dependence of the peak altitude shown in Fig. 7.11 as retrieved from SCIAMACHY measurements is not found in the in situ rocket data.

In addition to the in situ measurements with rockets, there are also airglow measurements available from rockets, the space shuttle and satellites. However, to retrieve density information from this method, radiative transfer models as well as inversion techniques and computational power are needed. As a result, slant column information rather than profiles has been retrieved and made available in the first remote sensing studies (see, e.g., Boksenberg and Gérard, 1973; Gérard and Monfils, 1974). Note, that the peak slant column density profile peaks a few km lower than the density profile. In Anderson and Barth (1971) (summer, $\approx 40°$ N) the region up to 106 km altitude is scanned during a sporadic E_s layer event, and the peak altitude of Mg^+ was not observed during this flight, i.e., it was higher than 106 km. No Mg signal above the instrumental noise was observed in these spectra.

The region above the peak altitude from 150 km up to the F-layer and above was investigated by Gérard and Monfils (1978), Fesen and Hays (1982a), Mende et al. (1985), Gardner et al. (1995) and Gardner et al. (1999) and typically shows less than 100 cm^{-3} Mg^+ ions at 150 km altitude. This is in good agreement with the profiles described in Sect. 7.3. In some cases, a higher density is observed in the profiles retrieved from the SCIAMACHY limb observations. This is explained by retrieval artifacts on the edge of the retrieval grid and vertical constraints tuned for the main peak. These effects result in small oscillations, compared to the main peak, in regions with lower density. Since these artifacts appear at the upper edge of the profile, it also implies that there is a significant density above the highest tangent altitude for Mg^+.

Minschwaner et al. (2007) show a combined NO and Mg^+ retrieval for satellite limb measurements, which is in good agreement with the results in Sect. 7.3 in terms of the Mg^+ concentrations at peak altitude and at the upper edge of the profile. However, the Mg^+ peak altitude is at 90 km. Taking into account the coarser

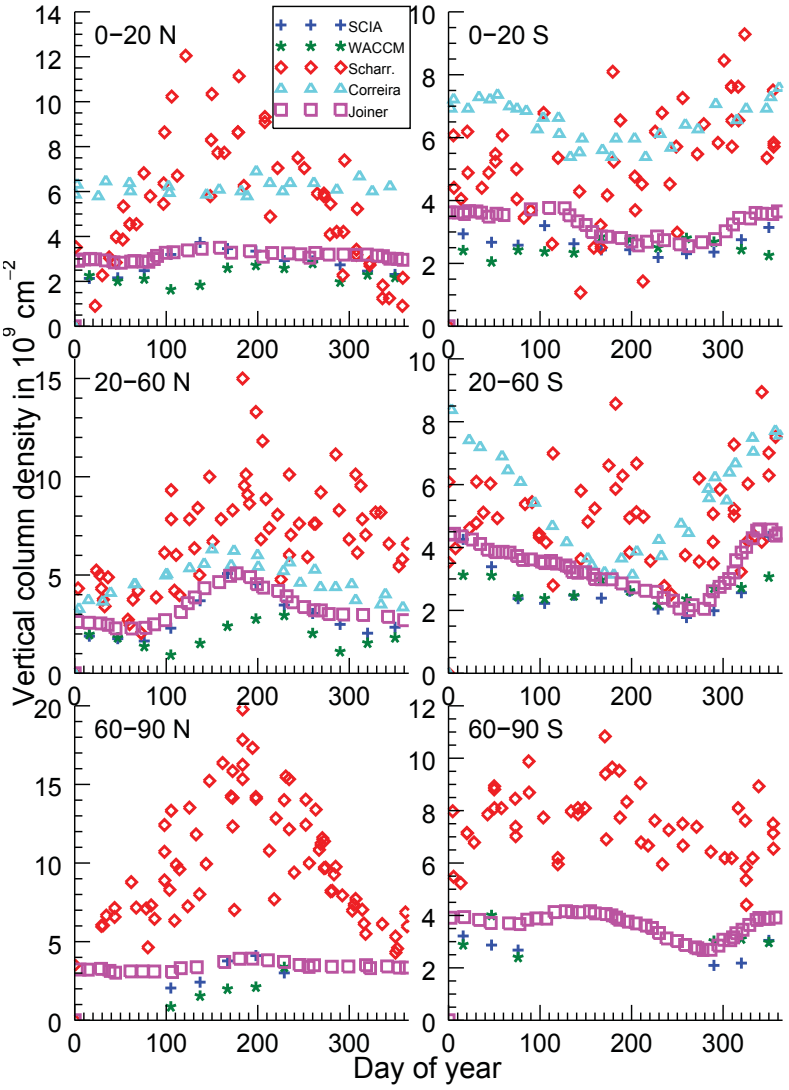

Fig. 7.15: Mg^+ VCD results for different latitudinal regions and from different sources. The left column shows the Northern Hemisphere and the right column the Southern Hemisphere. From top to bottom results for low, mid and high latitudes are shown. Note, that different sources originally used different time periods. The data used are from SCIAMACHY (blue) and WACCM (green) from this work (0–20° N/S, 20–60° N/S, 60–90° N/S, WACCM data co-located to SCIAMACHY coverage), from Scharringhausen (2007) (Figs. 7.3 and 7.4, red), from Correira et al. (2008) (Fig. 1, cyan) and Joiner and Aikin (1996) (Figs. 10 and 11, magenta).

sampling (every 7 km), higher statistical errors and a tangent height offset (±4 km) in Minschwaner et al. (2007), the agreement is reasonable. Accurate tangent height determination was an issue for SCIAMACHY, too. Offsets of a similar magnitude were initially observed for SCIAMACHY data products (see, e.g., von Savigny et al., 2005). However, this error source was minimized and the tangent height knowledge was improved to ±200 m (von Savigny et al., 2009). The NO band emission, which overlaps with the Mg^+ lines and the Mg line, is of the same order of magnitude as the Mg^+ lines in the study by Minschwaner et al. (2007) and even bigger than the Mg emission, which made a NO correction necessary. The SCIAMACHY MLT data set does not show these strong NO lines at 280 and 285 nm. NO in this region is very sensitive to solar activity. Results on NO retrievals from the same SCIAMACHY level 1 data set as used in this study are reported by Bender et al. (2013). Only at high latitudes in summer and winter and from late 2011 to 2012 do the results of Bender et al. (2013) show NO density of the same magnitude as the equatorial density plot for NO in Minschwaner et al. (2007). However, we did not observe clear NO signals in the vicinity of the Mg/Mg^+ lines during this period.

A time series of Mg^+ vertical columns covering several years and retrieved from SBUV nadir measurements was presented by Joiner and Aikin (1996). These measurements were performed approximately 1 day per month, with a spectral resolution of 1.13 nm and a spectral sampling every 0.2 nm (compare to SCIAMACHY with ≈ 0.22 nm resolution and sampling every ≈ 0.11 nm). The results in Joiner and Aikin (1996) are in very good agreement with the results obtained in this study, especially when comparing Figs. 10 and 11 in Joiner and Aikin (1996) with Fig. 7.12 in this study. Figs. 7.15 and 7.16 show redrawn VCD time series from different sources for a quick and easy comparison of the results of Mg^+ and Mg (see the original sources for more details).

Mg was also investigated in Joiner and Aikin (1996). However, the average VCD for these profiles, where the signal was significant, is $4 \times 10^{10}\,cm^{-2}$, which is a factor of 10 more than the VCD in Fig. 7.7. These large discrepancies must be investigated in the future.

Satellite measurements with long time series and daily coverage are available from GOME and GOME-2 in nadir mode and from SCIAMACHY in nadir and nominal limb mode. All three instruments have a similar spectral resolution. The VCD of Mg and Mg^+ was retrieved from the GOME data set by Correira et al. (2008, 2010) and Correira (2009). Figures 1 and 2 in Correira et al. (2008) show Mg and Mg^+ VCD for 1996 and 1997 as well as the Mg^+ to Mg ratio for latitude intervals from 0–10° and 30–40° for both hemispheres. For the low latitudes, where there is less seasonality, the VCD for Mg^+ is about $6–7\times10^9\,cm^{-2}$, which is higher than in Fig. 7.12. The Mg density is about $3\times10^9\,cm^{-2}$, which is similar to that in Fig. 7.7. In the equatorial region the Mg column agrees better, and a higher VCD in the nadir results for Mg^+ can be explained by the thermospheric part of Mg^+, which is not part of the VCD in Fig. 7.12. At mid latitudes Mg^+ shows a strong seasonal cycle with a summer maximum in Correira et al. (2008). This seasonality is quite symmetric for both hemispheres with higher VCD in the Southern Hemisphere. Furthermore, the summer maximum at mid latitudes does not exceed the VCD at low latitudes. For Mg a summer maximum is observed, which is more pronounced

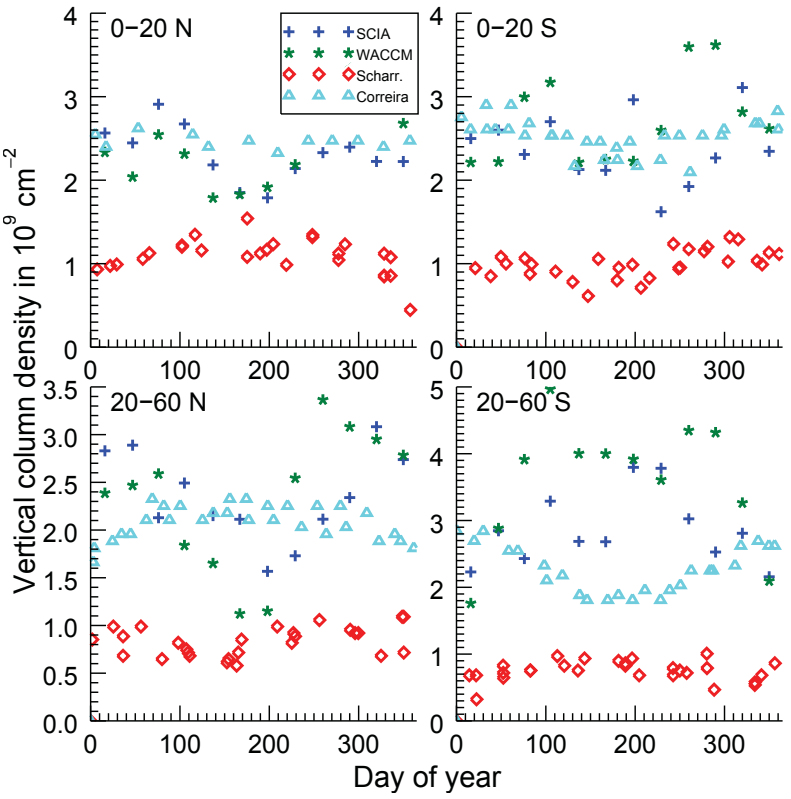

Fig. 7.16: Mg VCD results for different latitudinal regions and from different sources. The left column shows the Northern Hemisphere and the right column the Southern Hemisphere. From top to bottom results for low and mid latitudes are shown. Note, that different sources originally used different time periods. The data used are from SCIAMACHY (blue) and WACCM (green) from this work (0–20° N/S, 20–60° N/S, 60–90° N/S, WACCM data co-located to SCIAMACHY coverage), from Scharringhausen (2007) (Figs. 7.3 and 7.4, red) and from Correira et al. (2008) (Fig. 1, cyan).

in the Southern Hemisphere. The Mg^+ to Mg VCD ratios in Correira et al. (2008) are in good agreement with our results shown in Fig. 7.13.

The nominal limb mode data set, covering the tangent height range from the surface up to 92 or 105 km (only until early 2003), respectively, combined with the nadir mode data set from SCIAMACHY from 2002 to 2007 was investigated in Scharringhausen et al. (2008b,a) and Scharringhausen (2007). The peak region for the ions could not be fully resolved with the nominal limb mode. However, qualitatively the agreement between the results by Scharringhausen (2007) (his page 69 and 70) and the results from this study for the ions is quite good, showing a similar seasonal cycle in the Northern Hemisphere. At high northern latitudes Scharringhausen (2007), however, retrieved up to a factor 3–4 larger density and the highest density is found at high latitudes, and not at mid latitudes, as in this study. In the Southern Hemisphere the seasonal cycle was not identified well in Scharringhausen (2007), which agreed with the results similar to Fig. 7.12 when

using the pre-flight polarization correction (not shown), which we now found to be wrong (see Sect. 6.5.6).

At the very beginning of the Scharringhausen VCD data set in 2002 (e.g., Figs. 69 and 70 in Scharringhausen (2007)), slightly higher VCD for both species are observed than for the other years. This may come from increased NO signals near the Mg and Mg^+ lines during solar maximum. Another reason for the higher Mg^+ VCD in 2002 may also be related to the change of the maximum tangent altitude from 105 to 92 km.

More differences between our retrievals and the results by Scharringhausen (2007) can be found in the Mg data. This is, because the data product retrieved using the Mg line is more affected by the radiative transfer improvements made in this study than the product retrieved using the Mg^+ lines. For example, there was no correction of the Ring effect, the filling in of Fraunhofer lines by inelastic Raman scattering in the Earth's atmosphere (Grainger and Ring, 1962), in Scharringhausen (2007), which led to high density below 80 km and a density maximum at the lower edge of the retrieval boundaries at 70 km. This inelastic scattering contribution additionally adds a seasonal variation to the data set.

Furthermore, the Mg line at 285.2 nm is much more affected by self-absorption of the emission, which was not considered in Scharringhausen (2007), so the Mg VCD was smaller than reported here. However, most of the important findings from the further analysis of the Scharringhausen data set are only weakly affected by these differences and are still valid.

In summary, it can be concluded that the Mg/Mg^+ results presented here are often in good agreement with previous satellite and rocket instruments. The most striking difference to previous measurements is the strong latitudinal dependence of the peak altitude of Mg^+ with differences of up to 10 km for different latitudes.

7.5 Comparison of Mg and Mg^+ observed by SCIAMACHY and modeled with the WACCM model

The WACCM-Mg Mg/Mg^+ data set was produced by our cooperation partners, which published also similar results for sodium (Marsh et al., 2013a) and iron (Feng et al., 2013). Where applicable, the text and tables in Sect. 7.5.1 are reformulated from Sect. 6.1 in Langowski et al. (2014b), which was originally written by John Plane and Wuhu Feng. The figures in this section are based on the WACCM-Mg data set, but created by myself.

7.5.1 WACCM model

The Whole Atmosphere Community Climate Model (WACCM) a is chemical-dynamical model with a large vertical extent, simulating altitudes from the Earth's surface up to 140 km altitude and is, e.g., described in Garcia et al. (2007), Marsh et al. (2007) and Marsh et al. (2013b). For the investigation of the metal layers, the specified dynamics version SD-WACCM is used. This is nudged by the GEOS5 meteorological data set (see, e.g., Suarez et al., 2008), including parameters like temperature, specific humidity, horizontal winds etc., below 60 km altitude. Nudging means, that

Table 7.7: Magnesium Photolysis Chemistry added into WACCM

No.	Reaction	Rate / s^{-1}	Reference
J1	$Mg + h\nu \rightarrow Mg^+ + e^-$	$j_1 = 4\times10^{-7}$	Swider (1984)

the model is not free running but a weighted averaged with the meteorological data set at certain time steps of the simulation. The weighting of the meteorological data set is typically very small (often less than 1%) in this process. The same version of the SD-WACCM model was also used in Marsh et al. (2013a) and Feng et al. (2013) to successfully simulate the sodium and iron species. For the development of the global model of meteoric magnesium in the upper atmosphere the injection of meteoric constituents into the atmosphere (see Marsh et al. (2013a) and Feng et al. (2013)), and the description of the neutral and ion-molecule chemistry of magnesium in the MLT region (see Whalley and Plane, 2010; Plane and Whalley, 2012) were combined with the WACCM model. A meteoroid injection function (MIF) with similar seasonal variability as those for Fe and Na is used, which has a minimum in spring and a maximum in autumn (see Marsh et al., 2013a, Fig. 1). The fluxes are scaled to match the observed Mg column density, which results in a flux of 1300 atoms $cm^{-2}s^{-1}$ in spring and 2400 $cm^{-2}s^{-1}$ in autumn. The peak ablation height is about 95 km (further details about the ablation process are shown, e.g., in Vondrak et al., 2008). The scaling for Mg leads to similar ablation fluxes of magnesium and sodium in the model. Magnesium has a 16 times higher abundance in chondrites (which are the most typical found meteoroids) than sodium, but the ablation flux of magnesium is by a similar factor 15 smaller than for sodium, so both factors nearly cancel each other out. The magnesium atom, ions and molecule species chemistry is summarized in tables 7.7–7.9.

Table 7.8: Neutral chemistry of magnesium added into WACCM.

No.	Reaction	Rate/cm^3 $molecule^{-1}$ s^{-1}	Reference
R1	$Mg + O_3 \rightarrow MgO + O_2$	$k_1 = 2.3\times10^{-10}\exp(-139/T)$	Plane and Helmer (1995)
R2	$MgO + O \rightarrow Mg + O_2$	$k_2 = 6.2\times10^{-10}\times(T/295)^{0.167}$	Whalley and Plane (2010)
R3	$MgO + O_3 \rightarrow MgO_2 + O_2$	$k_3 = 2.2\times10^{-10}\exp(-548/T)$	Plane and Helmer (1995)
R4	$MgO_2 + O \rightarrow MgO + O_2$	$k_4 = 7.9\times10^{-11}\exp(T/295)^{0.167}$	Whalley and Plane (2010)
R5	$MgO + H_2O + M \rightarrow MgO_2H_2 + M$	$k_5 = 1.1\times10^{-26}\times(T/200)^{-1.59}$	Rollason and Plane (2001)
R6	$MgO_3 + H_2O \rightarrow MgO_2H_2 + O_2$	$k_6 = 1.0\times10^{-12}$	Rollason and Plane (2001)
R7	$MgO + O_2 + M \rightarrow MgO_3 + M$	$k_7 = 3.8\times10^{-29}\times(T/200)^{-1.59}$	Rollason and Plane (2001)
R8	$MgO + CO_2 + M \rightarrow MgCO_3 + M$	$k_8 = 5.9\times10^{-29}\times(T/200)^{-0.86}$	Rollason and Plane (2001)
R9	$MgCO_3 + O \rightarrow MgO_2 + CO_2$	$k_9 = 6.7\times10^{-12}$	Plane and Whalley (2012)
R10	$MgO_2 + O_2 + M \rightarrow MgO_4 + M$	$k_{10} = 1.8\times10^{-26}\times(T/200)^{-2.5}$	Plane and Whalley (2012)
R11	$MgO_4 + O \rightarrow MgO_3 + O_2$	$k_{11} = 8.0\times10^{-14}$	Plane and Whalley (2012)
R12a	$MgO_2H_2 + H \rightarrow MgOH + H_2O$	$k_{12} = 1.0\times10^{-11}\times\exp(-600/T)$	Plane and Whalley (2012)
R12b	$MgOH + H \rightarrow Mg + H_2O$	faster than R12a	see text
R13	$MgOH + MgOH \rightarrow Mg_2O_2H_2$	$k_{13} = 9.0\times10^{-10}$	see text
R14	$MgO_3 + H \rightarrow MgOH + O_2$	$k_{14} = 2.0\times10^{-12}$	Plane and Whalley (2012)
R15	$MgO_3 + O \rightarrow MgO_2 + O_2$	$k_{15} = 1.0\times10^{-13}$	Plane and Whalley (2012)

The following statement for the reactions in table 7.8 was noted in Langowski et al. (2015b) and is quoted directly to prevent confusions due to the reformulation: "An important magnesium reservoir on the underside of the Mg layer is $Mg(OH)_2$. This is reduced back to Mg via MgOH by reaction with H atoms (R12a). The

subsequent reaction with H atoms (R12b) is most likely faster than R12a, and so the reaction rate is not explicitly [listed; (verb was missing here)] in table 7.8 (Plane and Whalley, 2012). The polymerization of MgO_2H_2 to form meteoric smoke is parameterized by a dimerization rate coefficient (reaction R13 in table 7.8). This reaction should be essentially at the high pressure limit even in the upper mesosphere because of the large number of atoms in the dimer and the large binding energy (268 kJ mol^{-1}, calculated at the B3LYP/6-311+g(2d,p) level of electronic structure theory using the Gaussian 09 program suite (see, e.g., Frisch et al., 2009). The capture rate coefficient is then increased to 9×10^{-10} cm^3 $molecule^{-1}$ s^{-1} to account for the concentration of other metallic species (e.g. $NaHCO_3$, FeOH etc) with which MgO_2H_2 can also polymerize." The here used model output is simulated for the

Table 7.9: Ion–molecule chemistry of magnesium added into WACCM.

No.	Reaction	Rate/cm^3 $molecule^{-1}$ s^{-1}	Reference
R16	$Mg^+ + O_3 \rightarrow MgO^+ + O_2$	$k_{16} = 1.2 \times 10^{-9}$	Whalley et al. (2011)
R17	$MgO^+ + O \rightarrow Mg^+ + O_2$	$k_{17} = 5.9 \times 10^{-10}$	Whalley and Plane (2010)
R18	$MgO^+ + O_3 \rightarrow Mg^+ + 2\ O_2$	$k_{18} = 1.8 \times 10^{-10}$	Whalley et al. (2011)
R19	$MgO^+ + O_3 \rightarrow MgO_2^+ + O_2$	$k_{19} = 3.3 \times 10^{-10}$	Whalley et al. (2011)
R20	$Mg^+ + N_2 + M \rightarrow MgN_2^+ + M$	$k_{20} = 1.8 \times 10^{-30}(T/200)^{-1.72}$	Whalley et al. (2011)
R21	$Mg^+ + O_2 + M \rightarrow MgN_2^+ + M$	$k_{21} = 2.4 \times 10^{-30}(T/200)^{-1.65}$	Whalley et al. (2011)
R22	$MgN_2^+ + e^- \rightarrow Mg + N_2$	$k_{22} = 3.0 \times 10^{-7}(T/200)^{-0.5}$	Plane and Whalley (2012)
R23	$MgCO_2^+ + e^- \rightarrow Mg + CO_2$	$k_{23} = 3.0 \times 10^{-7}(T/200)^{-0.5}$	Plane and Whalley (2012)
R24	$MgH_2O^+ + e^- \rightarrow Mg + H_2O$	$k_{24} = 3.0 \times 10^{-7}(T/200)^{-0.5}$	Plane and Whalley (2012)
R25	$MgO^+ + e^- \rightarrow Mg + O$	$k_{25} = 3.0 \times 10^{-7}(T/200)^{-0.5}$	Plane and Whalley (2012)
R26	$MgO_2^+ + e^- \rightarrow Mg + O_2$	$k_{26} = 3.0 \times 10^{-7}(T/200)^{-0.5}$	Plane and Whalley (2012)
R27	$Mg^+ + e^- \rightarrow Mg + h\nu$	$k_{27} = 1.0 \times 10^{-12}$	Plane and Helmer (1995)
R28	$Mg + O_2^+ \rightarrow Mg^+ + O_2$	$k_{28} = 1.2 \times 10^{-9}$	Rutherford et al. (1971)
R29	$Mg + NO^+ \rightarrow Mg^+ + NO$	$k_{29} = 8.2 \times 10^{-10}$	Rutherford et al. (1971)
R30	$MgN_2^+ + O_2 \rightarrow MgO_2^+ + N_2$	$k_{30} = 3.5 \times 10^{-12}$	Whalley et al. (2011)
R31	$MgO_2^+ + O \rightarrow MgO^+ + O_2$	$k_{31} = 6.5 \times 10^{-10}$	Whalley and Plane (2010)
R32	$Mg^+ + H_2O + M \rightarrow MgH_2^+ + M$	$k_{32} = 2.3 \times 10^{-28}(T/200)^{-2.53}$	Martínez-Núñez et al. (2010)
R33	$Mg^+ + CO_2 + M \rightarrow MgCO_2^+ + M$	$k_{33} = 4.6 \times 10^{-29}(T/200)^{-1.42}$	Whalley and Plane (2010)

years 2005 to 2011. The year to year differences are rather small compared to the differences between the model and the SCIAMACHY results, which will be shown in the next section, and every individual year from the sample would lead to similar conclusions in the comparison.

7.5.2 Comparison of WACCM model and SCIAMACHY measurement results

Figure 7.17 shows the monthly mean results for 7 years of WACCM simulations of Mg from 2005 to 2011 and Fig. 7.18 shows the same for Mg^+. Figure 7.19 shows the VCDs for both Mg and Mg^+. Mg shows a clear seasonal cycle with a winter maximum which is most pronounced at the poles. The peak altitude is nearly constant, but small seasonal variations can be found. For example, in February the peak altitude at the poles is roughly 85 km, while it is 5 km higher in the equatorial region. Mg^+ shows a seasonal cycle with a summer maximum. In addition, there is an increased Mg^+ density at the poles even in the winter hemisphere. The Mg^+ peak altitude is close to 95 km and shows no strong variation with latitude and time. The seasonal variations in the VCD profiles for Mg and Mg^+ are similar to Na and Na^+ profiles (Marsh et al., 2013a, Fig. 3). In contrast to the SCIAMACHY

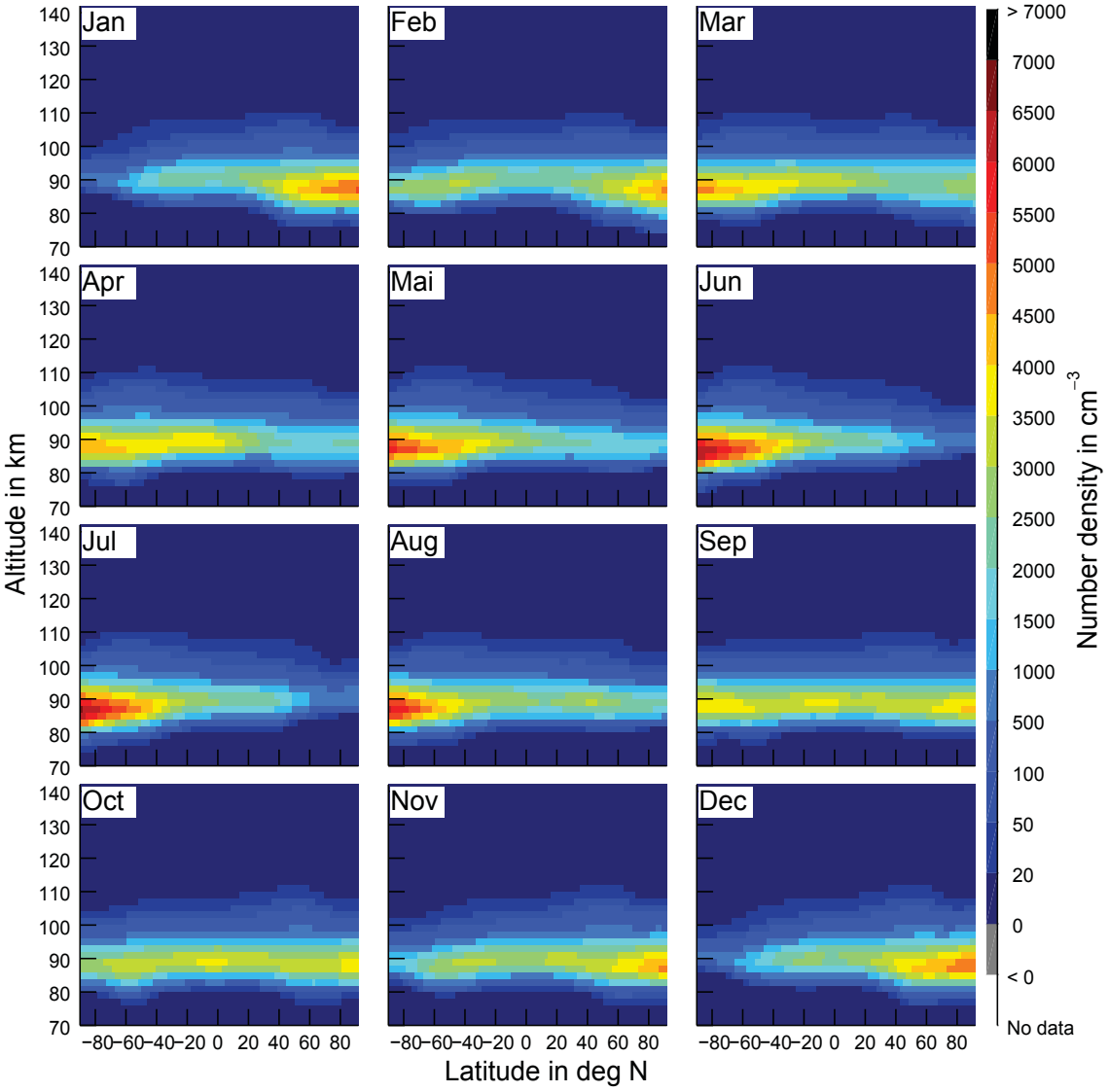

Fig. 7.17: Monthly mean WACCM results for Mg averaged over 7 years from 2005 to 2011. The model results show a clear seasonal cycle with a winter maximum, most pronounced at high latitudes.

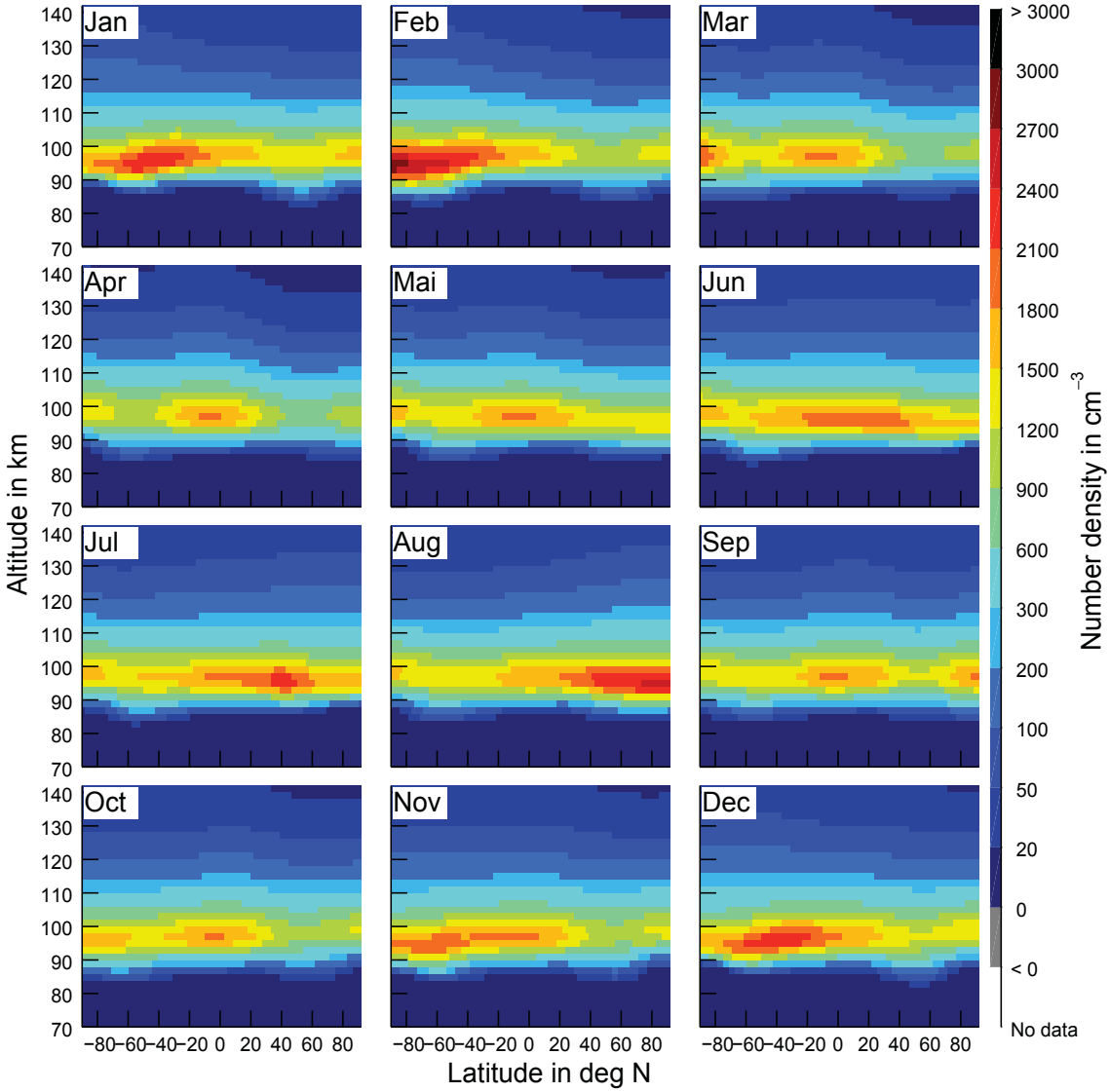

Fig. 7.18: Monthly mean WACCM results for Mg^+ averaged over 7 years from 2005 to 2011. The model results show a clear seasonal cycle with a summer maximum, most pronounced at high latitudes. There is also a second smaller maximum at high latitudes for the winter hemisphere.

measurements, which are made at one particular local time, daily averaged output is used for WACCM. We also collocated the WACCM data set to SCIAMACHY local time and latitudinal coverage, but found only small differences to the non collocated daily average mean of WACCM. There is a diurnal variation of the vertical column density in the WACCM data, but nearly no diurnal variation for the vertical profile shape. Figure 7.20 and Fig. 7.21 show the direct comparison of SCIAMACHY and WACCM annual means for Mg and Mg^+. We used the collocated data for both plots, since these are the only plots with significant differences between not-collocated daily means and collocated data, for latitudes higher than 40°. Figures 7.22 and 7.23 show the comparison for the vertical profile only.

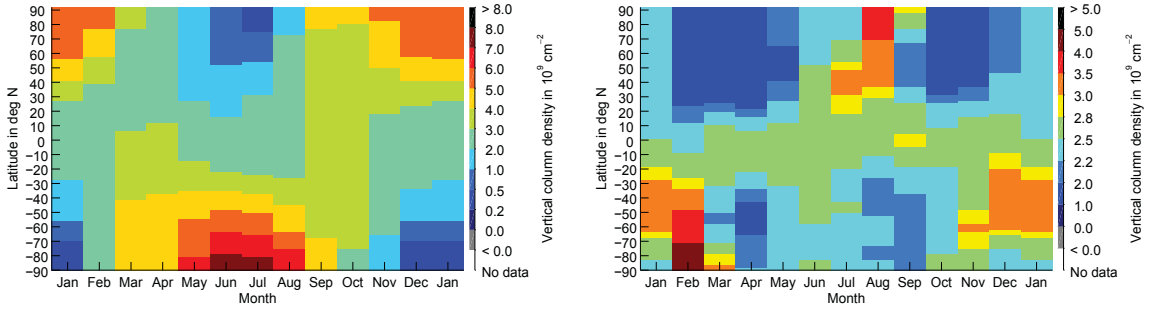

Fig. 7.19: WACCM VCD timelines for different latitudes for Mg (left) and Mg^+ (right). Mg VCD varies between $0.5 \times 10^9\,\mathrm{cm}^{-2}$ and $8 \times 10^9\,\mathrm{cm}^{-2}$ and shows a clear seasonal cycle with a winter maximum, which is most pronounced at the poles. Mg^+ VCD varies from $1 \times 10^9\,\mathrm{cm}^{-2}$ to $5 \times 10^9\,\mathrm{cm}^{-2}$ with most VCD values between $2 \times 10^9\,\mathrm{cm}^{-2}$ and $3 \times 10^9\,\mathrm{cm}^{-2}$. A clear seasonal cycle with a summer maximum for Mg^+ is observed. SCIAMACHY results of VCD are shown in Figs. 7.7 and 7.12.

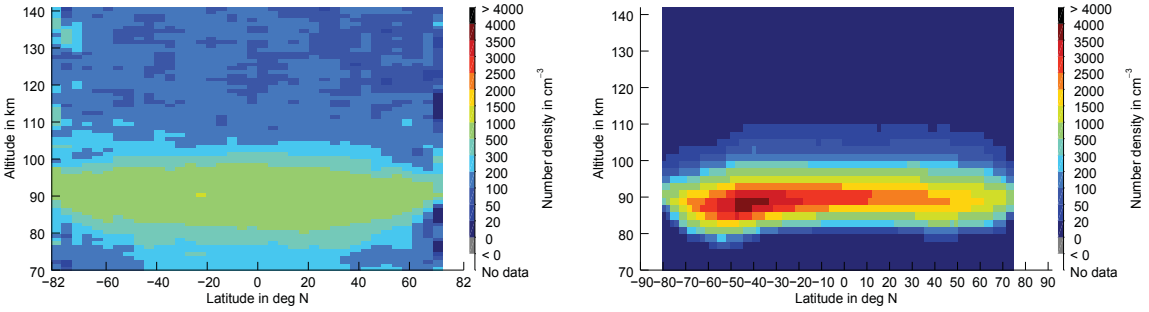

Fig. 7.20: Annual means of Mg number density for SCIAMACHY (left) and WACCM (right). SCIAMACHY measures smaller peak density having a wider vertical profile than WACCM. The averaging for WACCM simulations is restricted to match the time periods, where SCIAMACHY measurements are available.

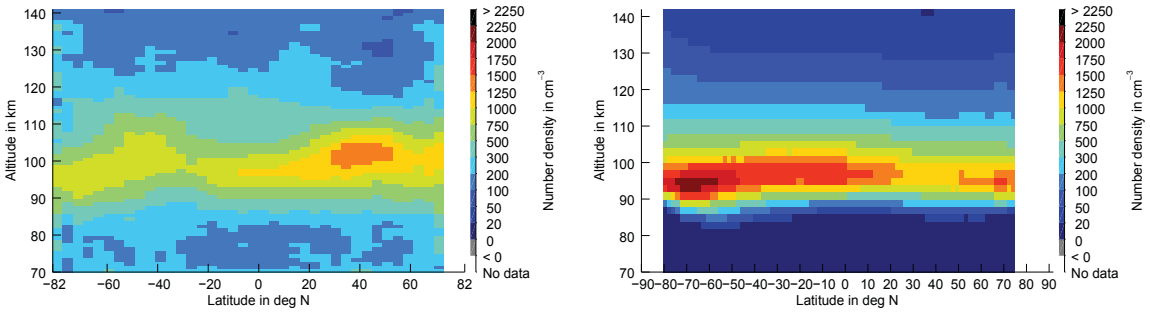

Fig. 7.21: Annual means of Mg^+ number density for SCIAMACHY (left, 280.4 nm) and WACCM (right). SCIAMACHY shows a wider vertical profile, and a stronger latitudinal dependence of the peak altitude. The peak density of both is in very good agreement. The averaging for WACCM simulations is restricted to match the time periods, where SCIAMACHY measurements are available.

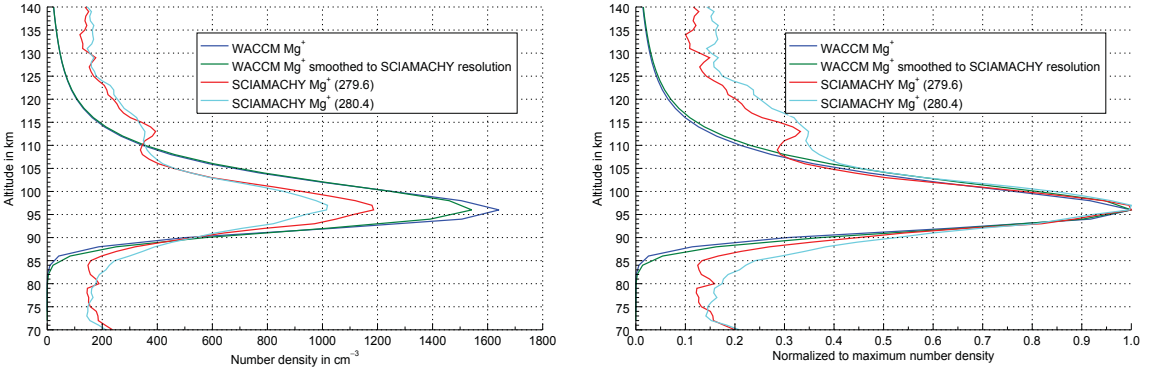

Fig. 7.23: Vertical profiles of Mg^+ for SCIAMACHY and WACCM at the equator. The right plot shows the profiles normalized to the maximum of each profile, to better compare the shape of the profiles. The VCD between 70 and 140 km at the equator is 2.7×10^9 cm^{-2} for SCIAMACHY (both 279.6 nm and 280.4 nm) and 2.4×10^9 cm^{-2} for WACCM. Note, that the statistical error for SCIAMACHY estimated by Langowski et al. (2014a) (Figs. 22 and 23) is $\approx \pm 150$ cm^{-3} for the 279.6 nm line and $\approx \pm 250$ cm^{-3} for the 280.4 nm line.

In the annual mean profile, the peak density of Mg for WACCM is higher than for SCIAMACHY. However, the peak width of the SCIAMACHY profile is larger than for WACCM and the VCDs are similar, which is expected, because the Mg MIF is scaled (see above). The peak altitude of WACCM (88 km) and SCIAMACHY (90 km) results agree well considering the 3.3 km step size in the SCIAMACHY measurements and the vertical sampling step size of 2 km for WACCM. One explanation for the wider SCIAMACHY peak is a strong Ring effect influence (Grainger and Ring, 1962) for the Mg spectral line, which has not adequately been corrected. At the lower and upper peak edge, the SCIAMACHY number densities are not zero. This may be related to the requirement in the retrievals that the densities must be positive, which is needed for stability in the retrieval. However, the densities

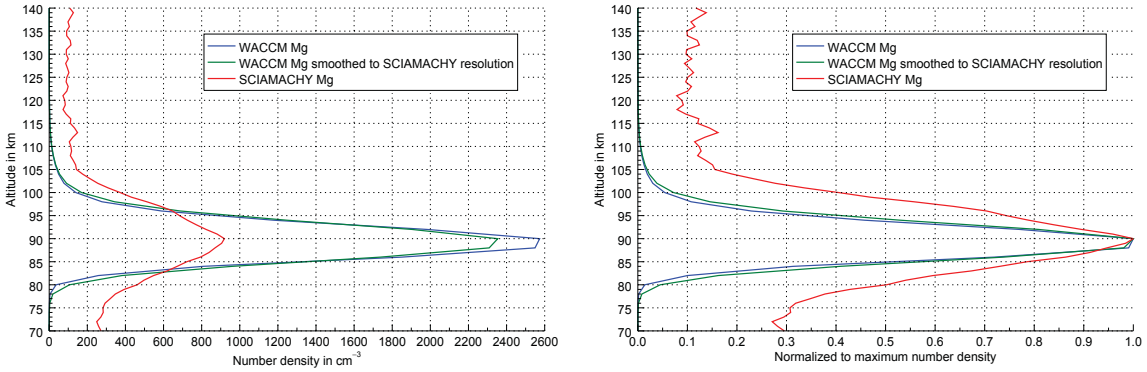

Fig. 7.22: Equatorial annual mean vertical profile of Mg for SCIAMACHY and WACCM. The right plot shows the profiles normalized to the maximum of each profile, to better compare the shape of the profiles. The smoothed profile is smoothed with a triangular function with a base width of 4 km. The VCD between 70 and 140 km is $2.3 \times 10^9\,\mathrm{cm}^{-2}$ for SCIAMACHY and $2.5 \times 10^9\,\mathrm{cm}^{-2}$ for WACCM. Note, that the statistical error for SCIAMACHY estimated by Langowski et al. (2014a) (Fig. 21) is $\approx \pm 200\,\mathrm{cm}^{-3}$ for the low-density region above 100 km and below 80 km and $\approx \pm 400\,\mathrm{cm}^{-3}$ at 90 km.

are smaller than the single day error there. The agreement of WACCM and SCIAMACHY is better for Mg^+. The WACCM Mg^+ peak and column densities are of similar magnitude as the SCIAMACHY results and also show the same seasonal variability with a summer maximum. There are also small differences: the maximum in the WACCM data is not reduced when passing the equatorial region during equinox and also has a stronger extension to the higher latitudes in summer. Furthermore, a second maximum appears in the winter hemisphere at high latitudes, which is not covered by SCIAMACHY measurements and so is not observed. The WACCM results also do not show the strong seasonality and latitude dependence of the peak altitude, that can be found in the SCIAMACHY results. The peak altitude at 95 km is in slightly better agreement with the rocket data than with SCIAMACHY. One reason for some of these discrepancies between model and observations is, that WACCM does not include the Lorentz $\vec{v} \times \vec{B}$-force, as discussed in Feng et al. (2013). In a global average of the profile, however, the differences are small. The lower peak edge of Mg^+ is in better agreement than for Mg, since, e.g., the Ring effect influence is much weaker for Mg^+ than for Mg. The agreement of the peak edge becomes worse with larger distance from the peak altitude. Nevertheless, the FWHM of the peak is in good agreement.

7.5.3 Summary on the investigations of Mg/Mg^+

Monthly averaged global vertical and latitudinal density distributions of Mg and Mg^+ retrieved from the SCIAMACHY limb MLT measurements from 2008 to 2012 have been shown in this chapter and were compared to observations in the literature as well as WACCM-Mg simulation results. Mg shows highly variable results and variability is mainly caused by the measurements error. No clear seasonal cycle is

observed for Mg, but densities are generally lower at latitudes above 40°. The peak altitude of Mg is nearly constant at ≈ 90 km for all latitudes and times. Mg^+ shows a clear seasonal cycle with a summer maximum in the peak density, most pronounced at mid latitudes. Furthermore, Mg^+ shows a seasonal variation of the peak altitude with higher altitudes in the summer, as well as a latitudinal dependence of the peak altitudes with up to $5 - 10$ km higher altitudes at mid latitudes compared to the equatorial peak altitude at 95 km.

The SCIAMACHY results are reasonably consistent with previously published rocket profiles. There is a good agreement for the peak density values as well as the seasonality for Mg^+. However, the rocket measurements do not show the latitudinal dependence of the peak altitude, but typically show peak altitudes at slightly below 95 km.

The SCIAMACHY results were compared with WACCM model results. There is a disagreement in the seasonal variation for Mg between SCIAMACHY and WACCM results, with WACCM showing a clear seasonal cycle with a winter maximum at the poles. Furthermore, the WACCM peak densities are roughly a factor 2 bigger than the SCIAMACHY peak densities. The peak altitude and peak shape are in good agreement with SCIAMACHY, showing a slightly wider profile. The combination of higher peak densities for WACCM, but wider peak profile for SCIAMACHY leads to very similar values for the VCDs. The agreement between WACCM and SCIAMACHY is better for Mg^+ than for Mg. Both SCIAMACHY and WACCM show a clear seasonal cycle for Mg^+ with a summer maximum and the peak densities are in good agreement. However, like the rocket measurements, WACCM does not show a latitudinal dependence of the peak altitude. The peak shape of Mg^+ agrees very well for WACCM and SCIAMACHY in the high density region, where the SCIAMACHY error is low.

8 Results of the Na number density retrieval

This section was concurrently worked out with Langowski et al. (2015a). The same data averaging process as for Mg/Mg^+ was used (see Sect. 7.1), but less data is cut off at the highest latitudes, because artifacts can be more easily ruled out for Na, because of the better signal to noise ratio compared to Mg and Mg^+.

8.1 Seasonal variations of Na

The monthly mean Na number densities as a function of latitude and altitude are shown in Fig. 8.1, where the equally weighted mean of the retrieval results from both Na D lines ($\frac{1}{2}(D_1+D_2)$) is used. The altitude of the density maximum is at

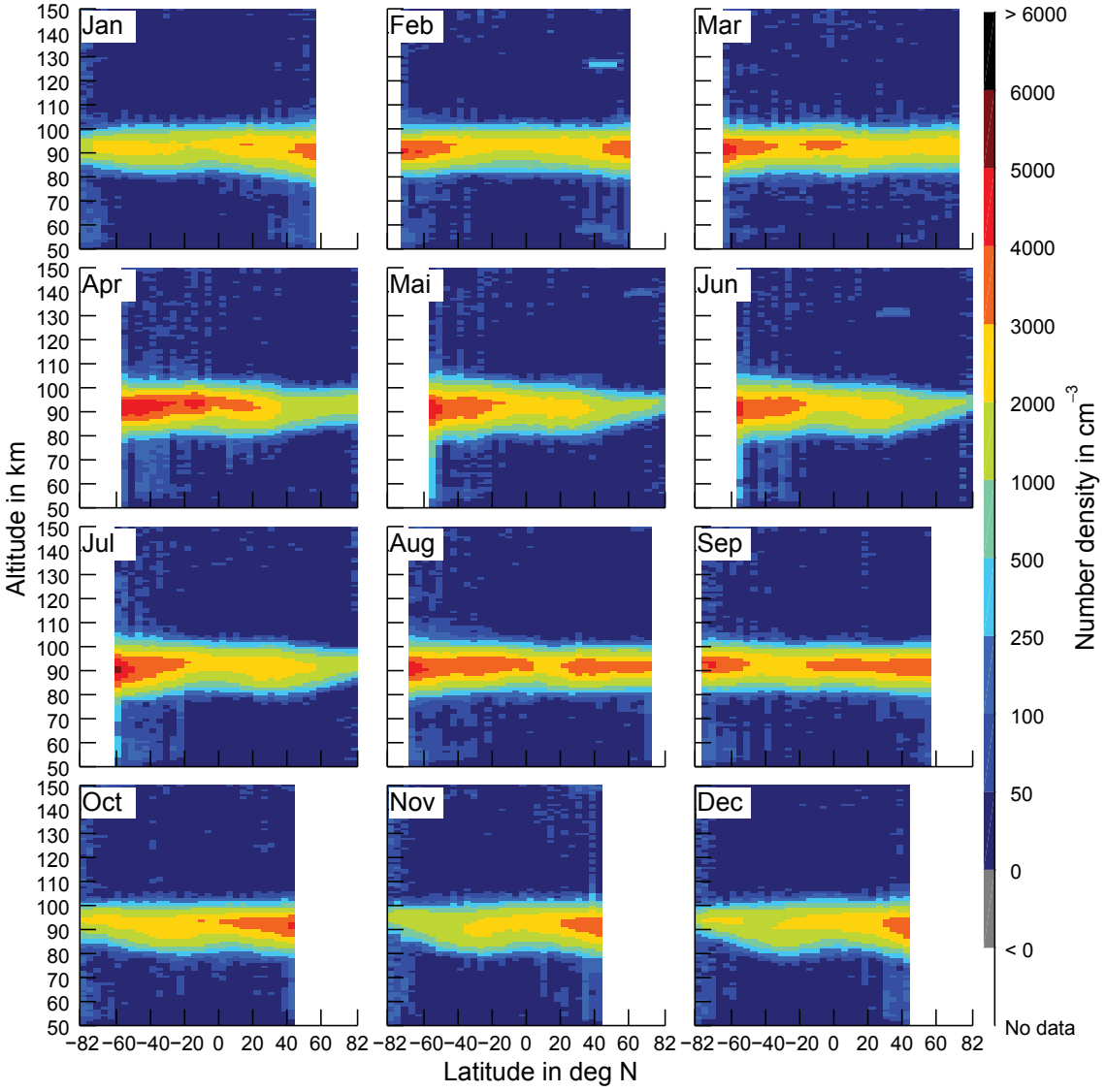

Fig. 8.1: Latitude-altitude distribution of the monthly mean Na densities. The average of the results for both Na lines is used.

around 92 km and varies only by a few km during the year. Na shows a seasonal cycle in number density with a winter maximum. The peak densities in the winter mid latitudes are up to 6000 cm^{-3}. In the summer, the maximum density decreases to only $\approx 1000\,cm^{-3}$ at high latitudes. The annual mean is shown in Fig. 8.2 and shows an average peak density of roughly 2000 cm^{-3} to 4000 cm^{-3} apart from the high latitudes, which are only covered in the summer period and, therefore, only show the small summer densities. The results for both hemispheres are nearly symmetric.

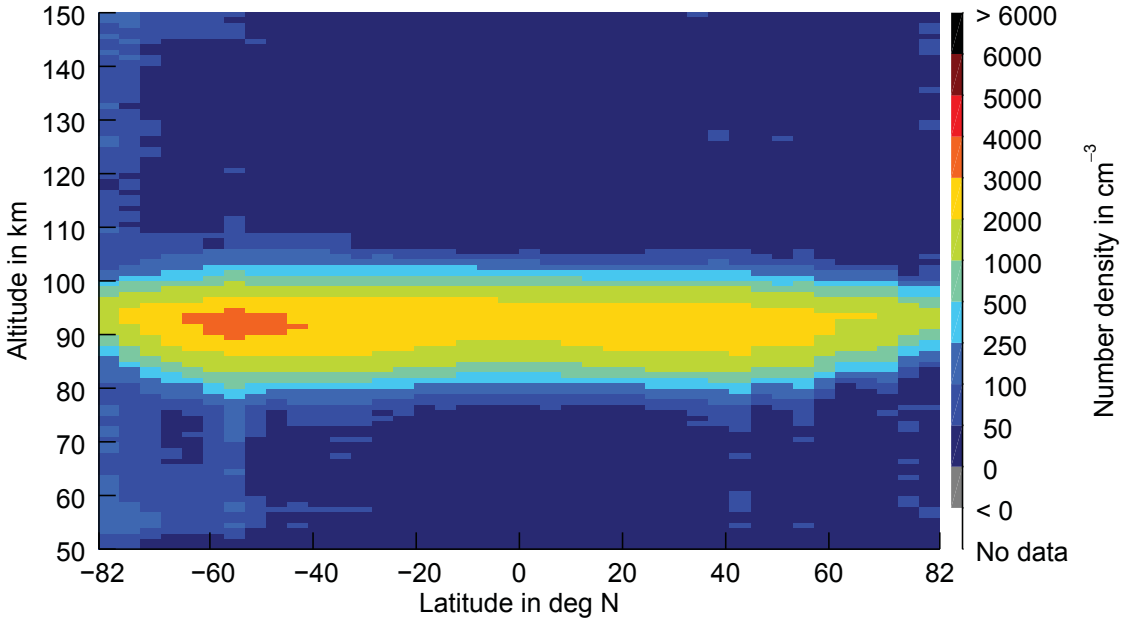

Fig. 8.2: Annual mean Na distribution. Note that the high latitudes are only covered in the hemispheric summer.

The seasonal variation of the vertical profile for low, mid and high latitudes is shown in Fig. 8.3, and vertical profiles for selected latitudes in July are shown in Fig. 8.4. At low latitudes, a semi-annual variation with maxima in March and September is observed. This variation is well correlated with the semi-annual variation in temperature (see, e.g., von Savigny and Lednyts'kyy, 2013), which shows a maximum during this time. The semi-annual oscillation for Na was also found in model studies by Marsh et al. (2013a). The Na profile has a reduced width at high northern latitudes in summer (this is in agreement with Fan et al., 2007), which may correspond to depletion of Na there, e.g., due the occurrence of noctilucent clouds or the reaction into molecular sink species due to low temperatures (see, e.g., Marsh et al. (2013a)). Figure 8.5 shows the vertical column densities (VCDs) for different months and latitudes, which also show the seasonal cycle with a summer minimum of slightly below $1 \times 10^9\,cm^{-2}$ at high latitudes and up to $1 \times 10^{10}\,cm^{-2}$ at the highest latitudes covered in the winter hemisphere.

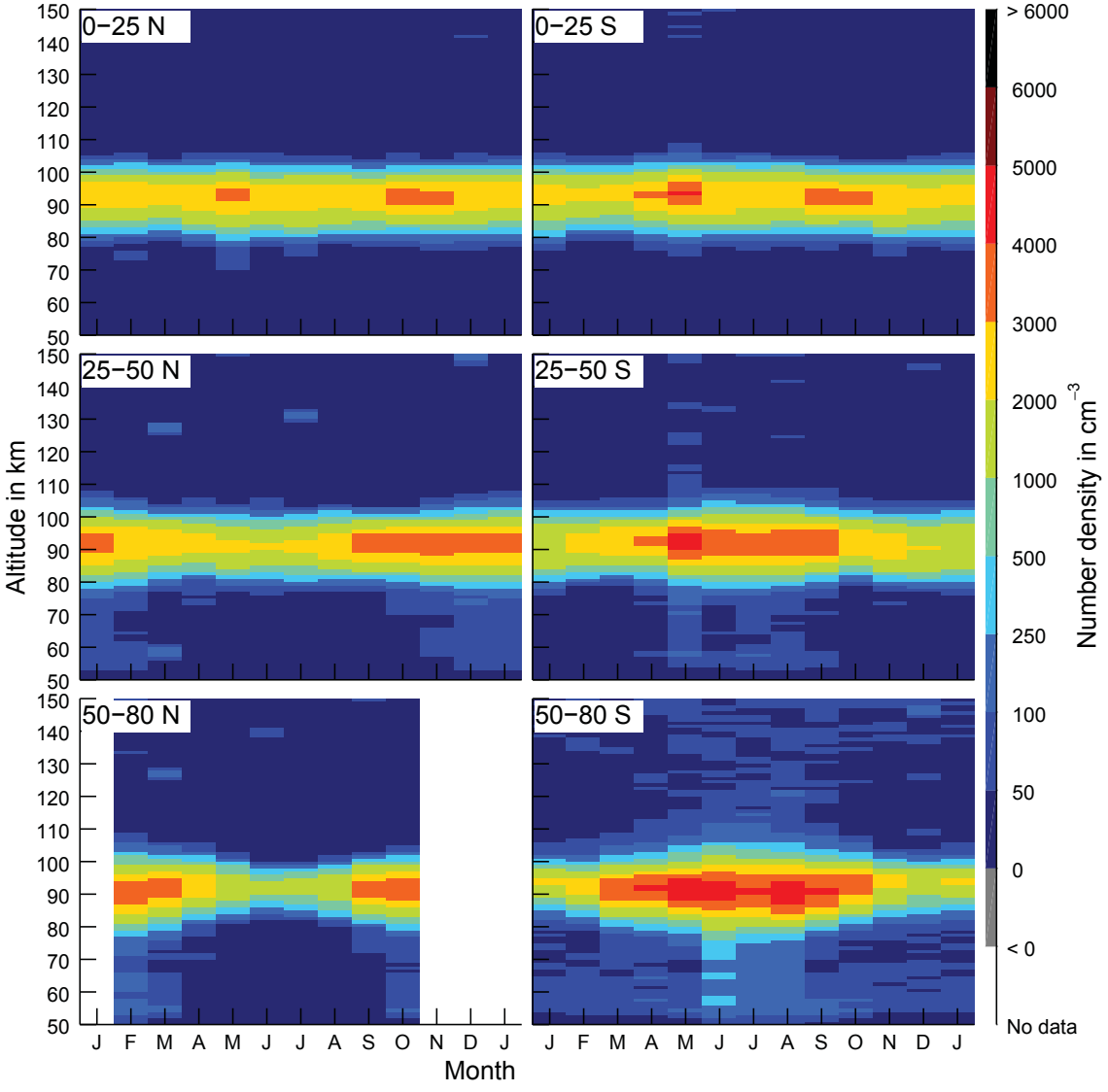

Fig. 8.3: Seasonal variation of the vertical Na density profile for low, mid and high latitudes.

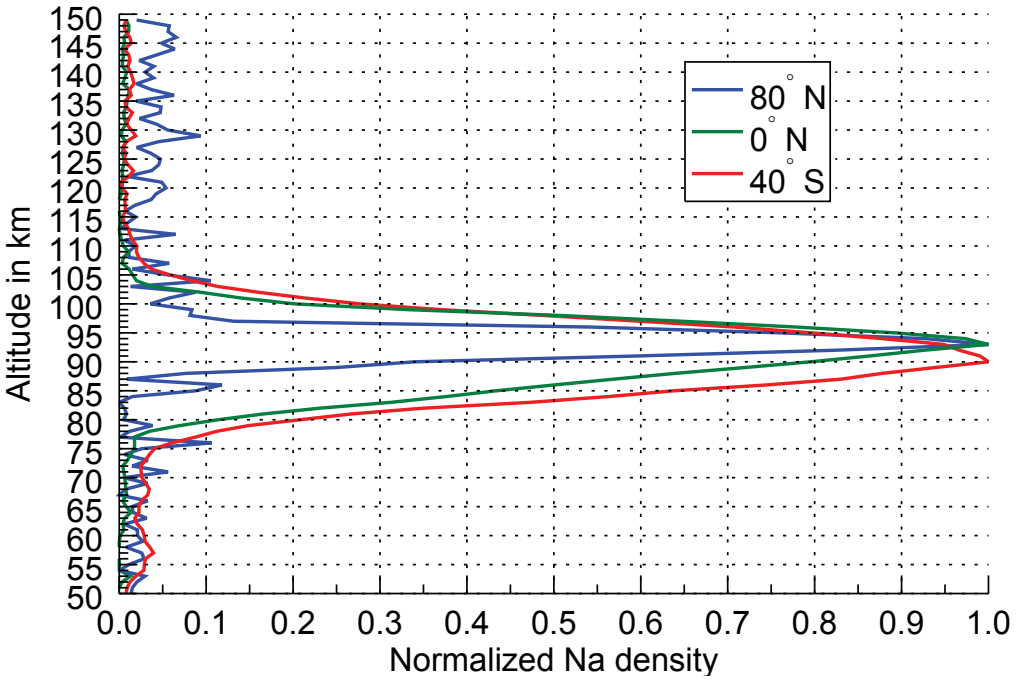

Fig. 8.4: Normalized vertical Na profile at selected latitudes in July. The densities are normalized to the peak value.

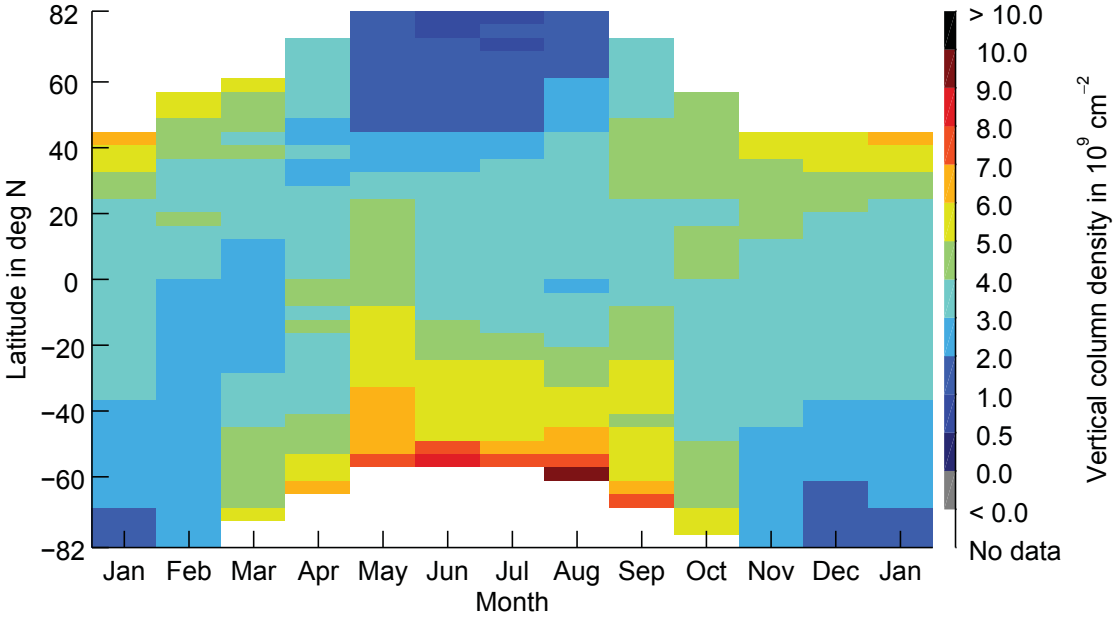

Fig. 8.5: Seasonal variation of the Na vertical column densities (VCDs).

8.2 Comparison with other measurements

In the following, the SCIAMACHY Na data set is compared to ground and space based measurements. Figure 8.6 shows a comparison to ground-based ALOMAR (Arctic Lidar Observatory for Middle Atmosphere Research) lidar-measurements at Andøya, Norway, at 69° recorded between 2008 and 2013. The ALOMAR data set was provided by Tim Dunker and Ulf-Peter Hoppe from the Arctic University of Norway in Tromsø. The majority of these data have been published by Dunker et al. (2015) and show a mean peak altitude of around 92 km, which is in good agreement with the peak altitudes found in the SCIAMACHY data. The SCIAMACHY error bars are the standard deviation of measurements in each month. The errors bars of the respective daily mean Na column density measured by the lidar denote the standard deviation, which is a measure of the geophysical variation on that day. The

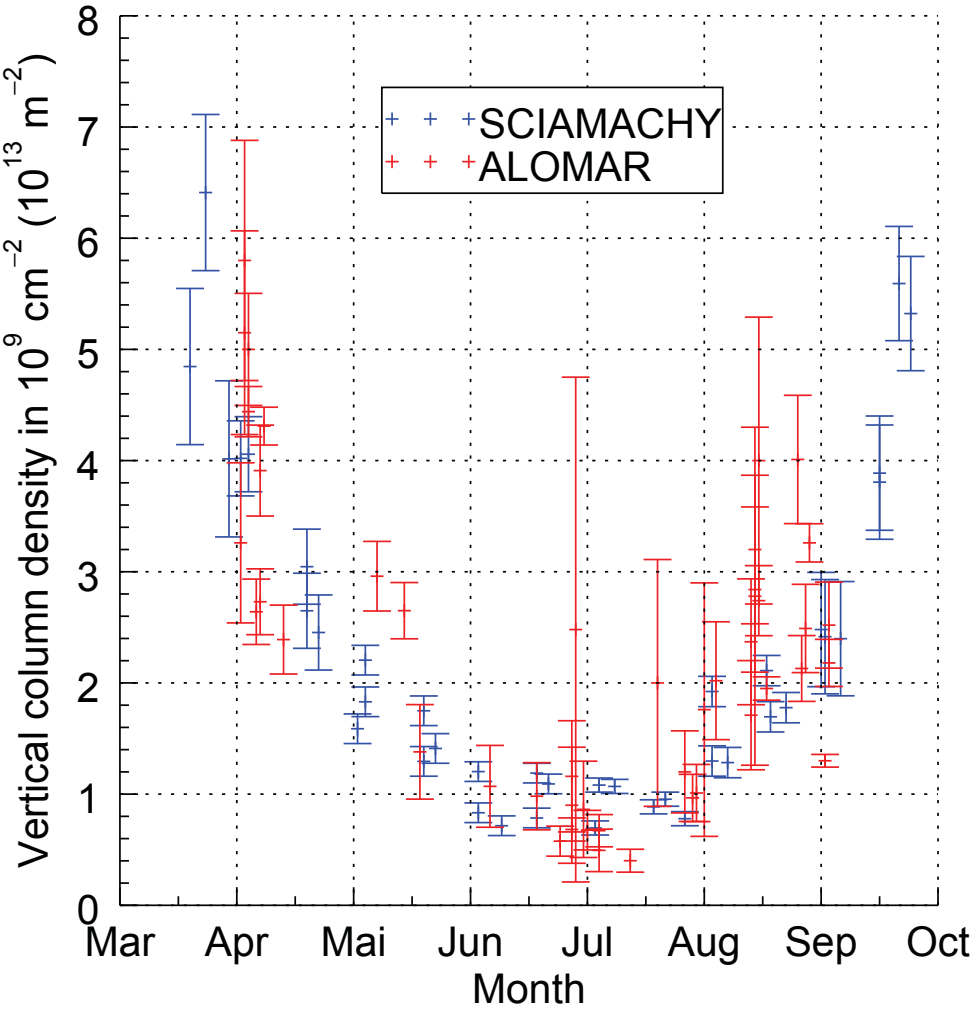

Fig. 8.6: Comparison of the SCIAMACHY Na VCD at 69°N with lidar observations at Andøya, Norway as a function of time. The ALOMAR data set was provided by Tim Dunker and Ulf-Peter Hoppe.

SCIAMACHY and the ALOMAR results are in the same order of magnitude (a factor 3, for the largest differences) and are overall in good agreement. In June and July the ALOMAR measurements show several times smaller, similarly large and also significantly larger VCDs than the SCIAMACHY measurements. The reason for the significantly larger number densities lies in the appearance of high density sporadic layers during the observation time, which are typically observed at night (see, e.g., Heinrich et al., 2008) and not during SCIAMACHY's local observation time. The ground based observations during the time of occurring sporadic layers have a larger error bar in Fig. 8.6, that rather comes from the larger natural variability than from statistical errors. SCIAMACHY does not see those strong variations in summer, as the much larger scanned air volume and the longitudinal averaging lead to more strongly averaged data. On average, larger VCDs are observed in April and in September than in June and July in the ALOMAR data set. This fits well to the seasonality observed by SCIAMACHY.

Figure 8.7 shows a comparison to ground based lidar-measurements by She et al. (2000) at around 40°N. The overall agreement of the seasonal variation and density is quite good.

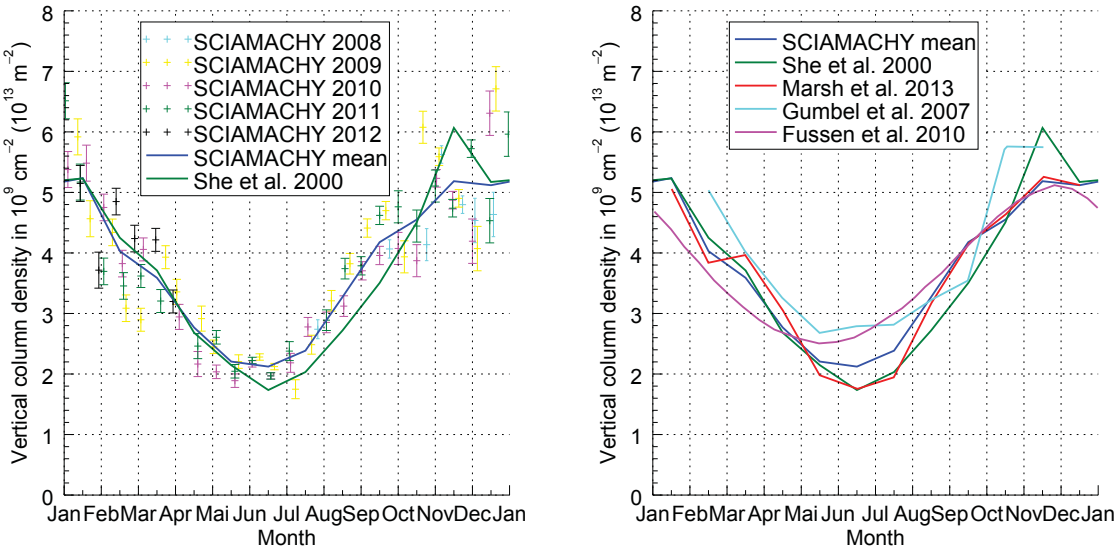

Fig. 8.7: left: Comparison of the SCIAMACHY VCD profile at 40°N and the results from She et al. (2000). right: Comparison of the same profiles with profiles from Fussen et al. (2010) (their Fig. 7) and Marsh et al. (2013a) (their Fig. 10).

The peak altitude obtained by She et al. (2000) is also 90-92 km. The full width at half maximum (FWHM) of the Na peak found be She et al. (2000) is 9 to 10 km (not shown here), which is slightly smaller than the 11 to 14 km for SCIAMACHY. However, SCIAMACHY scans a much larger volume of space, which might in part explain this difference. The largest differences between SCIAMACHY and the the values presented by She et al. (2000) are about 20% from July to November. The data set of She et al. (2000) was also used for comparison with other global satellite measurements by Fussen et al. (2010) (their Fig. 7) and the WACCM model by Marsh et al. (2013a) (their Fig. 10). There is a good qualitative and quantitative

global agreement in seasonal variation between the SCIAMACHY results presented here and the measurement results reported by Fussen et al. (2010) and Hedin and Gumbel (2011) as well as with model results by Marsh et al. (2013a). The reduction of the profile width in the summer high northern latitudes is tentatively explained by Marsh et al. (2013a) by an increased ionization rate at the upper edge of the vertical profile and the Na reaction into $NaHCO_3$. The latter is favored by low temperatures in the polar summer mesopause region at the bottom edge of the vertical profile. A further reduction process at the bottom edge of the vertical profile, which is, e.g., discussed by Gardner et al. (2005), Fig. 9, is the uptake of atomic Na by NLC or smaller ice particles. This process also depends on low temperatures so that both reduction processes at the lower edge of the vertical profile are correlated and hardly distinguishable.

9 Metal interaction with noctilucent clouds

Parts of this chapter also have been shown in Langowski et al. (2015b,a). As described in Sect. 2.4, the summer polar mesopause is the coldest region in the Earth's atmosphere and the temperatures are low enough, that clouds are regularly formed there. These polar mesospheric clouds (PMCs) between 82-85 km altitude are optically thin, but reflect a small part of the incoming solar radiation. Because of their high altitude, the PMCs are still illuminated for solar zenith angles significantly larger than 90 degrees, which results in reflection of solar radiation to the ground when it is already/still dark at the ground in the night. Therefore, the PMCs are also called noctilucent clouds (NLCs). The NLCs are regularly observed above 50° latitude during the summer. Ice particles are present between 82-88 km. Only the lower part of the vertical ice-particle distribution includes larger ice particles, that lead to a significant backscattered radiation, compared to the Rayleigh-scattering background, which is identified as the NLC. However, a larger fraction of the ice-particle distribution, which includes also ice particles at higher altitudes than the NLCs, is observed as so called polar mesospheric summer echoes (PMSEs) in radar observations (see, e.g. Rapp and Lübken, 2004). These echoes may also result from other atmospheric features, so that not every PMSE is caused by ice particles.

The NLCs were first observed by Leslie (1885), Backhouse (1885) and Jesse (1885) after the volcanic eruption of the Krakatoa volcano in 1883. Afterwards, NLCs have regularly been observed, but were not as bright. It is still an open question, whether the NLCs existed before 1883. As these clouds apparently are only observed since the industrial age, it is in open discussion, whether the NLCs are an indicator for climate change. Long term investigations show a systematical increase of the NLC occurrence rate over the past decades (see, e.g., Deland et al., 2007; Shettle et al., 2009), but recently these observations were called into question by Stevens et al. (2007) due to local time issues with these observations.

First NLC observations with SCIAMACHY have already been reported in 2004 by von Savigny et al. (2004), which was followed by several other publications. NLC signals can also be found in the MLT observations. For the following quantitative analysis of the NLCs, two consecutive algorithms are used. The first algorithm is focused on separating measurements with and without NLCs, while the second algorithm quantifies the strength of the NLC signal, only for the cases where NLCs are detected. Figure 9.1 shows some example plots of the NLC test, which is subsequently discussed.

For both algorithms, a wavelength window (458 nm - 552 nm) is used, which has no strong spectral emission or absorption features. The use of this large wavelength window results in negligibly small statistical errors. Rayleigh single-scattering of air molecules is proportional to the air density, which decreases nearly exponentially with tangent altitude. As the decrease of the Rayleigh scattered background signal with altitude is steep, a cloud can easily be identified by a positive altitude gradient in the averaged signal. However, a certain minimum cloud signal strength is needed, to fulfill this condition, and certain instrumental issues, e.g., different effects of stray light demand better criteria for a robust NLC detection than just a positive gradient. For the MLT measurements in this wavelength region, the signal nearly

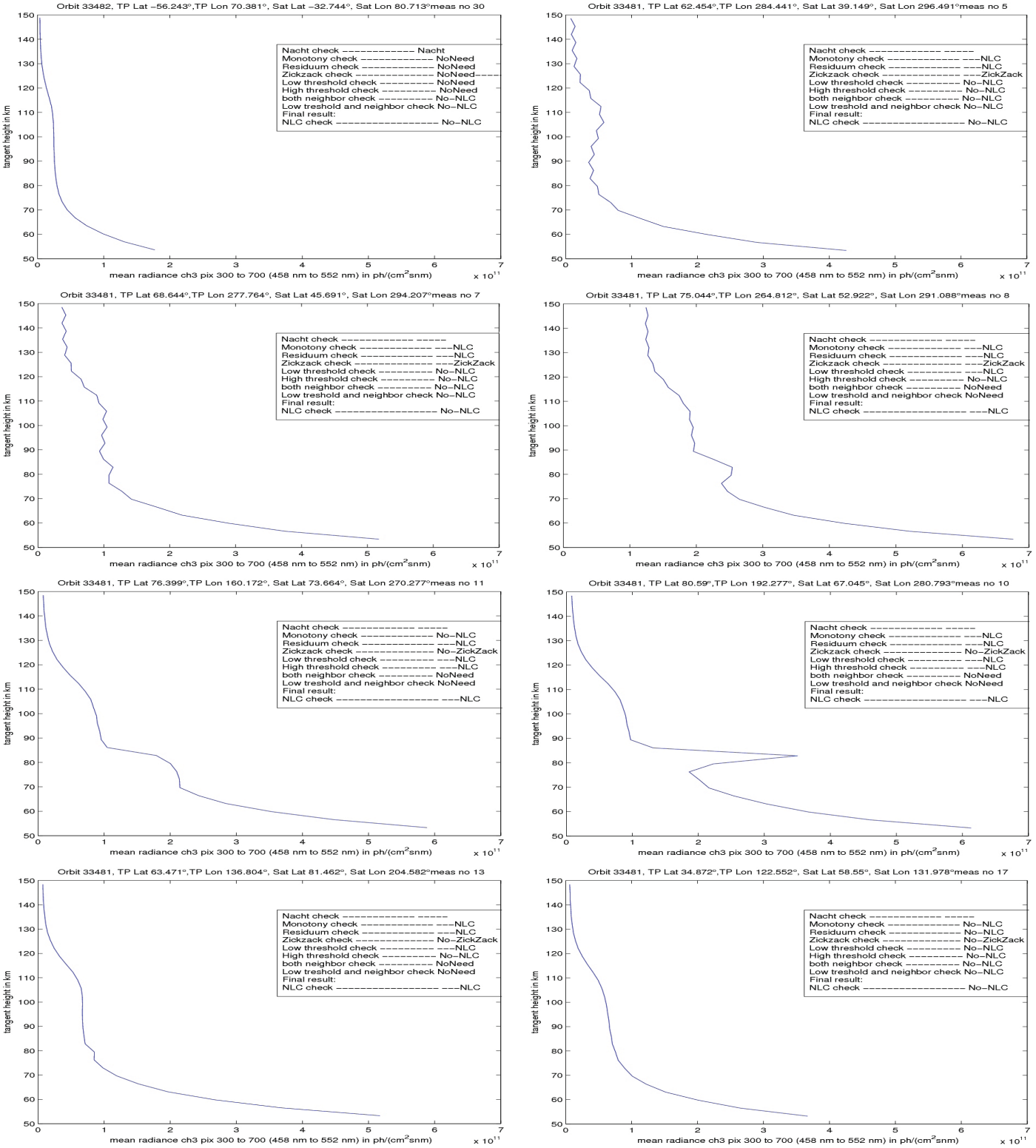

Fig. 9.1: Test for polar mesospheric clouds with different scenarios:
row 1 left: night measurement, right: zigzag no NLC,
r2 l: zigzag not detected too weak NLC, r: zigzag NLC,
r3 l: no monotony but NLC, r: clear NLC,
r4 l: weak NLC, r: no NLC

exponentially falls to a constant value at around 110 km (the reason may be external stray light) and then falls again nearly exponentially for higher altitudes. Around this plateau, value the vertical gradient may even be positive, when no NLC signature in the vertical profile, typically a peak at 84 km tangent altitude, is seen by eye. The criterion used here is to first fit an 8th order polynomial P_8 to the vertical profile M with 30 tangent altitude steps i (starting from the highest altitude) and then form the following quantity:

$$x = \sqrt{\sqrt{\sum_{i=15}^{26} \left(\frac{(P_8(i) - M(i))^2}{M(i)} \right)^4}}, \tag{9.1}$$

which is the residuum of the fit, just with a higher weighting of larger discrepancies than the usual χ^2 approach, leaving out the edge points of the fit, as they usually do not fit well for polynomial fits (steps 15 and 26 correspond to about 70 km and 100 km altitude). The median of this quantity is formed for all measurements of an orbit and a threshold criterion (5×median) has to be passed to identify a NLC. Under the assumption, that there are more non NLC measurements in an orbit than NLC measurements, the median criterion ensures that a non NLC measurement is used to find the threshold. Even for weak NLCs, this method reproduces same result as my check of the vertical profile by eye. However, this method fails when the

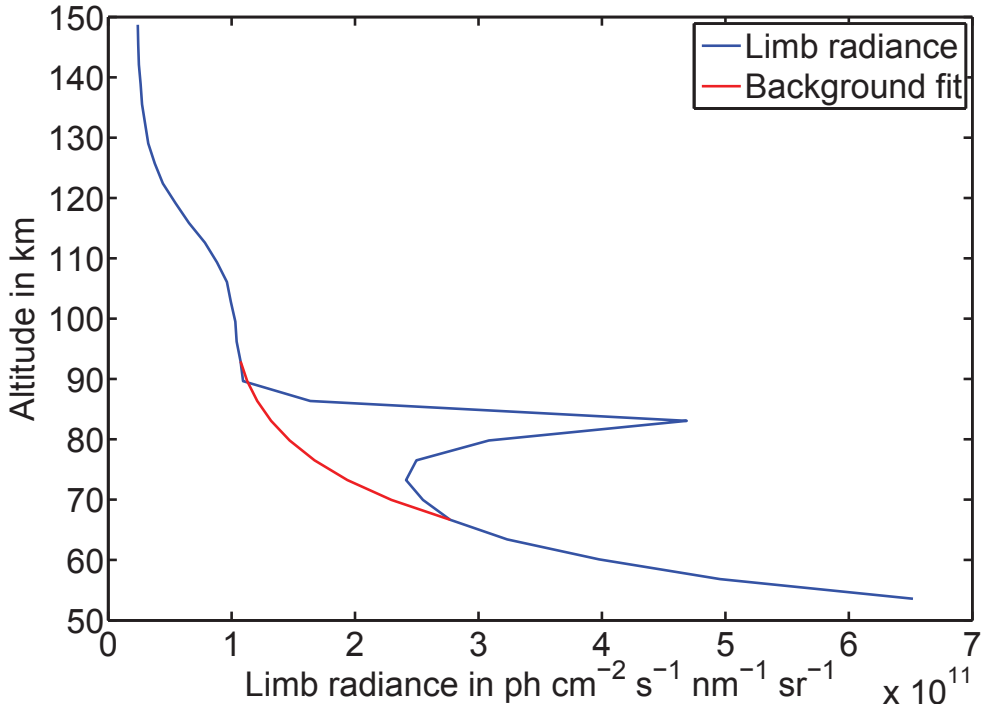

Fig. 9.2: Averaged SCIAMACHY limb MLT signal between 458 and 552 nm showing a polar mesospheric cloud signature at 84 km and a background fit. The occurrence of the noctilucent clouds is quantified as the summed up differences between the signal and the background fit in the vicinity of the NLC signal.

vertical profile shows noise-like patterns. For the first daylight measurements of an orbit, strong stray light signatures can be observes in the dark signal measurement, which lead to a zig-zag pattern in the vertical profile. The reason for this zig-zag pattern apparently is, that the sun is partly in the instruments total clear field of view (TCFOV). For two consecutive limb MLT scans the TCFOV moves slightly downwards and there are several subscans done with the TCFOV moving to the left (for all subscans) or back to the right. A small horizontal offset in the TCFOV, caused by alternating left and right scanning, leads to a zig-zag pattern in the averaged signal, due to the different solar stray light influence. Fortunately, this zig-zig-pattern is easily identified, and if it is present, a simpler test for a threshold value of the gradient, that has to be a bit stronger than the average zig-zig to identify a NLC, is used. As fine tuning is needed to find the right and robust parameters, the results of all tests were plotted and were checked for the daily averaged MLT measurements (roughly $84 \times 50 \approx 4000$ plots, which is still handable number for check by eye).

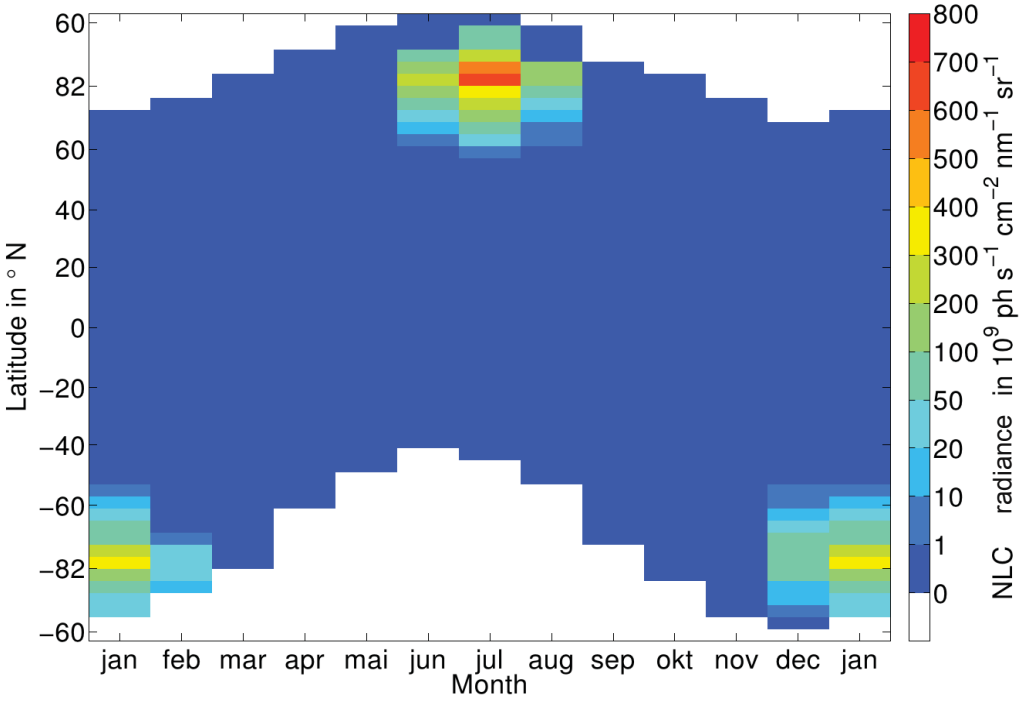

Fig. 9.3: Monthly average NLC radiance in the SCIAMACHY MLT measurements for different latitudes.

To quantify the "strength" of the NLC signal, the background signal in the NLC measurements is determined and the difference of the full signal and the background signal is summed up as a slant column for the tangent altitude region around 84 km. In the following this retrieved quantity is called the NLC radiance. To find the background signal, the edge tangent altitudes, where the signal is not significantly affected by the NLC are determined. The search is limited to the region be-

tween 63 and 90 km, and the upper altitude is reduced/the lower altitude increased, as long as a threshold criterion for the ratio of signals of consecutive tangent altitudes is fulfilled (the chosen threshold is 0.9). 3 data points (the lower altitude value as well as the upper altitude value and its consecutive value (for minimizing the zig-zag-effect)) are used for the fit. A 2nd order polynomial is fitted to the logarithm of the signal as a function of tangent altitude. Figure 9.2 shows a typical background fit for a strong NLC signal. As the algorithm may fail, if no NLC is detected (e.g. in the search of the 2 points of the signals peak edge), the NLC radiance is set to 0, if no NLC was initially detected.

The average NLC signal for the 4 years dataset of MLT measurements is shown in Fig. 9.3. Significant NLC radiances are detected at latitudes higher than 50° in three summer months in both hemispheres (June, July, August in the North and December, January and February in the South). The NLC radiance is largest at the highest latitude and the largest latitudinal spread and radiance is found in July and January.

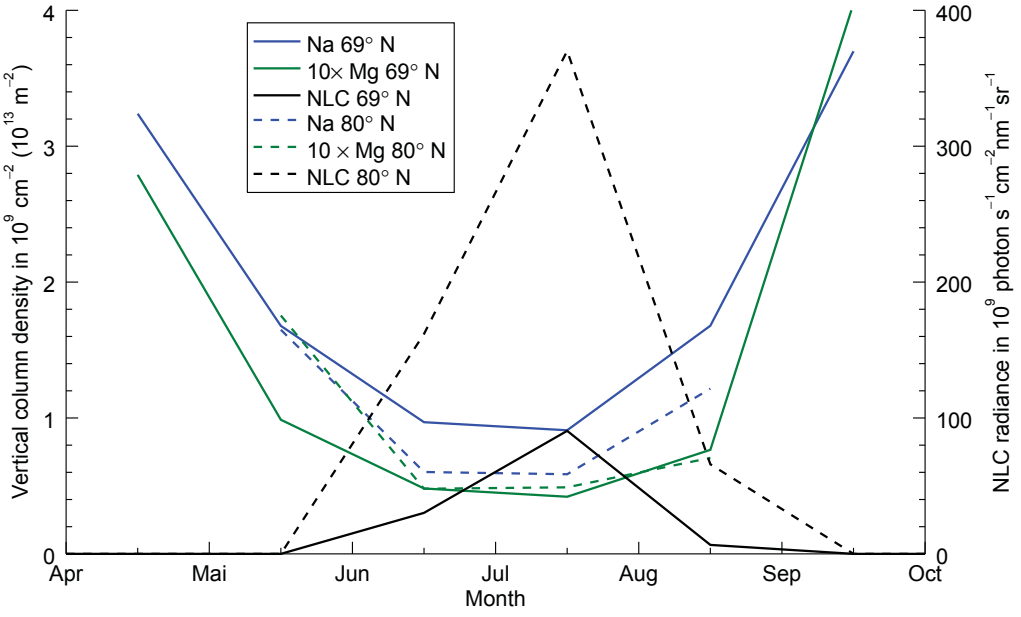

Fig. 9.4: Sodium and Mg number densities and NLC radiance at high northern latitudes.

It is interesting to investigate the observed metal layers for changes during the NLC season, as the neutral metal layers form at around the same altitude as the NLCs and meteoric smoke particles are supposed to be involved in the cloud formation. Model simulations, e.g., by Gardner et al. (2005) show a clear reduction of Na at the lower peak edge and a bite out in Fe density, as the Fe layer peaks at lower altitudes than Na. In contrast to lidar observations, the SCIAMACHY observations cover a much larger air volume and a certain fraction of NLC always covers the field of view. Thus, a direct comparison between observation with and without NLC at nearly the same date, time and latitude can hardly be done. For the two metals observed in this study, Na shows a stronger seasonal variations at high

latitudes than Mg. The VCD for Na and Mg as well as the monthly NLC radiance at the highest covered latitudes of around 80° as well as at around 70° N, a latitude which includes NLCs but also is covered in month where no NLCs are present, are shown in Fig. 9.4.

The Na monthly variability is stronger. For both, Na and Mg the minimum VCD is found in July, where the NLCs are strongest. However, the seasonal

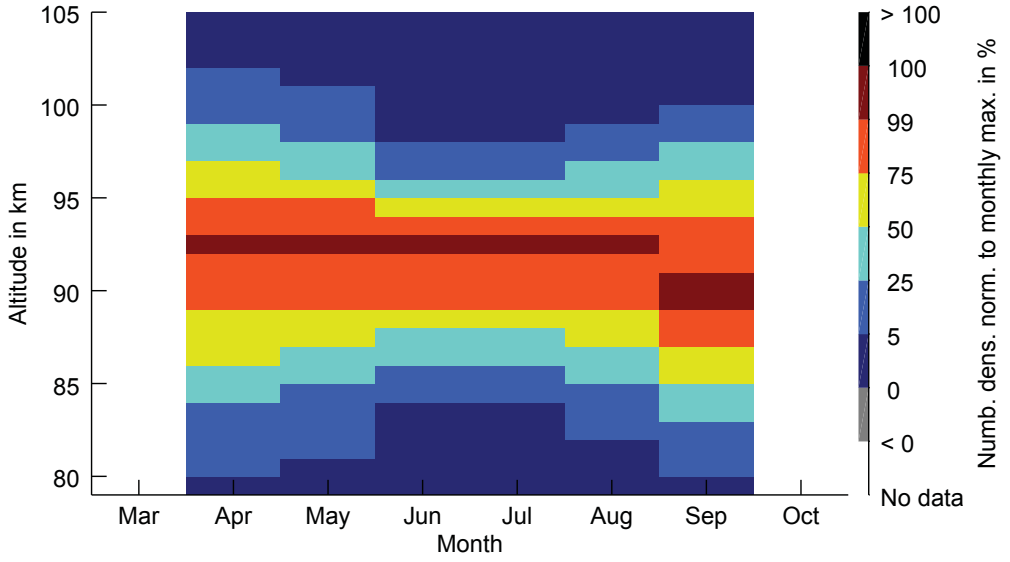

Fig. 9.5: Sodium number densities at 70° divided by the peak densities of the vertical profiles for each month.

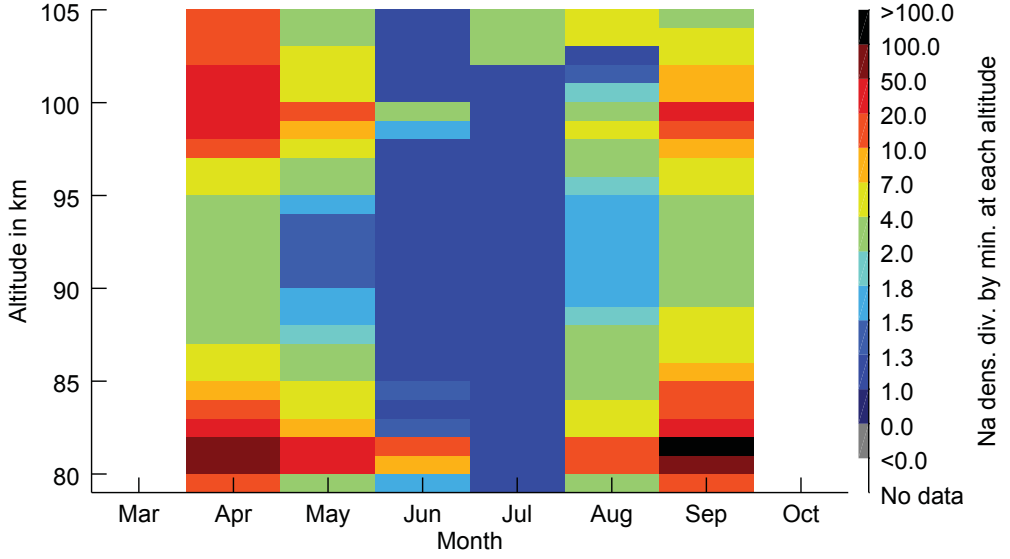

Fig. 9.6: Sodium number densities at 70° divided by minimum density for each altitude.

variation of Na also can be observed at lower latitudes, where no NLCs are present. To investigate, whether a reduction of the density can be seen at lower peak edge, Fig. 9.5 shows the monthly vertical profiles normalized to its monthly peak values,

and the width of the profile can be read from this figure. Figure 9.6 shows the monthly vertical profiles normalized to the minimum density at each altitude (in July; Note that if the minimum was 0 the minimum density was set to 1 cm^{-3}), showing at which altitude the seasonal variations are most pronounced. The FWHM of the profile shows a clear seasonal variation with around 11 km in spring and autumn and only 7 km in summer. However, the variations at the lower and the upper peak edge are similarly strong and it is rather the whole profile that is thinned out in summer than just one peak edge. Marsh et al. (2013a), who model very similar Na densities compared to the SCIAMACHY results shown here, also observe the thinning out, and it is explained by ionisation at the upper peak edge, and increased reaction of Na into the reservoir species $NaHCO_3$ due to cold temperatures at the lower edge. Uptake of Na by NLCs are not included in the model by Marsh et al. (2013a). This means, that there are two thinning out effects at the lower edge of the profile, which both are strongly correlated with the decrease of the mesopause temperature, and the weighting of both effects needs to be evaluated.

10 Estimation of the daily input of interstellar dust particles

In this chapter, a rough estimation of the daily input of interstellar dust particles based on the SCIAMACHY results is made. A similar study was made by Scharringhausen (2007) in his chapter 9. As introduced in Sect. 3, estimations for this quantity cover a wide range between 2 to 300 tons per day. The number density in the mesospheric metal atom layer is determined by the dynamical equilibrium of meteoric input on the one hand side and chemical loss processes on the other hand side. For this very simple estimation, we only take into account the most important loss reactions, which in the case of Mg are:

$$\begin{aligned} R1 \quad & \mathrm{Mg} + \mathrm{O_3} \quad \rightarrow \mathrm{MgO} + \mathrm{O_2} \\ R2 \quad & \mathrm{Mg} + \mathrm{O_2^+} \quad \rightarrow \mathrm{Mg^+} + \mathrm{O_2} \\ R3 \quad & \mathrm{Mg} + \mathrm{NO^+} \quad \rightarrow \mathrm{Mg^+} + \mathrm{NO} \end{aligned}$$

with their corresponding reaction rates k_1, k_2, and k_3, which can be taken from tables 7.8 and 7.9 (R1, R28 and R29). k_1 is temperature dependent, and the temperatures from Fig. 2.1 are used to account for this. The vertical profiles of NO^+,

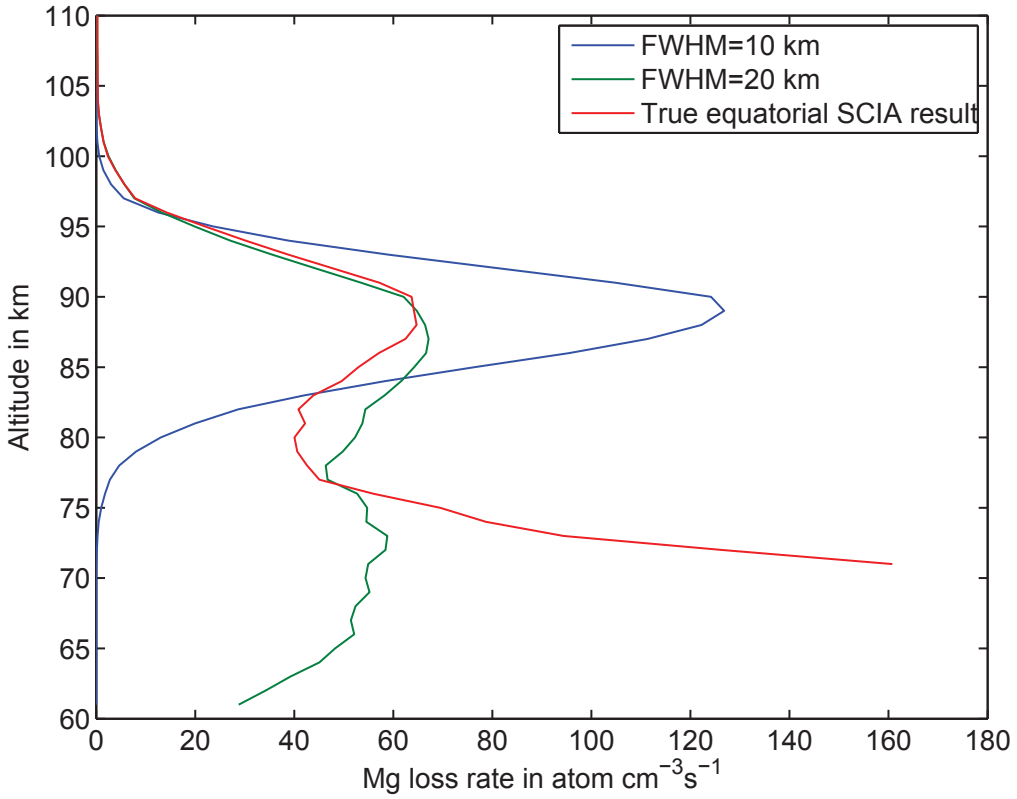

Fig. 10.1: First order loss rate of Mg atoms.

O_2^+ and are taken from Plane (2003), its Fig. 4, for this study, while the O_3 profile

is taken from Smith et al. (2013), its Fig 2. At each altitude level, the total loss rate of Mg is determined by:

$$\frac{dn(\mathrm{Mg})}{dt} = -k_1 n(\mathrm{Mg}) n(\mathrm{O_3}) - k_2 n(\mathrm{Mg}) n(\mathrm{O_2^+}) - k_3 n(\mathrm{Mg}) n(\mathrm{NO^+}) \tag{10.1}$$

with respective concentrations n and time t. Figure 10.1 shows the particle loss per cubic centimeter and second at each altitude for a Gaussian Mg profile with a FWHM of 20 km, respectively 10 km, a peak altitude of 90 km and a VCD of $2.52 \times 10^9\,\mathrm{cm^{-2}}$, as well as for the actual Mg profile retrieved from SCIAMACHY, shown in Fig. 7.22. For the here used data, the Mg loss nearly completely comes from reaction R1. The chosen Gaussian profiles are very similar to the SCIAMACHY and WACCM-Mg profiles in Fig. 7.22. For the thinner profile, the loss mainly occurs at around 90 km, while for the wider profile, the loss occurs at the whole lower profile edge. For the actually retrieved profile, the loss is especially large at the lower edge of the profile, which has known issues with the Ring effect correction. The peak Mg density is nearly $1200\,\mathrm{cm^{-3}}$ for the 20 km wide profile (also for the retrieved SCIAMACHY profile) and nearly $2400\,\mathrm{cm^{-3}}$ for the 10 km wide profile, which means that even at the peak altitude more than 5% of the Mg atoms are converted into MgO within one second.

Integrating the loss rate along the vertical direction results in an "column loss rate" in $\frac{\text{particle}}{\mathrm{cm^2\,s}}$. For the Gaussian profile with the FWHM of 20 km, the column loss rate is 1830 $\frac{\text{Mg atoms}}{\mathrm{cm^2\,s}}$, while for a Gaussian profile with the same VCD and altitude but a FWHM of 10 km the loss rate is smaller (1170 $\frac{\text{Mg atoms}}{\mathrm{cm^2\,s}}$), because the loss process is most effective where O_3 densities are high and O_3 density increases with decreasing altitude. The column loss rate for the measured SCIAMACHY Mg profile is 1570 $\frac{\text{Mg atoms}}{\mathrm{cm^2\,s}}$. However, this value is only for the column above 70 km and would increase to 11000 $\frac{\text{Mg atoms}}{\mathrm{cm^2\,s}}$, assuming a constant Mg profile with the value at 70 km between 60-70 km. Such a profile, however, could only be explained by a much lower ablation altitude of the metals or a very fast downward acceleration and transport speed of Mg, which seems to be rather unlikely. Therefore, only the results for the Gaussian profile are further discussed, noting, that a better Ring effect correction at the lower edge of the layer would significantly reduce the systematic errors in this estimation.

The results are on the same order of magnitude and close to the actual average input into the WACCM-Mg model of 1850 $\frac{\text{Mg atoms}}{\mathrm{cm^2\,s}}$. Multiplying the column loss rate with the Earth's surface area results in the injection rate of Mg atoms per second and multiplying the mass of Mg to this quantity results in the global mass injection rate for Mg to the upper atmosphere. As already discussed in Sect. 7.5.1 for 1850 $\frac{\text{Mg atoms}}{\mathrm{cm^2\,s}}$ this results in nearly $32.5\,\mathrm{kg\,d^{-1}}$.

Similar considerations as for Mg can be made for the estimation of the daily input of Na into the MLT. Reaction rates for the gas-phase neutral and ion-molecule reactions of Na species can be found, e.g., in Plane (2004). The reaction rates for the reaction with O_2^+ and NO^+ are nearly the same as in the case of Mg, but the reaction rate with O_3 is a factor 6 to 6.5 larger between 70 and 100 km altitude for Na compared to Mg. For this estimation a Na profile with a VCD of $5 \times 10^9\,\mathrm{cm^{-2}}$ is used and the profile has a FWHM of 5 to 15 km and peaks at 92 km, which comes

close to the annual average profile for most latitudes shown in Fig. 8.2. The Na column loss rates are relatively independent from the width of the profile varying between 11800 and 12300 $\frac{\text{Na atoms}}{\text{cm}^2\,\text{s}}$, which is a roughly a factor 6.5 more than for Mg, and roughly 25% of the incoming Na is converted to NaO within a second. The daily input rate of Na is 217 $\text{kg}\,d^{-1}$. In Marsh et al. (2013a), where a more sophisticated chemistry and transport model is used, the input rate of Na is estimated to be very similar to the one of Mg with 2000 $\frac{\text{Na atoms}}{\text{cm}^2\,\text{s}}$, respectively $35\,\text{kg}\,\text{d}^{-1}$. However, knowing that the differences between different estimation methods of the meteoric input vary by 2 orders of magnitude, a factor 6 is a relatively small difference.

The total daily mass input rate of extraterrestrial material into the Earth's atmosphere can be estimated from the daily mass input rate of Na and Mg into the upper atmosphere. However, further assumptions have to be made, which increase the uncertainties in those estimations. One approach is, to scale the upper atmospheric injection rates with its mass percentage in meteorites found on the Earth surface. A large fraction of these meteorites are so called Chondrites, which consist to roughly 10% of Mg and 5% Na (see, e.g., McDonough and Sun, 1995). Assuming that the bulk mass of the meteoroids fully ablates this would result in 0.3 to $0.7\,\text{t}\,\text{d}^{-1}$, which is on the lower edge of the wide range of estimations between 2 to 300 d^{-1} (see, e.g., Plane, 2012). However, the meteoroids entering the Earth's atmosphere are exposed to extreme thermal and pressure conditions, which may drastically change their composition. For example, it is likely, that the more volatile components ablate more strongly, while a higher percentage of non volatile matter is found in the meteorites. Following the discussion in Scharringhausen (2007) roughly 87% of the mass of meteoroids representing 98% of the particles in the velocity distribution function, which are slow ($11.2\,\text{km}\,\text{s}^{-1}$), is not ablated. Multiplying our estimates above by a factor of 10 would increase the total input to a value of 3 to 7 $\text{t}\,\text{d}^{-1}$.

11 Summary and outlook

The global metal atom and ion layers in the upper Earth's atmosphere have been investigated in this study. The general vertical structure of the Earth atmosphere has been discussed with a focus on the mesosphere and the thermosphere, where the metal layers reside. The MLT can not easily be probed by in-situ measurements and investigations mainly rely on remote sensing methods. As strong emitters of resonance fluorescence, the metals are easily accessed by remote sensing and yield valuable information on wave propagation and other transport processes at their altitude. The source of the mesospheric metals are meteoroids ablating in the upper atmosphere. The bulk mass of these meteoroids are about 220 µm sized particles entering the Earth atmosphere with a speed of around 11 km/s. The total mass of daily input is very uncertain and ranges from around 2 to 300 tons per day.

During their collision with the air particles in the upper atmosphere, the meteoroids slow down and undergo frictional heating. The meteoroids fully or partially ablate at around 100 km altitude and thereby inject metals (mostly atoms) into the MLT. The metal atoms become partly ionized in the lower thermosphere by charge exchange with the major species NO^+, O_2^+. The initial metal atoms and ions undergo chemical reactions, mainly started by reactions with O and O_3, and are transformed into stable reservoir species (hydroxides, carbonates) and other larger molecules, forming mesospheric smoke particles (MSP). The MSP play an important role as nucleation nuclei for the heterogenous nucleation of (other) aerosol particles as well as clouds in the middle atmosphere (NLCs and PSCs).

In this thesis, data from SCIAMACHY/Envisat was used to retrieve number densities of the mesospheric metal atoms and ions. SCIAMACHY operated in limb, nadir and occultation mode from 2002 to 2012. Since the highest tangent altitude of the nominal limb mode is around 92 km, which is just in the maximum density region of the metal atom layers, but below the peak altitude of the metal ion layers, a special limb MLT mode was introduced in 2008 covering the altitude region from 50 to 150 km. A survey of suitable emission lines for the number density retrieval was carried out, and so far magnesium atoms (Mg) and ions (Mg^+) as well as sodium atoms (Na) have been retrieved and investigated.

Several new features have been added to the existing metal number density retrieval algorithm. A Ring effect correction has been implemented, improving significantly the accuracy at the lower peak edge of the Mg layer and explaining, why unexpectedly high Mg and Mg^+ densities at altitudes below 70 km were found in the spectra before the correction. The emissivity calculation of the emission lines now includes several high resolution spectral features, that cannot be fully resolved by SCIAMACHY. High resolution solar spectra are used and several shifts and splitting mechanisms of the mesospheric absorption cross sections are considered now. These shifts and line splittings are important for another new included feature, the treatment of self absorption. Self absorption is the re-absorption of the radiance, which is emitted by the metals, by the metals itself (same species, but different individual particles), which leads to a saturation of the detected emission signal and an underestimation of the metal number densities, if not corrected for. To also retrieve Na at 590 nm, which lies in a spectral region, where multiple scattering has to be

considered, a correction for the additional backscattered radiation from below has been added to the retrieval algorithm.

New data sets of number density fields on a 2D latitudinal and vertical grid for Mg, Mg^+ and Na have been produced. Of these data sets the Mg data set shows the strongest unsystematic month to month and latitude to latitude variations, which can be attributed to the fact, that it has the lowest signal to noise ratio of the retrieved species. Nevertheless, some significant key features are found. The Mg layer has a nearly constant altitude at around 90 km and shows only low seasonal variations, which can hardly be separated from noise. However, a weak latitudinal dependence is present, showing slightly lower densities at higher latitudes. At high latitudes, Mg, furthermore, shows a weak but significant summer minimum in its peak density and VCD.

The Mg^+ data set shows a much clearer seasonal cycle than the Mg data set with a summer maximum in number density, which is most pronounced at mid latitudes. Furthermore, the peak density altitude varies with season and latitude. In summer, the Mg^+ layer is 5 to 10 km higher than in winter for the high and mid latitudes. In addition, there is a similar strong dependence on latitude, showing a higher layer at mid latitudes than at equatorial and polar latitudes. The highest density above 110 km is found at the equator. This fits well to the model of an equatorial outflow at the equator and transport to higher latitudes due to meridional winds, which can also explain the steep peak altitude gradient between low and mid latitudes.

The Na data set shows a clear seasonal cycle in number densities with a summer minimum most pronounced at high latitudes. At low latitudes, a semi-annual oscillation is observed. The altitude of the maximum number density is rather constant for all latitudes and seasons and is about 92 km. The width (FWHM) of the Na layer shows a seasonal variation, which is most pronounced at high latitudes with a lower vertical profile width in summer. Comparing the metal atom layers of Mg and Na, both show a constant altitude for the maximum density and lower densities during summer. However, the seasonal variations in the Na data set are much more pronounced than in the Mg data set.

All retrieved metal species have been compared to other observations and are in good agreement with these. The retrieved data sets were compared to WACCM model results. The SCIAMACHY Na data set agrees very well in terms of seasonality and absolute number densities with the WACCM model results and other measured data sets. For the Mg data set, the model predicts a much clearer seasonal variation (similar to Na) and, furthermore, predicts a peak number density, which is a factor 2 larger than for the observations. However, the observations show also a much wider profile than the model, so that the vertical column densities are of the same magnitude. For Mg^+ model and observations agree very well in terms of number densities, profile width and seasonal variations, but there are small differences: the model predicts the highest seasonal variations and densities at high latitudes, while these are observed at mid latitudes. Furthermore, the model shows a nearly constant peak altitude at around 95 km. However, the model does not include Lorentz-force driven ion transport, which is a possible explanation for the differences.

Since noctilucent clouds (NLCs) are formed at around 82-85 km altitude, which

is at the lower edge of the neutral metals layer, and mesospheric metals, in the form of meteoric smoke particles, are involved in their formation process, it was investigated whether a direct connection between both can be found in the metal data sets. The neutral metals indeed show a significant reduction during the NLC season and also a thinning out of the layer at the right altitudes. However, the WACCM model results for Na show a similar reduction and thinning out of the layer, without including NLC microphysics. There are two mechanism explaining the reduction of density at the lower peak edge (i.e., NLC formation, and chemical reactions into reservoir species), which both depend on low temperatures and, therefore, are both strongly correlated, which makes a differentiation complicated.

The metal number density retrieval algorithm has been refined in this work and it seems that all significant issues have been found and treated properly. As a starting step for an ongoing project the retrieval can be extended to either further metal species and/or extended to a larger data set. The most promising species to add to the SCIAMACHY data set are potassium (K) from the line at 770 nm, and with an improved Ring-effect correction, the several Fe and Ca^+ lines in the SCIAMACHY spectra might yield reasonable density profiles. For this work, the limb MLT data set from 2008 to 2012 was used. However, the data set can be extended to the nominal limb data set, which spans the time period from 2002 to 2012 and comprises daily measurements. This, e.g., can be done using the MLT data set as an a-priori, for the profile width and altitude. As the focus of this work was more on the retrieval algorithm itself, many of the further analyses of the data from Scharringhausen (2007) (e.g., correlation with meteor showers, solar flux, solar proton events etc.) can be repeated for the whole data set. Furthermore, the retrieval algorithm is also easily applicable to other similar satellite and rocket data, which will hopefully be available in the near future.

Part IV
Appendix

A Resonance fluorescence

This section gives a short theoretical overview on resonance fluorescence, which is the process causing the metal emissions. The classical derivation, based on the Lorentz-oscillator named after Hendrik Antoon Lorentz, can be found in similar forms in various textbooks, e.g., Demtröder (2004) and Chamberlain (1961). The quantum mechanical treatment is, e.g., shown in Demtröder (2000) and Chamberlain (1961). Additional source for the discussion of polarization and the phase function are, e.g., Chandrasekhar (1960) and Rödel and Wagner (2011). Unfortunately, these sources are either not showing the full derivation and/or are hard to understand, and some of the most comprehensive sources from the world wide web are unfortunately not properly citable, because they are either lecture scripts or presentations. The intention of this section is to give a comprehensive and complete overview of the classical treatment of this effect and also to provide information on the quantum mechanical treatment. This will, hopefully, make it easier to understand the original sources for the quantum mechanical treatment, e.g., by Weisskopf (1931) and Hamilton (1947).

A.1 classical theory

A.1.1 Driven and damped harmonic oscillator

The classical theory starts with the idea, that an electron with charge e and mass m_e, which is bound in a distance $\vec{r}$ to the atomic core by a restoring force $\vec{F}$, which follows Hooke's law ($\vec{F} \sim \vec{r}$), is excited by an external periodical electromagnetic field $\vec{E(t)} = \vec{E_0}e^{i\omega t}$. Due to the external force, the distance of the electron $\vec{r}$ to the core of the atom starts to vary, which results in an oscillating dipole. This dipole is called the Hertzian dipole, which is named after Heinrich Hertz. The external drive is damped either by the emission of photons due to electron transitions or due to collisions of electrons. The equation for the driven damped harmonic oscillator is:

$$\underbrace{m_e\ddot{\vec{r}}(t)}_{\text{force due to inertia}} + \underbrace{m_e\Gamma\dot{\vec{r}}(t)}_{\text{damping by scattering or radiation}} + \underbrace{m_e\omega_0^2\vec{r}(t)}_{\text{Hooke's law force}} = \underbrace{e\vec{E}_0e^{i\omega t}}_{\text{external force}} \tag{A.1}$$

After a short time of transition, during which the dipole changes it's oscillation frequency, the dipole is assumed to oscillate with the same frequency as the the driving external electromagnetic field. Then, the typical ansatz $\vec{r(t)} = \vec{r_0}e^{i\omega t}$ is made.

$$\begin{aligned} \vec{r}(t) &= \vec{r_0}e^{i\omega t} \\ \dot{\vec{r}}(t) &= i\omega\vec{r}(t) \\ \ddot{\vec{r}}(t) &= -\omega^2\vec{r}(t) \end{aligned} \tag{A.2}$$

Substituting this into A.1:

$$(-\omega^2 + i\gamma\omega + \omega_0^2)\vec{r} = \frac{\mathrm{e}\vec{E_0}}{m_e}e^{i\omega t}. \tag{A.3}$$

Rearranging the equation leads to:

$$\vec{r} = \frac{\mathrm{e}\vec{E_0}}{m_e} e^{i\omega t} \frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega}. \tag{A.4}$$

A.1.2 Total radiated Power P

From classical electrodynamics it is known, that accelerated charges are losing energy by emission of radiation. The total power P radiated by an accelerated particle with charge e is given by the Larmor formula (Larmor, 1897):

$$P = \frac{2}{3} \frac{1}{4\pi\epsilon_0} \frac{\mathrm{e}^2}{c^3} \ddot{\vec{r}}^2 \tag{A.5}$$

For deriving $\ddot{\vec{r}}$ we note, that we actually need the real part of $\vec{r}$. Therefore, we first expand Eq. (A.4), so that the complex number is in the numerator of the fraction:

$$\vec{r} = \frac{\mathrm{e}\vec{E_0}}{m_e} e^{i\omega t} \frac{\omega_0^2 - \omega^2 - i\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}. \tag{A.6}$$

The real part is:

$$\mathrm{Re}(\vec{r}) = \frac{\mathrm{e}\vec{E_0}}{m_e} \left(\frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \cos(\omega t) + \frac{\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \sin(\omega t) \right). \tag{A.7}$$

The 2nd derivative after time of this is:

$$\ddot{\vec{r}} = -\frac{\mathrm{e}\vec{E_0}}{m_e} \left(\frac{(\omega_0^2 - \omega^2)\omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \cos(\omega t) + \frac{\gamma\omega^3}{(\omega_0^2 - \omega^2) + \gamma^2\omega^2} \sin(\omega t) \right). \tag{A.8}$$

The square of this is:

$$\begin{aligned} \ddot{\vec{r}}^2 = \left(\frac{\mathrm{e}\vec{E_0}}{m_e} \right)^2 \Bigg(& \frac{(\omega_0^2 - \omega^2)^2\omega^4}{((\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2)^2} \cos^2(\omega t) \\ & + \frac{\gamma^2\omega^6}{((\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2)^2} \sin^2(\omega t) \\ & + \frac{2\gamma\omega^5(\omega_0^2 - \omega^2)^2}{((\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2)^2} \sin(\omega t)\cos(\omega t) \Bigg). \end{aligned} \tag{A.9}$$

We are interested in the radiated power averaged over one period. Using

$$\int_0^{2\pi} \sin^2(\omega t) dt = \frac{1}{2}, \int_0^{2\pi} \cos^2(\omega t) dt = \frac{1}{2} \text{ and } \int_0^{2\pi} \sin(\omega t)\cos(\omega t) dt = 0,$$

We derive:

$$< \ddot{\vec{r}}^2 >= \left(\frac{\mathrm{e}\vec{E_0}}{m_e} \right)^2 \left(\frac{(\omega_0^2 - \omega^2)^2\omega^4}{((\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2)^2} \frac{1}{2} + \frac{\gamma^2\omega^6}{((\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2)^2} \frac{1}{2} \right) \tag{A.10}$$

This can be further simplified:

$$< \ddot{\vec{r}}^2 >= \left(\frac{\mathrm{e}\vec{E}_0}{m_e}\right)^2 \frac{\omega^4}{2} \frac{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}{((\omega_0^2-\omega^2)^2+\gamma^2\omega^2)^2} \tag{A.11}$$

Canceling:

$$< \ddot{\vec{r}}^2 >= \frac{1}{2}\left(\frac{\mathrm{e}\vec{E}_0}{m_e}\right)^2 \frac{\omega^4}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}. \tag{A.12}$$

The radiated power averaged over one period then is:

$$< P >= \frac{1}{3}\frac{\mathrm{e}^4 E_0^2}{4\pi\epsilon_0 m_e^2 c^3} \frac{\omega^4}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}. \tag{A.13}$$

A.1.3 Total cross section σ

The ratio of the radiated power averaged over one period divided by the incoming intensity averaged over one period has the dimension of an area and is called the total cross section σ. For the incoming electromagnetic wave, we assume a plane parallel wave:

$$\vec{E} = \vec{E}_0 e^{i\omega t} \tag{A.14}$$

$$\vec{B} = \vec{B}_0 e^{i\omega t} \tag{A.15}$$

Using

$$\vec{B} = \frac{1}{c}\vec{k} \times \vec{E} \tag{A.16}$$

with the direction of propagation vector $\vec{k}$ of unit length, $\vec{B}_0$ is perpendicular to $\vec{E}_0$ and $B_0 = \frac{E_0}{c}$. The intensity averaged over one period of the incoming radiation $| < u_{\mathrm{in}} > |$ is the absolute value of the Poynting vector averaged over one period:

$$\begin{aligned} | < u_{\mathrm{in}} > | =& | < \frac{\vec{E} \times \vec{B}}{\mu_0} > | \\ =& \frac{1}{2} c\epsilon_0 E_0^2. \end{aligned} \tag{A.17}$$

Therefore, the total cross section is:

$$\begin{aligned} \sigma =& \frac{< P >}{| < u_{\mathrm{in}} > |} \\ =& \frac{8\pi}{3} r_e^2 \frac{\omega^4}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}, \end{aligned} \tag{A.18}$$

with r_e being the classical electron radius $r_e = \frac{e^2}{4\pi\epsilon_0 m_e c^2}$.

There are three special cases to discuss: $\omega << \omega_0$, $\omega >> \omega_0$, and $\omega \approx \omega_0$. The case $\omega >> \omega_0$ is called Thomson scattering and leads to a wavelength independent total cross section of

$$\sigma_{\mathrm{Thomson}} = \frac{8\pi}{3} r_e^2. \tag{A.19}$$

The case $\omega << \omega_0$ is called Rayleigh scattering and leads to

$$\sigma_{\text{Rayleigh}} = \frac{8\pi}{3} r_e^2 \frac{\omega^4}{\omega_0^4}. \tag{A.20}$$

Rayleigh scattering shows the characteristical $\sim \omega^4$ angular frequency dependency, which is well known for causing the blue color in a cloud free sky. The cross section is largest in the resonant case $\omega \approx \omega_0$, which is further discussed below.

A.1.4 Classical radiative decay rate of a single electron oscillator γ_{cl}

The damping force of the damped and driven harmonic oscillator is $\vec{F} = -\gamma m_e \dot{\vec{r}}$. To derive a power, the force is multiplied with the velocity:

$$P(t) = -\gamma m_e \dot{\vec{r}}^2 \tag{A.21}$$

The left hand side of the equation is replaced by the Larmor formula (Eq. (A.5)).

$$\frac{2}{3} \frac{1}{4\pi\epsilon_0} \frac{\mathrm{e}^2}{c^3} \ddot{\vec{r}}^2(t) = -\gamma m_e \dot{\vec{r}}^2 \tag{A.22}$$

Once again the ansatz $\vec{r} = \vec{r_0} e^{i\omega t}$ is made:

$$\frac{2}{3} \frac{1}{4\pi\epsilon_0} \frac{\mathrm{e}^2}{c^3} \omega^4 = \gamma m_e \omega^2 \tag{A.23}$$

Rearranging this equation and assuming that $\omega \approx \omega_0$ yields the classical radiative decay rate of a single electron oscillator γ_{cl}:

$$\gamma_{\text{cl}} = \frac{2}{3} \frac{e^2}{4\pi\epsilon_0 m_e c^3} \omega_0^2 = \frac{2}{3} \frac{r_e}{c} \omega_0^2. \tag{A.24}$$

This approximation $\omega \approx \omega_0$ is reasoned in the next section.

A.1.5 Lorentzian-distribution

Simplifying approximations can be made, when γ_{cl} is much smaller than ω_0. γ_{cl} is of the same dimension as ω_0 and can also be rewritten in the form:

$$\gamma_{\text{cl}} = a\omega_0 \tag{A.25}$$

with the dimensionless factor a being:

$$a = \frac{2}{3} \frac{e^2}{4\pi\epsilon_0 m_e c^3} \omega_0 \tag{A.26}$$

Setting $a = 1$ yields the frequency, where $\gamma_{\text{cl}} = \omega_0$.

$$\omega_0 = \frac{3}{2} \frac{4\pi\epsilon_0 m_e c^3}{e^2} \tag{A.27}$$

$$\omega_0 = 1.596 \cdot 10^{23}\,\mathrm{s}^{-1} \tag{A.28}$$

This corresponds to a wavelength of $\lambda_0 = 1.18 \cdot 10^{-14}\,\mathrm{m} = 1.18 \cdot 10^{-5}\mathrm{nm}$. The typical atomic emissions wavelength lie in the UV/Vis spectral region with wavelengths on the order of 100 to several hundreds of nm. Thus, it can be concluded, that $a << 1$ and $\gamma_{\mathrm{cl}} << \omega_0$ for our considerations. This also means, that the profile function is only non-negligibly small, for $\omega \approx \omega_0$. This means, that $\omega^4 \approx \omega_0^4$, $\omega^2 \approx \omega_0^2$, and $(\omega_0^2 - \omega^2)^2 = ((\omega_0 + \omega)(\omega_0 - \omega))^2 \approx (2\omega_0(\omega_0 - \omega))^2 = 4\omega_0^2(\omega_0 - \omega)^2$. The profile function then is:

$$\begin{aligned} \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} &\approx \frac{\omega_0^4}{4\omega_0^2(\omega_0-\omega)^2+\gamma^2\omega_0^2} \\ &= \frac{\omega_0^2}{4(\omega_0-\omega)^2+\gamma^2} \\ &= \frac{\omega_0^2/\gamma^2}{(\frac{2}{\gamma}(\omega_0-\omega))^2+1} \end{aligned} \tag{A.29}$$

Through a substitution, this profile can be brought into a form:

$$g(x) = Cf(x) \tag{A.30}$$

with C being a constant. and $f(x)$ being:

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}, \tag{A.31}$$

which is a normalized distribution function, as the indefinite integral of $f(x)$ is $F(x) = \frac{1}{\pi}\arctan(x) + const.$ $f(x)$ is called the Cauchy distribution, Lorentz distribution, Cauchy-Lorentz distribution or Breit-Wigner distribution, depending on the mathematical or physical problem discussed. To calculate the FWHM of the profile we search for the two angular frequencies, for which the distribution has half of the value as the maximum value of the distribution:

$$\frac{1}{2}\max\left(\frac{\omega_0^2/\gamma^2}{(\frac{2}{\gamma}(\omega_0-\omega))^2+1}\right) = \frac{\omega_0^2/\gamma^2}{(\frac{2}{\gamma}(\omega_0-\omega))^2+1} \tag{A.32}$$

This leads to two solutions:

$$\omega_{1,2} = \omega_0 \pm \frac{\gamma}{2}, \tag{A.33}$$

so that the simple solution for the FWHM is FWHM $= \gamma$.

A.1.6 Integrated total scattering cross section

The classical line width is extremely small and usually can not be resolved with spectrometers. Therefore, we are interested in the integrated linewidth. This quantity also has the advantage, that the integrated value can be easily redistributed to a broadened profile, which in our case is the Doppler-broadened profile in the mesosphere. We define the following 3 integrated total scattering cross sections: the angular frequency integrated cross section σ_ω, the frequency integrated cross section

σ_ν and the wavelength integrated cross section σ_λ:

$$\sigma_\omega = \int_{-\infty}^{\infty} \sigma(\omega) d\omega \tag{A.34}$$

$$\sigma_\nu = \int_{-\infty}^{\infty} \sigma(\nu) d\nu \tag{A.35}$$

$$\sigma_\lambda = \int_{-\infty}^{\infty} \sigma(\lambda) d\lambda. \tag{A.36}$$

Starting with

$$\sigma(\omega) = \frac{8\pi}{3} r_e^2 \frac{\omega_0^2/\gamma^2}{(\frac{2}{\gamma}(\omega_0 - \omega))^2 + 1}, \tag{A.37}$$

we substitute $x = \frac{-2}{\gamma}(\omega_0 - \omega)$. This leads to:

$$\begin{aligned} dx &= \frac{2}{\gamma} d\omega \\ d\omega &= 2\pi d\nu \\ d\omega &= -\frac{2\pi c}{\lambda^2} d\lambda \end{aligned}$$

Note that the sign for $d\lambda$ is canceled by the switch of the integration boundaries, so that we derive a positive value for all 3 integrated cross sections. The integration for the 3 integrated cross sections leads to:

$$\sigma_\omega = 2\pi^2 c r_e \tag{A.38}$$

$$\sigma_\nu = \pi c r_e \tag{A.39}$$

$$\sigma_\lambda = \pi r_e \lambda_0^2 \tag{A.40}$$

Note that Eq. (A.40), is already very close to the value used in Eq. (6.7) beside the factor f, which results from the quantum mechanical treatment of the problem.

A.1.7 Polarization and phase function

Until now, only the total scattering cross section was discussed, but nothing was said about the polarized components of the radiation. Classically, all three special cases of the cross section are based on the same model, the radiation of a Hertzian dipole, which is excited by an incoming radiation field. Lord Rayleigh (Strutt, 1871) first developed a theory, describing the scattering experiments by John Tyndall on small spherical particles. Beside the frequency dependency, he also found out, that while the light scattered into the same direction as the incident light is fully unpolarized, it is fully polarized, when scattered perpendicular to the direction of the incident light. Later it was shown, that this is also true for Thomson scattering. The classical explanation for this is, that a Hertzian dipole can not radiate into the direction of it's dipole axis. Figure A.1 illustrates the case for incident light polarized parallel and perpendicular to the scattering plane. For the incident light polarized perpendicular to the scattering plane, the dipole is also induced perpendicular to the scattering plane, so that the emission into the scattering plane is always perpendicular to the dipole axis regardless of the scattering angle θ. This means, that this component

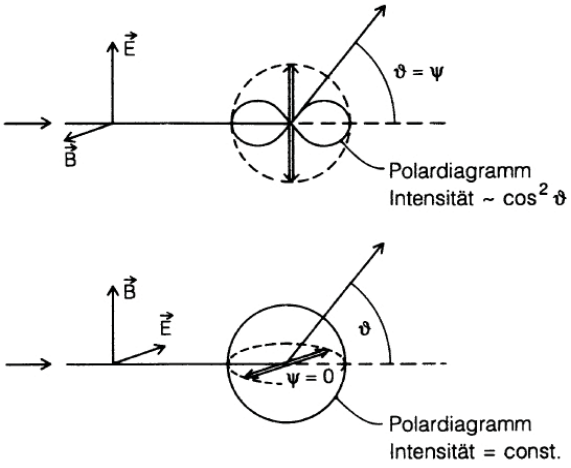

Abb. 1.17 Polardiagramm der Rayleigh-Streuung für die beiden Polarisationsrichtungen E-Feld in der Streuebene (oben) und E-Feld senkrecht zur Streuebene (unten). Als Streuebene ist jeweils die Zeichenebene anzusehen

Abb. A.1: Rayleigh scattering for different polarized light from Rödel and Wagner (2011). The upper sketch shows the case for incident light polarized in the the scattering plane (which lies in the drawing plane), while the lower sketch shows the case of incident light polarized perpendicular to the scattering plane.

is conserved ($I_{\perp,\,\mathrm{out}} = I_{\perp,\,\mathrm{in}}$). When the incoming light is fully polarized parallel to the scattering plane, the dipole axis also lies in the scattering plane. Then, there is no emission for the light component, that is parallel to the dipole axis. Therefore, only the projection of the dipole perpendicular to the direction of the outgoing light results in radiation, so that the outgoing electric field is modified with $\cos(\theta)$. As $I \sim E^2$, $I_{\parallel,\,\mathrm{out}} = I_{\parallel,\,\mathrm{in}} \cos^2(\theta)$. The Phase function of the total intensity $I = I_{\perp,\,\mathrm{out}} + I_{\parallel,\,\mathrm{out}}$ is $P(\theta) = \frac{3}{2}(1+\cos^2(\theta))$, which is normalized to a factor of 4π ($\int P d\theta = 4\pi$).

A.2 Quantum mechanical extension

While the classical theory was fully sufficient for Rayleigh-scattering, resonance fluorescence was one of the key problems to solve during the time of the development of quantum mechanics. As will be shown in the next sections the total scattering cross section has to be extended by a factor f, the so called f-value or oscillator strength. Furthermore, the line is broadened through the natural linewidth of both, the upper and the lower state involved in the transition. Additionally, the polarization of the emitted radiation is also different than in the classical case. This latter phenomenon is connected with the Hanle-effect, experimentally first observed by Wood and Ellett (1923), Gaviola and Pringsheim (1924), and Hanle (1924), which describes the depolarization of resonantly scattered light in presence of an external magnetic field. Later the Hanle-effect was generalized to the so called level crossing (see, e.g., Breit, 1933). Theoretically the depolarization was, e.g., discussed by Bohr (1924) and Heisenberg (1925), who, as noted by Weisskopf (1931), first formulated a rule on the directional distribution of the polarized light for resonance fluorescence, which was derived by perturbation theory. Weisskopf (1931) used Dirac's theory to

formulate a theory of resonance fluorescence. His equations led to some phenomenologically easy to comprehend consequences, e.g. that the theoretical linewidth of the "scattered" radiation is as wide as the incoming radiation, and thus is much smaller than the natural linewidth of the spontaneous emission, when no Doppler-broadening is considered. This also means, that resonance fluorescence is not an independent sequence of absorption and re-emission.

A.2.1 Einstein coefficients

In his contribution to the solution of the problem of black body radiation and the derivation of Planck's law, Einstein (1916) postulated three possible interaction processes of a photon and a quantum mechanical system with an upper state 2 and a lower state 1 for electrons: absorption, spontaneous emission, and stimulated emission. In the absorption process an incoming photon of the transition wavelength vanishes and an electron is lifted from the lower state 1 to the upper state 2. During the spontaneous emission process, an electron in the upper state transits from the upper state 2 to the lower state 1, and a photon of the transition wavelength is emitted. In the stimulated emission process a photon with the transition wavelength passes the electron state, and initiates a transition from the upper state 2 to the lower state 1, which results in a second photon with the same wavelength as the initializing photon. Each process has its own rate coefficient: B_{12} for absorption, A_{21} for spontaneous emission, and B_{21} for stimulated emission. The transition rate is proportional to the number of electrons in the initial state N. Furthermore, the transition rate of absorption and stimulated emission is proportional to the energy density of the present radiation field $u_\nu(\nu, T)$ (with unit $\mathrm{Jm^{-3}Hz^{-1}}$). The equation of the loss rates are:

$$\frac{dN_1}{dt} = -B_{12}N_1u_\nu(\nu,T) \text{ for absorption} \tag{A.41}$$

$$\frac{dN_2}{dt} = -B_{21}N_2u_\nu(\nu,T) \text{ for stimulated emission} \tag{A.42}$$

$$\frac{dN_2}{dt} = -A_{21}N_2 \quad \text{for spontaneous emission} \tag{A.43}$$

$$\frac{dN_1}{dt} = -\frac{dN_2}{dt} \quad \text{considering all three processes} \tag{A.44}$$

In an equilibrium state the three processes cancel each other out, so that the time derivatives are 0.

$$B_{12}N_1u_\nu(\nu,T) = B_{21}N_2u_\nu(\nu,T) + A_{21}N_2 \tag{A.45}$$

Reformulating this equation for $u_\nu(\nu, T)$ yields

$$u_\nu(\nu,T) = \frac{A_{21}N_2}{B_{12}N_1 - B_{21}N_2} \tag{A.46}$$

In thermodynamical equilibrium the population of the states follows the Boltzmann relation:

$$\frac{N2}{N1} = \frac{g_2}{g_1}\frac{e^{-\frac{E_2}{k_BT}}}{e^{-\frac{E_1}{k_BT}}} = \frac{g_2}{g_1}e^{-\frac{h\nu}{k_BT}}, \tag{A.47}$$

with g_1 and g_2 being the degeneracy factors of the respective states. Substituting this into the equation above results in

$$u_\nu(\nu, T) = \frac{A_{21}}{\frac{g_1}{g_2} B_{12} e^{\frac{h\nu}{k_B T}} - B_{21}}. \tag{A.48}$$

Now we have one equation but 4 unknowns: $u_\nu(\nu, T)$, A_{21}, B_{21}, and B_{12}. Therefore, some additional assumptions must be made. We demand that the energy density should be infinite for infinite temperature: $\lim_{T \to \infty} u_\nu(\nu, T) = \infty$, which results in $B_{21} = \frac{g_1}{g_2} B_{12}$. For small T we can approximate $e^x = 1 + x$, which results in

$$u_\nu(\nu, T) = \frac{A_{21}}{B_{21}} \frac{k_B T}{h\nu} \tag{A.49}$$

For small T Planck's law transits into the Rayleigh-Jeans-law and from a comparison of coefficients the relation of the two remaining unknown Einstein coefficients is derived:

$$A_{21} = \frac{8\pi h \nu^3}{c^3} B_{21} \tag{A.50}$$

$$A_{21} = \frac{8\pi h}{\lambda^3} B_{21} \tag{A.51}$$

To avoid losing the focus on the actual problem we are interested in, the Rayleigh-Jeans law is not discussed here. Note, that the Einstein coefficients have different units, as one depends on the energy density and the other does not. For typical solar spectral intensity in the visible region it is estimated, that the emission of the metals comes purely from the spontaneous emission. This is, e.g., the reason why in a laser electrons need to be pumped from the lower to the upper state (note, that for continuous wave lasers, this pumping is not resonant). In consequence, only the spontaneous emission process is of importance for our discussion of resonance fluorescence.

As usually radiances rather then energy densities are measured, Planck's law can also be defined for the the spectral radiance $B_\nu(\nu, T)$. By replacing $u_\nu(\nu, T)$ by

$$B_\nu(\nu, T) = \frac{c}{4\pi} u_\nu(\nu, T) \tag{A.52}$$

As the Einstein coefficients B_{21} and B_{12} depend on $u_\nu(\nu, T)$, sometimes a different conversion factor between A_{21} and B_{21} can be found in various textbooks (e.g., in Chamberlain, 1961, his Eq. (1.41)). However, the unit of A_{21} is independent of this, so that the unit of B_{21} shows, which convention is used in the definition of the Einstein coefficient.

A.2.2 Transition dipole moment

For a quantum mechanical discussion, we first reformulate the equation for the total scattering cross section (Eq. (A.18)) in the way necessary to derive a quantum mechanical expectation value. This equation uses the quantum mechanical expectation value of the dipole moment, which is e multiplied with the quantum mechanical

expectation value of the distance between electron and core. In the Larmor formula Eq. (A.5), we use $\ddot{\vec{r}} = -\omega^2\vec{r}$ so that$(\ddot{\vec{r}})^2 = \omega^4\vec{r}^2$. Following the entire calculation of the total cross section from above we derive

$$\sigma(\omega) = \frac{16\pi}{3}\frac{e^2\omega^4}{(4\pi\epsilon_0)^2c^4E_0^2} < \vec{r}^2 >, \tag{A.53}$$

with $< \vec{r}^2 >$ being classically $\frac{1}{2}r_0^2$, so that the classical formula is

$$\sigma(\omega) = \frac{8\pi}{3}\frac{\omega^4}{(4\pi\epsilon_0)^2c^4E_0^2}p_0^2, \tag{A.54}$$

or

$$\sigma(\omega) = \frac{8\pi}{3}\frac{\mathrm{e}^2\omega^4}{(4\pi\epsilon_0)^2c^4E_0^2}r_0^2, \tag{A.55}$$

with $r_0^2 = \vec{r_0}^2$ and $\vec{r} = \vec{r_0}e^{i\omega t}$ ($\vec{r_0} = \frac{\mathrm{e}\vec{E_0}}{m_e}\frac{1}{\omega_0^2-\omega^2+i\gamma\omega}$, see Eq. (A.6)).

For a quantum mechanical transition from an upper state 2 and a lower state 1, instead of $\vec{r_0}$ the quantum mechanical expectation value $\int d^3r\psi_2^*\vec{r_0}\psi_1$ is used. To derive the transition dipole moment M_{21}, the expectation value of $\vec{r_0}$ is multiplied by e.

$$M_{21} = e\int d^3r\psi_2^*\vec{r_0}\psi_1 \tag{A.56}$$

p_0^2 in Eq. (A.54) is replaced by $|M_{21} + M_{12}|^2 = 4|M_{21}|^2$ with $M_{21} = M_{12}$:

$$\sigma(\omega) = \frac{32\pi}{3}\frac{\omega^4}{(4\pi\epsilon_0)^2c^4E_0^2}|M_{21}|^2. \tag{A.57}$$

$S_{21} = |M_{21}|^2$ is also called the line strength.

A.2.3 Quantum mechanical linewidth

The classical linewidth γ_{cl} is typically much smaller than the quantum mechanical damping constant. Knowing the Einstein coefficients for spontaneous emission A_{ij} for an upper state i and all possible lower states j, the lifetime of the upper state i τ_i is derived by $\frac{1}{\tau_i} = \Gamma_i = \sum_j A_{ij}$. Γ_i is called the quantum mechanical damping constant. The linewidth of both the upper and the lower state are summed up to obtain the linewidth of the transition. Note, that also the ground state has a finite linewidth, as it has a limited life time, due to the absorption process.

A.2.4 f-value, Einstein coefficient and line strength

As discussed in Sect. A.1.6 we are interested in the integrated cross sections, because other broadening effects, e.g., Doppler-broadening dominate the linewidth of the emission lines we are interested in. Furthermore, the linewidth, which is measured by the instrument, is even broader than the Doppler width. The transition factor

from the classical values to the the quantum mechanical values is called the f-value or oscillator strength:

$$\sigma_\omega = 2\pi^2 c r_e f_{12} \tag{A.58}$$
$$\sigma_\nu = \pi c r_e f_{12} \tag{A.59}$$
$$\sigma_\lambda = \pi r_e f_{12} \lambda_{21}^2 \tag{A.60}$$

Eq. (A.60) is used in Eq. (6.7).

There are clear connections between the f-value, the Einstein coefficient A_{21} and the line strength S_{21}. Multiplying Eq. (A.57) with the absolute value of the incoming Poynting-flux Eq. (A.17) results in the average radiated Power P

$$P = \frac{4}{3} \frac{\omega^4}{4\pi\epsilon_0 c^3} S_{21}. \tag{A.61}$$

of the Hertzian dipole. On the other hand side the same quantity can be derived with the rate coefficient of the spontaneous emission for N_2 particles,

$$P = N_2 \hbar \omega A_{21}. \tag{A.62}$$

Setting both equations equal for one particle results in:

$$A_{21} = \frac{2}{3} \frac{\omega_{21}^3}{\epsilon_0 c^3 h} S_{21} \tag{A.63}$$

When, furthermore, degenerated states are considered, multiple lines with equal energy have to be considered. Let the upper level be 2, composed of $g_2 = 2J_2+1$ states (a) and the lower level be 1, composed of $g_1 = 2J_1 + 1$ states (b). Then:

$$S_{21} = \sum_a \sum_b S(a, b) \tag{A.64}$$

N_2 is then $N_2 = g_2 N_a$, so that

$$A_{21} = \frac{2}{3} \frac{\omega_{21}^3}{\epsilon_0 c^3 h} \frac{S_{21}}{g_2} \tag{A.65}$$

Now the connection of the f-value to the other two quantities is discussed. On the one hand side we know that the integrated cross section is given by Eq. (A.59). The classical Beer-Lambert absorption law for scattering into the solid angle $d\Omega$ is:

$$dI(\nu) = -N_1 \sigma_\nu I(\nu) dx d\Omega \tag{A.66}$$

In quantum mechanics the absorption is described by Eq. (A.41) and we keep in mind that for the conversion from spectral energy density to spectral intensity we need a factor $\frac{4\pi}{c}$, and furthermore another factor $\frac{1}{4\pi}$ for the normalization of the angular distribution, so that we get the following equation:

$$N_1 \sigma_\nu I(\nu) = N_1 \frac{1}{4\pi} \frac{4\pi}{c} B_{12} I(\nu) h\nu. \tag{A.67}$$

We insert Eq. (A.41) and $B_{12} = \frac{g_2}{g_1} B_{21}$ and divide both sides by $N_1 I(\nu)$

$$\pi c r_e f_{12} = \frac{g_2}{g_1} \frac{1}{4\pi} \frac{4\pi}{c} B_{21} h\nu. \tag{A.68}$$

We divide by $\pi c r_e$ and replace the Einstein coefficient of stimulated emission by the Einstein coefficient of spontaneous emission (Eq. A.50).

$$f_{12} = \frac{g_2}{g_1} \frac{c^3}{8\pi h\nu^3 \pi c r_e} \frac{1}{c} A_{21} h\nu. \tag{A.69}$$

We replace r_e by Eq. (A.24) and cancel, and derive:

$$f_{12} = \frac{g_2}{g_1} \frac{A_{21}}{3\gamma_{\mathrm{cl}}} \tag{A.70}$$

f_{12} is called the absorption oscillator strength, which is just the conversion factor f between the quantum mechanical and the classical integrated cross sections. There is also a quantity f_{21}, which is called the emission oscillator strength, which is derived by:

$$f_{21} = -\frac{g_1}{g_2} f_{12} = -\frac{A_{21}}{3\gamma_{\mathrm{cl}}} \tag{A.71}$$

Using Eq. (A.65) we derive the conversion factor between the absorption oscillator strength f_{12} and the line strength S_{21}:

$$f_{12} = \frac{g_2}{g_1} \frac{1}{3} \underbrace{\frac{3}{2} \frac{4\pi\epsilon_0 m_e c^3}{e^2 \omega_{21}^2}}_{\frac{1}{\gamma_{\mathrm{cl}}}} \underbrace{\frac{2}{3} \frac{\omega_{21}^3}{\epsilon_0 c^3 h} \frac{S_{21}}{g_2}}_{A_{21}}. \tag{A.72}$$

$$f_{12} = \frac{4\pi\omega_{21} m_e}{he^2} \frac{S_{21}}{g_1} \tag{A.73}$$

More connections to other useful quantities in this context can be found in Hilborn (1982).

A.2.5 Polarization and phase function

In contrast to the classical theory, the light observed at a scattering angle of 90° is not necessarily fully polarized anymore in the case of resonance fluorescence. Hanle (1924) explained this by a rotation of the polarization plane between the absorption and re-emission process by a gyration process in an external magnetic field. Weisskopf (1931) showed, that this rotation is only strong in the resonant case, but not present in the case of Rayleigh scattering. He also claimed that the emitted light is in a fixed phase relation with the absorbed light, which comes from the conservation of the spectrum of the absorbed light. Due to the spin-orbit coupling, a certain magnetic field is usually present in an atom, so that no external magnetic field is necessary to depolarize the light. The answer to the question, how strong the depolarization is for a transition line, lies in the evaluation of the line strength, for different scattering angles and polarization states of incoming light. For dipole

radiation, Hamilton (1947) derived easy to calculate formulas for the treatment of the polarization, which only depend on the angular momentum j of the lower state and the change of the angular momentum Δj through the absorption process, which are used in this work (see Eq. 6.8, Table 6.4 and Eq. 6.51).

References

Aikin, A. C. and Goldberg, R. A. (1973). Metallic Ions in the Equatorial Ionosphere. J. Geophys. Res., 78:734 – 745.

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., and Sorensen, D. (1999). LAPACK Users Guide. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition.

Anderson, G. P. and Hall, L. A. (1989). Solar irradiance between 2000 and 3100 Angstroms with spectral bandpass of 1.0 Angstroms. J. Geophys. Res., 94:6435 – 6441.

Anderson, J. G. and Barth, C. A. (1971). Rocket Investigation of the Mg I and Mg II Dayglow. J. Geophys. Res., 76:3723–3732.

Backhouse, T. W. (1885). The luminous cirrus cloud of June and July. Meteorolog. Mag., 20:133.

Bates, D. R. and Nicolet, M. (1950). Atmospheric hydrogen. Publ. Astron. Soc. Pac., 62:106 – 110.

Batteiger, V., Knünz, S., Herrmann, M., Saathoff, G., Schüssler, H. A., Bernhardt, B., Wilken, T., Holzwarth, R., Hänsch, T. W., and Udem, T. (2009). Precision spectroscopy of the $3s$-$3p$ fine-structure doublet in Mg^+. Phys. Rev. A, 80:022503.

Bender, S., Sinnhuber, M., Burrows, J. P., Langowski, M., Funke, B., and López-Puertas, M. (2013). Retrieval of nitric oxide in the mesosphere and lower thermosphere from SCIAMACHY. Atmos. Meas. Tech., 6:2521 – 2531.

Blanco-Muriel, M., Alarcon-Padilla, D. C., Lopez-Moratalla, T., and Lara-Coira, M. (2001). Computing the solar vector. Solar Energy, 70:431 – 441.

Bohr, N. (1924). Zur Polarisation des Fluorescenzlichtes. Naturwissenschaften, 12:1115–1117.

Boksenberg, A. and Gérard, J. C. (1973). Ultraviolet Observations of Equatorial Dayglow above the F_2 Peak. J. Geophys. Res., 78:4641 – 4650.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H. (1999). SCIAMACHY: Mission Objectives and Measurement Modes. J. Atmos. Sci., 56(2):127 – 150.

Boyer, R., Henoux, J. C., and Sotirovski, P. (1971). Isotopes of magnesium in the solar atmosphere. Solar Physics, 19:330 – 337.

Brasseur, G. P. and Solomon, S. (2005). Aeronomy of the Middle Atmosphere: Chemistry and Physics of the Stratosphere and Mesosphere, 3. edition. , Springer, Dordrecht.

Breit, G. (1933). Quantum Theory of Dispersion (Continued). Part VI and VII. Reviews of Modern Physics, 5:91–140.

Burrows, J. P., Hölzle, E., Goede, A. P. H., Visser, H., and Fricke, W. (1995). SCIAMACHY–scanning imaging absorption spectrometer for atmospheric chartography. Acta Astronautica, 35(7):445 – 451.

Ceplecha, Z., Borovička, J., Elford, W. G., Revelle, D. O., Hawkes, R. L., Porubčan, V., and Šimek, M. (1998). Meteor phenomena and bodies. Space Science Reviews, 84:327 – 471.

Chamberlain, J. W. (1961). Physics of the Aurora and Nightglow. Academic Press Inc., London, 1st edition.

Chamberlain, J. W., Hunten, D. M., and Mack, J. E. (1958). Resonance scattering by atmospheric sodium-IV abundance of sodium in twilight. J. Atmos. Terr. Phys., 12(2-3):153 – 165.

Chance, K. and Kurucz, R. L. (2010). An improved high-resolution solar reference spectrum for earth's atmosphere measurements in the ultraviolet, visible and near infrared. J. Quant. Spectrosc. Radiat. Transfer, 111:1289 – 1295.

Chandrasekhar, S. (1960). Radiative Transfer. Dover Publ.

Chapman, S. (1930). A theory of upper atmospheric ozone. Memoirs of the Royal Meteorological Society, 3:103–125.

Chen, F. F. (1984). Introduction to Plasma Physics and Controlled Fusion, 2nd Edition. , Plenum Press, New York.

Correira, J. (2009). Temporal and Spatial Distribution of Metallic Species in the Upper Atmosphere. Ph.D. thesis.

Correira, J., Aikin, A. C., Grebowsky, J. M., and Burrows, J. P. (2010). Metal concentrations in the upper atmosphere during meteor showers. Atmos. Chem. Phys., 10(3):909 – 917.

Correira, J., Aikin, A. C., Grebowsky, J. M., Pesnell, W. D., and Burrows, J. P. (2008). Seasonal variations of magnesium atoms in the mesosphere-thermosphere. Geophys. Res. Lett., 35:330 – 337.

Crutzen, P. J. (1970). The influence of nitrogen oxide on the atmospheric ozone content. Quarterly Journal of the Royal Meteorological Society, 96:320 – 327.

Dawkins, E. C. M., Plane, J. M. C., Chipperfield, M. P., Feng, W., Gumbel, J., Hedin, J., Höffner, J., and Friedman, J. S. (2014). First global observations of the mesospheric potassium layer. Geophys. Res. Lett., 41:5653 – 5661.

de Wit, R. J., Hibbins, R. E., Espy, P., Orsolini, Y. J., Limpasuvan, V., and E., K. D. (2014). Observations of gravity wave forcing of the mesopause region during the January 2013 major Sudden Stratospheric Warming. Geophys. Res. Lett., 41:4745 – 4752.

Deland, M. T., Shettle, E. P., Thomas, G. E., and Olivero, J. J. (2007). Latitude-dependent long-term variations in polar mesospheric clouds from SBUV version 3 PMC data. J. Geophys. Res., 112:D10315.

Demtröder, W. (2000). Experimentalphysik 3 – Atome, Moleküle und Festkörper. Springer, Berlin, Heidelberg, New York, Barcelona, Hongkong, London, Mailand, Paris, Singapur, Tokio, 2nd edition.

Demtröder, W. (2004). Experimentalphysik 2 – Elektrizität und Optik. Springer, Berlin, Heidelberg, New York, Hongkong, London, Mailand, Paris, Tokio, 3rd edition.

Dunker, T., Hoppe, U.-P., Feng, W., Plane, J., and Marsh, D. (2015). Mesospheric temperatures and sodium properties measured with the ALOMAR Na lidar compared with WACCM. J. Atmos. Sol.-Terr. Phys., 127:111 – 119.

EID (2014). Earth Impact Database `http://www.unb.ca/passc/EarthImpactDatabase/` (Accessed: 04/March/2014).

Einstein, A. (1911). Über den Einfluß der Schwerkraft auf die Ausbreitung des Lichtes. Annalen der Physik, 340:898 – 908 (905).

Einstein, A. (1916). Zur Quantentheorie der Strahlung. Mitteilungen der Physikalischen Gesellschaft Zürich, 18.

Fan, Z. Y., Plane, J. M. C., Gumbel, J., Stegman, J., and Llewellyn, E. J. (2007). Satellite measurements of the global mesospheric sodium layer. Atmos. Chem. Phys., 7:4107–4115.

Feng, W., Marsh, D. R., Chipperfield, M. P., Janches, D., Hoffner, J., Yi, F., and Plane, J. M. C. (2013). A global atmospheric model of meteoric iron. J. Geophys. Res. Atmos., 118:9456 – 9474.

Fentzke, J. T. and Janches, D. (2008). A semi-empirical model of the contribution from sporadic meteoroid sources on the meteor input function observed at arecibo. J. Geophys. Res. Space Physics, 113:A03304.

Fentzke, J. T., Janches, D., Strelnikova, I., and Rapp, M. (2009). Meteoric smoke particle properties derived using dual-beam Arecibo UHF observations of D-Region spectra during different seasons. J. Atmos. Sol.-Terr. Phys., 71:1982–1991.

Fesen, C. G. and Hays, P. B. (1982a). Mg^+ Morphology From Visual Airglow Experiment Observations. J. Geophys. Res., 87:9217 – 9223.

Fesen, C. G. and Hays, P. B. (1982b). Two-dimensional inversion technique for satellite airglow data. Applied Optics, 21:3784 – 3791.

Fesen, C. G., Hays, P. B., and Anderson, D. N. (1983). Theoretical Modelling of Low-Latitude Mg^+. J. Geophys. Res., 88:3211 – 3223.

Forbes, J. M. (1981). The equatorial electrojet. Rev. Geophys., 19:469 – 504.

Fricke, K. H. and von Zahn, U. (1985). Mesopause temperatures derived from probing the hyperfine structure of the D2 resonance line of sodium by lidar. J. Atmos. Terr. Phys., 47(5):499 – 512.

Frisch, M., Trucks, G., Schlegel, H., Scuseria, G., Robb, M., Cheeseman, J., Scalmani, G., Barone, V., Mennucci, B., Petersson, G., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H., Izmaylov, A., Bloino, J., Zheng, G., Sonnenberg, J., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., J. A. Montgomery, J., Peralta, J., Ogliaro, F., Bearpark, M., Heyd, J., Brothers, E., Kudin, K., Staroverov, V., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J., Iyengar, S., Tomasi, J., Cossi, M., Rega, N., Millam, J., Klene, M., Knox, J., Cross, J., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R., Yazyev, O., Austin, A., Cammi, R., Pomelli, C., Ochterski, J., Martin, R., Morokuma, K., Zakrzewski, V., Voth, G., Salvador, P., Dannenberg, J., Dapprich, S., Daniels, A., Farkas, O., Foresman, J., Ortiz, J., Cioslowski, J., and Fox, D. (2009). Gaussian 09, Revision A.1. Gaussian Inc, www.gaussian.com.

Fussen, D., Vanhellemont, F., Tétard, C., Mateshvili, N., Dekemper, E., Loodts, N., Bingen, C., Kyrölä, E., Tamminen, J., Sofieva, V., Hauchecorne, A., Dalaudier, F., Bertaux, J.-L., Barrot, G., Blanot, L., Fanton d'Andon, O., Fehr, T., Saavedra, L., Yuan, T., and She, C.-Y. (2010). A global climatology of the mesospheric sodium layer from GOMOS data during the 2002–2008 period. Atmos. Chem. Phys., 10:9225–9236.

Garcia, R. R., Marsh, D. R., Kinnison, D. E., Boville, B. A., and Sassi, F. (2007). Simulation of secular trends in the middle atmosphere, 1950-2003. J. Geophys. Res., 112:D09301.

Gardner, C. S., Plane, J. M. C., Pan, W., Vondrak, T., Murray, B. J., and Chu, X. (2005). Seasonal variations of the Na and Fe layers at the South Pole and their implications for the chemistry and general circulation of the polar mesosphere. J. Geophys. Res., 110:D10302.

Gardner, J. A., Broadfoot, A. L., McNeil, W. J., Lai, S. T., and Murad, E. (1999). Analysis and modeling of the GLO-1 observations of meteoric metals in the thermosphere. J. Atmos. Sol.-Terr. Phys., 61:542–562.

Gardner, J. A., Viereck, R. A., Murad, E., Knecht, J., Pike, C. P., Broadfoot, A. L., and Anderson, E. R. (1995). Simultaneous observations of neutral and ionic magnesium in the thermosphere. Geophys. Res. Lett., 22:2119 – 2122.

Gaviola, E. and Pringsheim, P. (1924). Über den Einfluß der Konzentration auf die Polarisation der Floureszenz von Farbstofflösungen. Zeitschrift für Physik, 24:24–36.

Gérard, J. C. and Monfils, A. (1974). Satellite Observations of the Equatorial Mg II Dayglow Intensity Distribution. J. Geophys. Res., 79:2544 – 2550.

Gérard, J. C. and Monfils, A. (1978). The MgII Equatorial Airglow Altitude Distribution. J. Geophys. Res., 83:4389 – 4391.

Gorshelev, V., Serdyuchenko, A., Weber, M., Chehade, W., and Burrows, J. P. (2013). High spectral resolution ozone absorption cross-sections – Part 1: Measurements, data analysis and comparison with previous measurements around 293 K. Atmos. Meas. Tech. Discuss., 6:6567 – 6611.

Grainger, J. and Ring, J. (1962). Anomalous Fraunhofer Line Profiles. Nature, 193:762.

Grebowsky, J. M. and Aikin, A. C. (2002). Chapter 8 in: Meteors in the Earth's atmosphere. edited by Murad, E. and Williams, I. P., Cambridge University Press, Cambridge.

Grebowsky, J. M., Goldberg, R. A., and Pesnell, W. D. (1998). Do meteor showers significantly perturb the ionosphere? J. Atmos. Sol.-Terr. Phys., 60:607–615.

Gumbel, J., Fan, Z. Y., Waldemarsson, T., Stegman, J., Witt, G., Llewellyn, E., She, C.-Y., and Plane, J. M. C. (2007). Retrieval of global mesospheric sodium densities from the Odin satellite. Geophys. Res. Lett., 34:L04813.

Hall, L. A. and Anderson, G. P. (1991). High resolution solar spectrum between 2000 and 3100 Angstroms. J. Geophys. Res., 96:927 – 931.

Hamilton, D. R. (1947). The resonance radiation induced by elliptically polarized light. Astrophys. J., 106:457 – 465.

Hanle, W. (1924). Über magnetische Beeinflussung der Polarisation der Resonanzfluoreszenz. Zeitschrift für Physik, 30:93–105.

Hanson, W. B. and Sterling, D. L. (1972). Source and Identification of Heavy Ions in the Equatorial F Layer. J. Geophys. Res., 77:5530 – 5541.

Hedin, J. and Gumbel, J. (2011). The global mesospheric sodium layer observed by Odin/OSIRIS in 2004–2009. J. Atmos. Sol.-Terr. Phys., 73:2221–2227.

Heinrich, D., Nesse, H., Blum, U., Acott, P., Williams, B., and Hoppe, U.-P. (2008). Summer sudden Na number density enhancements measured with the ALOMAR Weber Na Lidar. Ann. Geophys., 26:1057 – 1069.

Heisenberg, W. (1925). Über eine Anwendung des Korrespondenzprinzips auf die Frage nach der Polarisation des Fluoreszenzlichtes. Zeitschrift für Physik, 31:617–626.

Herrmann, U., Eberhardt, P., Hidalgo, M. A., and Kopp, E. (1978). Metal ions and isotopes in sporadic E-layer during Perseid meteor shower. Space Res., 18:249 – 252.

Hervig, M. E., Deaver, L. E., Bardeen, C. G., Russel III, J. M., Bailey, S. M., and Gordley, L. L. (2012). The content and composition of meteoric smoke in the mesospheric ice particles from SOFIE observations. J. Atmos. Sol.-Terr. Phys., 84-85:1–6.

Hilborn, R. C. (1982). Einstein coefficients, cross sections, f values, dipole moments, and all that; revised version from 2002. American Journal of Physics, 50:982–986.

Huang, W. (2011). Simultaneous and Common-Volume Lidar Observations of Mesospheric Na and Fe Layers at Boulder Colorado (40N, 105W) in 2010. Joint CEDAR/GEM Workshop 2011, Santa Fe, NM, 26.6.-1.7., Poster MLTL-04.

Hunten, D. M. (1956). Resonance scattering by atmospheric sodium–III Supplementary considerations. J. Atmos. Terr. Phys., 9(4):179 – 183.

Istomin, V. (1963). Ions of extra-terrestrial origin in the earth's ionosphere. Planet. Space Sci., 11:173 – 181.

Janches, D., Heinselman, C., Chau, J., Chandran, A., and Woodman, R. (2006). Modeling the global micrometeor input function in the upper atmosphere observed by high power and large aperture radars. J. Geophys. Res., 111:A07317.

Jesse, O. (1885). Auffallende Abenderscheinungen am Himmel. Meteorologische Zeitschrift, 2:311 – 312.

Johnson, C. Y. and Meadows, E. B. (1955). First Investigation of Ambient Positive-Ion Compostion to 219 km by Rocket-Borne Spectrometer. J. Geophys. Res., 60:193 – 203.

Joiner, J. and Aikin, A. C. (1996). Temporal and spatial variations in upper atmospheric Mg^+. J. Geophys. Res, 101:5239 – 5250.

Kendall, P. C. and Windle, D. W. (1965). The Appleton Anomaly. Nature, 207:964 – 965.

Kopp, E. (1997). On the abundance of metal ions in the lower ionosphere. J. Geophys. Res, 102:9667 – 9674.

Kopp, E., Eberhardt, U., Herrmann, U., and Björn, L. (1985a). Positive ion composition of the high latitude summer D-region with noctilucent clouds. J. Geophys. Res., 90:13041–13053.

Kopp, E. and Herrmann, U. (1984). Ion composition in the lower ionosphere. Ann. Geophys., 2:83 – 94.

Kopp, E., Ramseyer, H., and Björn, L. (1985b). Positive ion composition and electron density in a combined auroral and NLC event. Adv. Space Res., 4:157 – 161.

Kramida, A., Ralchenko, Y., Reader, J., and NIST ASD Team (2012) (2012). NIST Atomic Spectra Database (version 5.0).

Krijger, J. M., Snell, R., van Harten, G., Rietjens, J. H. H., and Aben, I. (2014). Mirror contamination in space I: approach. Atmos. Meas. Tech. Discuss., 7:1213 – 1246.

Kyte, F. T. and Wasson, J. T. (1986). Accretion Rate of Extraterrestrial Matter: Iridium Deposited 33 to 67 Million Years Ago. Science, 232:1225 – 1229.

Langowski, M., Sinnhuber, M., Aikin, A., von Savigny, C., and Burrows, J. P. (2014a). Retrieval algorithm for densities of mesospheric and lower thermospheric metal atom and ion species from satellite-borne limb emission signals. Atmos. Meas. Tech., 7:29 – 48.

Langowski, M., von Savigny, C., Burrows, J. P., Feng, W., Plane, J. M. C., Marsh, D. R., Janches, D., Sinnhuber, M., and Aikin, A. (2014b). Global investigation of the Mg atom and ion layers using SCIAMACHY/Envisat observations between 70 km and 150 km altitude and WACCM-Mg model results. Atmos. Chem. Phys. Discuss., 14:1971 – 2019.

Langowski, M. P., von Savigny, C., Burrows, J. P., Dunker, T., Hoppe, U.-P., Sinnhuber, M., and Aikin, A. C. (2015a). Retrieval of sodium number density profiles in the mesosphere and lower thermosphere from SCIAMACHY limb emission measurements. Atmos. Meas. Tech. Discuss., 8:7909 – 7952.

Langowski, M. P., von Savigny, C., Burrows, J. P., Feng, W., Plane, J. M. C., Marsh, D. R., Janches, D., Sinnhuber, M., Aikin, A., and Liebing, P. (2015b). Global investigation of the Mg atom and ion layers using SCIAMACHY/Envisat observations between 70 km and 150 km altitude and WACCM-Mg model results. Atmos. Chem. Phys., 15:273 – 295.

Larmor, J. (1897). LXIII. On the theory of the magnetic influence on spectra; and on the radiation from moving ions. Philosophical Magazine Series 5, 44:503–512.

Le Boiteux, S., Klein, A., Rios Leite, J. R., and Ducloy, M. (1988). Doppler-free spectroscopy and isotopic shift of the Mg I resonance line at 285 nm . Le Journal de Physique, 49(6):885 – 887.

Leslie, R. C. (1885). Sky Glows. Nature, 32:245.

Liebing, P., Bramstedt, K., Noël, S., Rozanov, V., Bovensmann, Â. H., and Burrows, J. P. (2013). Polarization data from SCIAMACHY limb backscatter observations compared to vector radiative transfer model simulations. Atmos. Meas. Tech., 6:1503 – 1520.

Lindholm, E. (1942). Über die Verbreiterung und Verschiebung von Spektrallinien: experimentelle und theoretische Beiträge. Ph.D. thesis.

Lockwood, M., Horwitz, J. R., Chandler, M. O., Waite, J. H., Moore, T. E., and Chappel, C. R. (1985). The Cleft Ion Fountain. J. Geophys. Res., 90:9736 – 9748.

Love, S. G. and Brownlee, D. E. (1993). A direct measurement of the terrestrial mass accretion rate of cosmic dust. Science, 262:550 – 553.

Marmet, P. (1989). Red shift of spectral lines in the Sun's chromosphere. IEEE Trans. Plasma Sci., 17:238 – 244.

Marsh, D. R., Garcia, R. R., Kinnison, D. E., Boville, B. A., Sassi, F., Solomon, S. C., and Matthes, K. (2007). Modeling the whole atmosphere response to solar cycle changes in radiative and geomagnetic forcing. J. Geophys. Res., 112:D23306.

Marsh, D. R., Janches, D., Feng, W., and Plane, J. M. C. (2013a). A global model of meteoric sodium. J. Geophys. Res. Atmos., 118(19):11442 – 11452.

Marsh, D. R., Mills, M. J., Kinnison., D. E., Lamarque, J. F., Calvo, N., and Polvani, L. M. (2013b). Climate Change from 1850 to 2005 Simulated in CESM1(WACCM). Journal of Climate, 26:7372 – 7391.

Martínez-Núñez, E., Whalley, C. L., Shalashilin, D., and Plane, J. M. C. (2010). Capture, Collisional Stabilization and Collision-Induced Dissociation. Journal of Physical Chemistry A, 114:6472 – 6479.

Martus, C. M. and Collins, R. L. (2013). Resonance lidar detection of mesospheric nickel layer. AGU Fall Meeting 2013, San Francisco, Ca, 9.12.-13.12., Poster SA11B-1924.

McDonough, W. F. and Sun, S.-s. (1995). The composition of Earth. Chemical Geology, 120:223 – 253.

McNutt, D. P. and Mack, J. E. (1963). Telluric Absorption, Residual Intensities, and Shifts in the Fraunhofer D Lines. J. Geophys. Res., 68(11):3419 – 3429.

Mende, S. B., Swenson, G. R., and Miller, K. L. (1985). Observations of E and F Region Mg^+ From Spacelab 1. J. Geophys. Res., 90:6667–6673.

Minschwaner, K., Herceg, D., Budzien, S. A., Dymond, K. F., Fortna, C., and P., M. R. (2007). Observations of middle ultraviolet emissions in the middle and lower thermosphere: NO, O_2, O and Mg^+. J. Geophys. Res., 112:A10311.

Molina, M. J. and S., R. F. (1974). Stratospheric sink for chlorofluoromethanes: chlorine atom-catalysed destruction of ozone. Nature, 249:810 – 812.

Murad, E. and Williams, I. P. (2002). Meteors in the Earth's atmosphere. , Cambridge University Press, Cambridge.

Narcisi, R. S. (1971). Composition studies of the lower ionosphere. In Physics of the Upper Atmosphere, pages 11–59.

Narcisi, R. S. and Bailey, A. D. (1965). Mass Spectrometric Measurements of Positive Ions at Altitudes from 64 to 112 Kilometers. J. Geophys. Res., 70:3687 – 3700.

Nesvorný, D., Jenniskens, P., Levison, H. F., Bottke, W., Vokrouhlický, D., and Gounelle, M. (2010). Cometary origin of the Zodiacal Cloud and carbonaceous micrometeorites. Implications for hot debris disks. Astrophys. J., 713:816 – 836.

Penney, C. M., Peters, R. L. S., and Lapp, M. (1974). Absolute rotational Raman cross sections for N_2, O_2 and CO_2. J. Opt. Soc. Am., 64:712 – 716.

Philbrick, C. R., Narcisi, R. S., Good, R., Hoffmann, H. S., Keneshea, T. J., A., M. M., Zimmermann, S. P., and Reinisch, B. W. (1973). The Aladdin Experiment - Part II, composition. Space Res., VII:441–449.

Picone, J. M., Hedin, A. E., Drob, D. P., and Aikin, A. C. (2002). NRLMSISE-00 empirical model of the atmosphere: Statistical comparison and scientific issues. J. Geophys. Res., 107:1468.

Plane, J. M. C. (2003). Atmospheric Chemistry of Meteroric Metals. Chem. Rev., 103:4963 – 4984.

Plane, J. M. C. (2004). A time-resolved model of the mesospheric Na layer: constraints on the meteor input function. Atmos. Chem. Phys., 4:627 – 638.

Plane, J. M. C. (2011). On the role of metal silicate molecules as ice nuclei. J. Atmos. Sol.-Terr. Phys., 73:2192–2200.

Plane, J. M. C. (2012). Cosmic dust in the earth's atmosphere. J. Atmos. Sol.-Terr. Phys., 41:6507 – 6518.

Plane, J. M. C., Feng, W., and Dawkins, E. C. M. (2015). The Mesosphere and Metals: Chemistry and Changes. Chem. Rev., 115:4497 – 4541.

Plane, J. M. C. and Helmer, M. (1995). Laboratory Study of Reactions Mg + O_3 and MgO + O_3, Implications for the Chemistry of Magnesium in the Upper Atmosphere. Faraday Discuss., 100:411 – 430.

Plane, J. M. C. and Whalley, C. L. (2012). A New Model for Magnesium Chemistry in the Upper Atmosphere. J. Phys. Chem. A, 116(24):6240 – 6252.

Popp, C., Wang, P., Brunner, D., Stammes, P., Zhou, Y., and Grzegorski, M. (2011). MERIS albedo climatology for FRESCO+ O_2 A-Band cloud retrieval. Atmos. Meas. Tech., 4:463–483.

Rapp, M. and Lübken, F.-J. (2004). Polar mesosphere summer echoes (PMSE): review of observations and current understanding. Atmos. Chem. Phys., 4:2601–2633.

Roddy, P. A., Earle, G. D., Swenson, C., Carlson, C. G., and Bullett, T. W. (2004). Relative concentrations of molecular and metallic ions in the midlatitude intermediate and sporadic-E layers. Geophys. Res. Lett., 31:L19807.

Rödel, W. and Wagner, T. (2011). Physik unserer Umwelt: Die Atmosphäre, 4. Auflage. , Springer, Heidelberg, Dordrecht, London, New York.

Rollason, R. J. and Plane, J. M. C. (2001). A kinetic study of the reactions of MgO with H_2O, CO_2 and O_2: implications for magnesium chemistry in the mesosphere. Phys. Chem. Chem. Phys., 3:4733 – 4740.

Rozanov, A., Rozanov, V., Buchwitz, M., Kokhanovsky, A., and Burrows, J. P. (2005). SCIATRAN 2.0 - A new radiative transfer model for geophysical applications in the 175 - 2400 nm spectral region. Adv. Space Res., 36(5):1015 – 1019.

Rozanov, V. V., Rozanov, A. V., Kokhanovsky, A. A., and Burrows, J. P. (2014). Radiative transfer through terrestrial atmosphere and ocean: Software package SCIATRAN. J. Quant. Spectrosc. Radiat. Transfer, 133:13–71.

Rutherford, J. A., Mathis, R. F., Turner, B. R., and Vroom, D. A. (1971). Formation of Magnesium ions by charge transfer. J. Chem. Phys., 55:3785 – 3793.

Saunders, R. W., Dhomse, S., Tian, W. S., Chipperfield, M. P., and Plane, J. M. C. (2012). Interactions of meteoric smoke particles with sulphuric acid in the Earth's stratosphere. Atmos. Chem. Phys., 12:4387 – 4398.

Scharringhausen, M. (2007). Investigation of Mesospheric and Thermospheric Magnesium Species from Space. Ph.D. thesis.

Scharringhausen, M., Aikin, A. C., Burrows, J. P., and Sinnhuber, M. (2008a). Global column density retrieval of mesospheric and thermospheric MgI and MgII from SCIAMACHY limb and radiance data. J. Geophys. Res., 113(3):D13303.

Scharringhausen, M., Aikin, A. C., Burrows, J. P., and Sinnhuber, M. (2008b). Space-borne measurements of mesospheric magnesium species / a retrieval algorithm and preliminary profiles. Atmos. Chem. Phys., 8(7):1963 – 1983.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P. (2013). High spectral resolution ozone absorption cross-sections – Part 2: Temperature dependence. Atmos. Meas. Tech. Discuss., 6:6613 – 6643.

She, C. Y., Chen, S., Hu, Z., Sherman, J., Vance, J. D., Vasoli, V., White, M. A., Yu, J., and Krueger, D. A. (2000). Eight-year climatology of nocturnal temperature and sodium density in the mesopause region (80 to 105 km) over Fort Collins, CO(41°N,105°W). Geophys. Res. Lett., 27(20):3289 – 3292.

Shettle, E. P., Deland, M. T., Thomas, G. E., and Olivero, J. J. (2009). Long term variations in the frequency of polar mesospheric clouds in the Northern Hemisphere from SBUV. Geophys. Res. Lett., 36:L02803.

Skupin, J., Weber, M., Bovensmann, H., and Burrows, J. P. (2004). The Mg II solar activity proxy indicator derived from GOME and SCIAMACHY. Proceedings of the ENVISAT and ERS Symposium (SP-572).

Slijkhuis, S. (2008). SCIAMACHY Level 0 to 1c Processing, Algorithm Theoretical Basis Document, ENV-ATB-DLR-SCIA-0041.

Slipher, V. M. (1929). Emissions in the spectrum of the light of the night sky. Publ. Astron. Soc. Pac., 41:262 – 263.

Smith, A. K., Harvey, V. L., Mlynczak, M. G., Funke, B., Garcia-Comas, M., Hervig, M., Kaufmann, M., Kyrölä, E., López-Puertas, M., McDade, I., Randall, C. E., Russel III, J. M., Sheese, P. E., Shiotani, M., Skinner, W. R., Suzuki, M., and Walker, K. A. (2013). Satellite observations of ozone in the upper mesosphere. J. Geophys. Res., 118:5803 – 5821.

Snow, M., McClintock, W. E., Woods, T. N., White, O. R., Harder, J. W., and Rottman, G. (2005). The Mg II Index from SORCE. In Rottman, G., Woods, T., and George, V., editors, The Solar Radiation and Climate Experiment (SORCE), pages 325–344. Springer New York.

Solomon, S. (1999). Stratospheric ozone depletion: a review of concepts and history. Rev. Geophys., 37,3:275 – 316.

Stevens, M. H., Englert, C. R., DeLand, M. T., and Bailey, S. M. (2007). Polar mesospheric cloud mass and the ice budget: 2. Application to satellite data sets. J. Geophys. Res., 112:D08205.

Strutt, J. W. (1871). XV. On the light from the sky, its polarization and colour. Philosophical Magazine Series 4, 41:107–120.

Suarez, M. J., Rienecker, M. M., Todling, R., Bacmeister, J., Takacs, L., Liu, H. C., Gu, W., Sinkiewicz, M., Kostar, R. D., Gelero, R., Stajner, I., and Nielsen, J. E. (2008). The GEOS-5 Data Assimilation System-Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. NASA/TM-2008-104606-VOL-27.

Swider, W. (1984). Ionic and neutral concentrations of Mg and Fe near 92 km. Planet. Space Sci., 32:307312.

Tikhonov, A. N. (1943). On the stability of inverse problems. Procedings of the USSR Academy of Science, 39(5):195 – 198.

von Savigny, C., Bovensmann, H., Bramstedt, K., Dikty, S., Ebojie, F., Jones, A., Noél, S., Rozanov, A., and Sinnhuber, B.-M. (2009). Indications for long-term trends and seasonal variations in the SCIAMACHY Level 1 version 6.03 tangent heigth information, Techn. Note-IUP-scia-pointing-2009-01. University of Bremen, Bremen, Germany, Issue 2.

von Savigny, C., Kaier, J. W., Bovensmann, H., Burrows, J. P., McDermid, I. S., and Leblanc, T. (2005). Spatial and temporal characterization of SCIMACHY limb pointing errors during the first three years of the mission. Atmos. Chem. Phys., 5:2593 – 2602.

von Savigny, C., Kokhanovsky, A., Bovensmann, H., Eichmann, K.-U., Kaiser, J., Noël, S., Rozanov, A. V., Skupin, J., and Burrows, J. P. (2004). NLC detection and particle size determination: first results from SCIAMACHY on Envisat. Adv. Space Res., 34:851–856.

von Savigny, C. and Lednyts'kyy, O. (2013). On the relationship between atomic oxygen and vertical shifts between OH Meinel bands originating from different vibrational levels. Geophys. Res. Lett., 40:5821 – 5825.

Vondrak, T., Plane, J. M. C., Broadley, S., and Janches, D. (2008). A chemical model of meteoric ablation. Atmos. Chem. Phys., 8:7015 – 7031.

Weisskopf, V. (1931). Zur Theorie der Resonanzfluoreszenz. Annalen der Physik, 401:23 – 66.

Whaley, R. C., Petitet, A., and Dongarra, J. J. (2001). Automated empirical optimizations of software and the atlas project. Parallel Computing, 27:3 – 15.

Whalley, C. L., Gómez Martín, J. C., Wright, T. G., and Plane, J. M. C. (2011). A kinetic study of Mg^+ and Mg-containing ions reacting with O_3, O_2, N_2, CO_2, N_2O and H_2O, implications for magnesium ion chemistry in the upper atmosphere. Phys. Chem. Chem. Phys., 13:6352 – 6364.

Whalley, C. L. and Plane, J. M. C. (2010). Meteoric ion layers in the Martian atmosphere. Faraday Discuss., 147:349 – 368.

Wood, R. W. and Ellett, A. (1923). On the Influence of Magnetic Fields on the Polarisation of Resonance Radiation. Proc. Roy. Soc., 103:396–403.

Zbinden, P. A., Hidalgo, M. A., Eberhardt, P., and Geiss, J. (1975). Mass spectrometer measurements of the positive ion composition in the D- and E-Regions of the ionosphere. Planet. Space Sci., 23:1621 – 1642.

List of Figures

List of Tables

List of abbreviations

ATLAS Automatic Tuned Linear Algebra Software
CFC Chlorofluorocarbon
CME coronal mass ejection
Envisat Environmental Satellite
EOARD European Office of Aerospace Research and Development
ESA European Space Agency
EUV extended ultraviolet
FWHM full width at half maximum
GOME Global Ozone Monitoring Experiment
GOMOS Global Ozone Monitoring by Occultation of Stars
HF high-frequency radio waves
LAPack Linear Algebra Package
LDEF Long Duration Exposure Facility
LF low-frequency radio waves
LFS line from sun
lidar (visible) light detection and ranging
LOS line of sight
MF mid-frequency radio waves
MIF meteoric input function
MLT mesosphere and lower thermosphere
MSP meteoric smoke particle
NIST National Institute of Standards and Technology
NLC noctilucent cloud
NRLMSISE-00 Naval Research Laboratory, Mass Spectrometer and Incoherent Scatter, ground through Exosphere, released in 2000 (earth atmosphere model)
OSIRIS Optical Spectrograph and InfraRed Imager System
PMC polar mesospheric cloud
PMSE polar mesospheric summer echo
PSC polar stratospheric cloud
radar radio detection and ranging
SAA South Atlantic Anomaly
SAO Smithsonian Astrophysical Observatory

SBUV Solar Backscatter Ultraviolet Instrument
SCD slant column density
SCE slant column emission
SCIAMACHY SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
SCIATRAN Radiative transfer model and retrieval algorithm first developed for SCIAMACHY measurements
SZA solar zenith angle
TCFOV total clear field of view
UV ultraviolet
VCD vertical column density
Vis visible
vmr volume mixing ratio
WACCM Whole Atmosphere Community Climate Model
ZHR zenithal hourly rate

Acknowledgements

At this point, I would like to acknowledge the participation of the many people that contributed to this work. First of all, I would like to thank John Burrows for giving me the opportunity to work at the IUP in Bremen. I would like to thank Christian von Savigny for the very good supervision of my work as a PhD student. I would like to thank Miriam Sinnhuber and Art Aikin for the very good introduction into the field and their continuous help with questions. In addition to the aforementioned persons, I would also like to thank Stefan Bender for the critical reading and commenting of this thesis. I would like to thank my predecessor in this project, Marco Scharringhausen, for leaving me his well commented and structured retrieval program, as well as for his very well written and easy to read PhD thesis, which was and still is a very good source of information for this work. I would like to thank Patricia Liebing for the many discussions of calibration issues of the SCIAMACHY data, especially the discussion of the polarization correction, which is different for the metals than for most other applications, and I would like to thank Vladimir Rozanov for his help in determining the multiply scattered part of the light in the visible region for the Na retrieval. I would like to thank my co-authors, in addition to the already mentioned people these were Wuhu Feng, John Plane, Dan Marsh, Diego Janches, Tim Dunker and Ulf-Peter Hoppe, for their contributions to my publications, which in parts were used in this thesis. I would also like to thank my many other co-workers at the IUP in Bremen for being nice colleagues and helping me with any kinds of questions and issues that I had. I would like to thank the European Office of Aerospace Research and Development (EOARD) for funding this project through grant# FA8655-09-1-3012 and providing travel funds through the Window on Science program.