Trace gas concentration retrieval from short-wave infrared nadir sounding spaceborne spectrometers

Philipp Hochstaffl

München 2021

Trace gas concentration retrieval from short-wave infrared nadir sounding spaceborne spectrometers

Philipp Hochstaffl

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Philipp Hochstaffl
aus Innsbruck, Österreich

München, den 06.09.2021

Erstgutachter: Prof. Dr. Mark Wenig

Zweitgutachter: Prof. Dr. Thomas Trautmann

Tag der mündlichen Prüfung: 20.01.2022

To my loving parents who gave me the opportunity to build on my talent and offered so much more

Zusammenfassung

Aus der Beobachtung reflektierter Sonnenstrahlung im kurzwelligen Infrarot (SWIR) können Spurengaskonzentrationen in der Erdatmosphäre abgeleitet werden, wobei die Lösung des inversen Problems eine Schätzung des wahren Atmosphärenzustands liefert. Die Inversionsmethode BIRRA (Beer InfraRed Retrieval Algorithm) ist einer von mehreren am DLR (Deutsches Zentrum für Luft- und Raumfahrt) am Institut für Methodik der Fernerkundung (IMF) entwickelten Algorithmen zur Bestimmung von Molekülkonzentrationen aus spektroskopischen Messungen. Das Vorwärtsmodell von BIRRA basiert auf dem ebenfalls am DLR-IMF entwickelten Generic Atmospheric Radiation Line-by-line Infrared Code (GARLIC).

Am Anfang stand die Validierung der mit BIRRA abgeleiteten Kohlenmonoxid (CO) Gesamtsäulen aus SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) Messungen im $2.3\,\mu m$ Bereich. Dazu wurden die BIRRA Gesamtsäulen mit jenen der bodengebundenen Beobachtungsstationen der Netzwerke TCCON (Total Carbon Column Observing Network) und NDACC (Network for the Detection of Atmospheric Composition Change) im Zeitraum von 2003–2011 verglichen. Die mit BIRRA ermittelten CO Konzentrationen zeigen eine $\approx 10\,\%$ negative Abweichung und stimmen mit den Ergebnissen ähnlicher Studien anderer Autoren weitgehend überein.

Nach erfolgter Validierung wurden Neuerungen des Strahlungstransportmodells GARLIC in das BIRRA Vorwärtsmodell eingebaut und die Ergebnisse des aktualisierten Inversionsalgorithmus mit jenen des Vorgängers verglichen. Auf Basis von SCIAMACHY Daten wurde numerische Übereinstimmung der Ergebnisse festgestellt.

Anschließend wurde das Vorwärtsmodell mit Blick auf die Verwendung neuester spektroskopischer Liniendaten, wie SEOM–IAS (Scientific Exploitation of Operational Missions – Improved Atmospheric Spectroscopy), erweitert. Um genauere Molekülabsorptionsquerschnitte berechnen zu können, musste das (klassische) Voigt-Absorptionslinienprofil erweitert werden. Der Einfluss der neuen Spektroskopie wurde zuerst auf Basis von SCIAMACHY Messungen untersucht und anhand von Vergleichsrechnungen auf Basis aktueller HITRAN (HIgh-resolution TRANsmission molecular absorption database) und GEISA (Gestion et Etude des Informations Spectroscopiques Atmosphériques) Daten bewertet. Es stellte sich heraus, dass die SEOM–IAS Liniendaten einen signifikanten Einfluss auf die Inversion haben: die Residuen werden kleiner und auch die abgeleiteten CO Konzentrationen unterscheiden sich leicht. Die gleiche Methodik wurde anschließend dazu verwendet, den Einfluss der Spektroskopie für das CO Retrieval aus TROPOMI Messungen zu bestimmen. Dabei zeigten sich die Auswirkungen noch deutlicher – signifikant kleinere Residuen und eine damit einhergehend höhere Genauigkeit (kleinere Fehler) der CO Säulen sowie der (mit-)abgeleiteten Parameter. Desweiteren besteht weitgehende Übereinstimmung mit den Resultaten der SCIAMACHY Studie.

Ein weiterer Teil der Arbeit beschäftigt sich mit Instrumentenfunktionen (auch bekannt

als Instrumentenprofile), speziell mit der Untersuchung einer passenden Parameterisierung der TROPOMI-Funktion im SWIR Band. Die tabellierten TROPOMI Instrumentenprofile können mit geeigneten Parameterisierungen gut modelliert werden. Darüber hinaus konnte der positive Einfluss der SEOM–IAS Spektroskopie auf die spektralen Residuen auch mit einem parameterisierten Instrumentenprofil nachgewiesen werden. Aufgrund der Flexibilität der vorgestellten Parameterisierungen könnten diese auch für zukünftige Sensoren zum Einsatz kommen.

Abschließend wird der Einfluss von Aerosolen im CO Retrieval analysiert. Auf Basis einer einfachen Parameterisierung wurde versucht, die Extinktion bzw. die optische Tiefe (mit) zu bestimmen. In diesem Zusammenhang wurden auch der (klassische) nichtlineare Least Squares und der separierbare Least Squares hinsichtlich des Konvergenzverhaltens beim Lösen des inversen Problems untersucht. Erste Ergebnisse zeigen ein stabiles CO Retrieval unter Verwendung der separierbaren Least Squares Methode, wobei die (mit-)abgeleiteten Aerosol- und Reflektivitätsparameter auf Probleme durch Entartung hinweisen.

Die vorliegende Arbeit hat gezeigt, wie das CO Retrieval aus SCIAMACHY Messungen verbessert werden kann. Mit dem weiterentwickelten BIRRA Code wurden darüberhinaus erfolgreich CO Konzentrationen aus wolkenfreien TROPOMI Messungen bestimmt. Viele Aspekte der Arbeit sind auch für die präzise Konzentrationsbestimmung anderer Moleküle wie CO_2 oder CH_4 von Bedeutung. Damit bietet die vorliegende Arbeit eine valide Grundlage für die Weiterentwicklung.

Abstract

The remote sensing of short wave infrared (SWIR) radiation reflected from the Earth allows to infer atmospheric trace gas concentrations by solving the inverse problem. The retrieval algorithm BIRRA (Beer InfraRed Retrieval Algorithm) has been developed at the DLR (Deutsches Zentrum für Luft- und Raumfahrt) Remote Sensing Technology Institute (IMF) since around 2005 and is one of multiple algorithms to infer molecular concentrations from calibrated radiance spectra. BIRRA's forward model is based on the Generic Atmospheric Radiation Line-by-line Infrared Code (GARLIC) which has also been developed at the DLR-IMF.

First, the BIRRA retrieved carbon monoxide (CO) columns from SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) $2.3\,\mu$m observations from 2003–2011 were validated against eighteen stations from the ground-based networks TCCON (Total Carbon Column Observing Network) and NDACC (Network for the Detection of Atmospheric Composition Change). The BIRRA inferred CO concentrations were found to be $\approx 10\,\%$ low biased which is in large agreement with other similar studies.

Next, the latest updates from the radiative transfer code GARLIC were incorporated in BIRRA's forward model and the physical results of both, the old (but validated) and the latest (updated) BIRRA algorithms were verified and found to be numerically consistent for SCIAMACHY input data.

Subsequently, the forward model was extended by upgrading its capabilities with respect to spectroscopy, i.e., enhanced line models were incorporated in order to utilize latest spectroscopic information from line lists such as the SEOM–IAS (Scientific Exploitation of Operational Missions – Improved Atmospheric Spectroscopy). More specifically, 'beyond Voigt' line profiles were implemented and the impact of the SEOM–IAS spectroscopy was studied with respect to latest compilations of HITRAN (HIgh-resolution TRANsmission molecular absorption database) and GEISA (Gestion et Etude des Informations Spectroscopiques Atmosphériques) for a large set of SCIAMACHY measurements. It was found that the SEOM–IAS line data and corresponding line models have significant impact on the spectral fitting: the residuals become smaller and the retrieved CO concentrations are also slightly different. The same methodology was then applied to study the spectroscopic impact on CO from S5P/TROPOMI measurements. The impact of the SEOM–IAS spectroscopy revealed to be even more pronounced, in particular with respect to the fitting residuals and smaller retrieval errors (higher precision) of the CO and co-retrieved parameters. Overall, the TROPOMI results are in agreement with that found for SCIAMACHY.

A subsequent part of the thesis examines instrument spectral response functions (ISRF), in particular appropriate parameterizations for the TROPOMI's SWIR band responses. A first assessment with tabulated instrument profiles indicates that the parameterized variants can mimic the tabulated responses within ≈ 3–$6\,\%$, depending on

the instrument model and spectral position. The positive impact of the SEOM–IAS spectroscopy on the spectral fitting residuals could also be identified with the parameterized response functions. Moreover, the presented instrument profiles are considered promising candidates for the description of responses from upcoming sensors due to their flexibility.

Finally, the co-retrieval of aerosol parameters in the CO fit is presented. Based on a simple model for the aerosol optical thickness the feasibility to co-retrieve aerosol extinction was investigated. In this context two different inverse solvers, namely the 'classical' nonlinear least squares and separable least squares, were examined with respect to convergence. First results show a stable CO retrieval for the separable least squares solver, however, the co-retrieved aerosol and reflectivity parameters indicate issues due to degeneracies.

This thesis improved the retrieval of CO from SCIAMACHY observations. Moreover, the upgraded BIRRA algorithm successfully retrieved CO concentrations from cloud-free TROPOMI measurements. Many aspects investigated in this study are also relevant for the retrieval of other atmospheric constituents, such such CO_2 or CH_4. The study does hence provide a proven basis for further developments.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Remote sensing methods make a wide range of measurements in order to study various aspects of the Earth system (*Solimini*, 2016). Nowadays, observations are made from various platforms such as satellite, aircraft, balloon and from the surface. In particular, satellite observations have become an important data source to determine the state of the Earth system, including its atmosphere (*Burrows et al.*, 2011).

1.1 Brief overview of Earth observation missions

Earth observations from space-based sensors have been used since the early years of space flight basically starting in April 1960 with the launch of NASA's (National Aeronautics and Space Administration) Television and Infrared Observation Satellite (TIROS; *Kidder and Vonder Haar*, 2010) into a low-Earth orbit. The evolution of remote sensing techniques in the following decades motivated the development of new spaceborne instruments. Particularly, in the last two decades important Earth observation missions were realized by various space agencies and private companies around the globe.

Highly valuable atmospheric scientific missions from the European Space Agency (ESA) include ENVISAT (ENVIronmental SATellite) and some of the Sentinel missions, each incorporating instruments dedicated for the monitoring of atmospheric chemistry and temperature. ENVISAT for example is hosting instruments such as SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY; *Bovensmann et al.*, 1999; *Gottwald and Bovensmann*, 2011), GOMOS (Global Ozone Monitoring by Occultation of Stars; *Kyrölä et al.*, 2004) and MIPAS (Michelson Interferometer for Passive Atmospheric Sounding; *Fischer et al.*, 2008). SCIAMACHY is a diffraction grating spectrometer that records atmospheric absorption from scattered and reflected solar radiation in various viewing geometries while MIPAS employs a Fourier transform spectrometer (FTS) to observe limb emission spectra from the middle and upper atmosphere. The Sentinel-5 Precursor (S5P; *Veefkind et al.*, 2012) is ESA's first satellite within the European Union's Earth observation program Copernicus that is dedicated to the monitoring of atmospheric chemistry. S5P's payload is the Tropospheric Monitoring Instrument (TROPOMI) which consists of a nadir viewing grating spectrometer and is building upon the heritage of its predecessors GOME (Global Ozone Monitoring Experiment; *Burrows et al.*, 1999), GOME-2 (*Munro et al.*, 2016), SCIAMACHY, and OMI (Ozone Monitoring Instrument; *Levelt et al.*, 2006).

Sensors from other space agencies such as the Michelson interferometer of the ACE

(Atmospheric Chemistry Experiment; *Bernath*, 2017) aboard SCISAT from the Canadian Space Agency (CSA) and GOSAT/GOSAT-2 (Greenhouse Gases Observing Satellite; *Kuze et al.*, 2009, 2016) from the Japan Aerospace Exploration Agency (JAXA) use FTS to record atmospheric spectra. The NASA instrument MOPITT (Measurement of Pollution in the Troposphere; *Vargas-Rodríguez and Rutt*, 2009; *Dakin et al.*, 2003) measures emitted and reflected radiance from the Earth using gas correlation spectroscopy. The NASA missions named OCO-2 (Orbiting Carbon Observatory-2; *Crisp et al.*, 2004) and OCO-3 (Orbiting Carbon Observatory-3; *Eldering et al.*, 2019; *Taylor et al.*, 2020), of which the latter is deployed aboard the International Space Station (ISS), employ nadir viewing grating spectrometers to image atmospheric absorption spectra. The Atmospheric Infrared Sounder (AIRS; *Chahine*, 1991; *Chahine et al.*, 2006) on NASA's Aqua satellite and the IASI (Infrared Atmospheric Sounding Interferometer; *Clerbaux et al.*, 2009) instruments aboard the three METOP (METeorological OPerational Satellite) satellites, operated by EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites) measure thermal emission in the nadir direction. In the recent past also private companies such as GHGSat launched missions dedicated for the remote sensing of greenhouse gases and other trace gases (*Varon et al.*, 2018).

1.2 Atmospheric remote sensing

The principle of atmospheric remote sensing measurements from space is based on the interaction of electromagnetic radiation with constituents in the atmosphere (*Solimini*, 2016). Most passive sensors measure the radiation coming from the Sun that is transmitted and scattered through Earth's atmosphere or emitted by the atmosphere itself. Other instruments use solar, lunar or stellar occultation measurements (*Gottwald and Bovensmann*, 2011; *Kyrölä et al.*, 2004).

According to Wien's displacement law the Sun strongly emits in the near-UV (Ultra Violet), VIS (VISible) while radiation emitted by the Earth peaks at much longer wavelengths in the IR (InfraRed) (*Zdunkowski et al.*, 2007). As indicated in Fig. 1.1 the solar radiation at Top of Atmosphere (ToA) which is proportional to the solid angle of the Sun at a distance of one astronomical unit d_{au} (E_{Sun}^{5777K} at ToA $\propto (r_{Sun}/d_{au})^2$, also see *Goody and Yung* (1995, Tab. A.9.1)) is dominant in the NIR (Near InfraRed) range (indicated in gray). Towards the thermal infrared (TIR) the emission from Earth's surface and atmosphere becomes larger and surpasses the solar irradiance at ToA at roughly $2000\,cm^{-1}$. Consequently, the relevance of different processes in radiative transfer through the atmosphere is depending on the observed wavenumber interval.

Interaction of electromagnetic radiation with atmospheric constituents such as trace gases and aerosols generally include absorption, emission, and scattering. However, in the absence of particles[1] scattering is of little importance in the SWIR (Short Wave InfraRed) and thermal emission from Earth's surface and atmosphere is also weak (see Fig. 1.1). Consequently, most of the photons detected by nadir looking space or airborne instruments have taken the path from the Sun to the surface and back, thereby providing nearly homogeneous sensitivity to all atmospheric levels (*Richter, A.*, 2010). Measurements that target the quantification of molecules such as carbon monoxide (CO), methane (CH_4) or carbon dioxide (CO_2) from the SWIR fulfill this ideal situation and are hence employed

---

[1] Solid and liquid, e. g., aerosols, clouds, haze, …
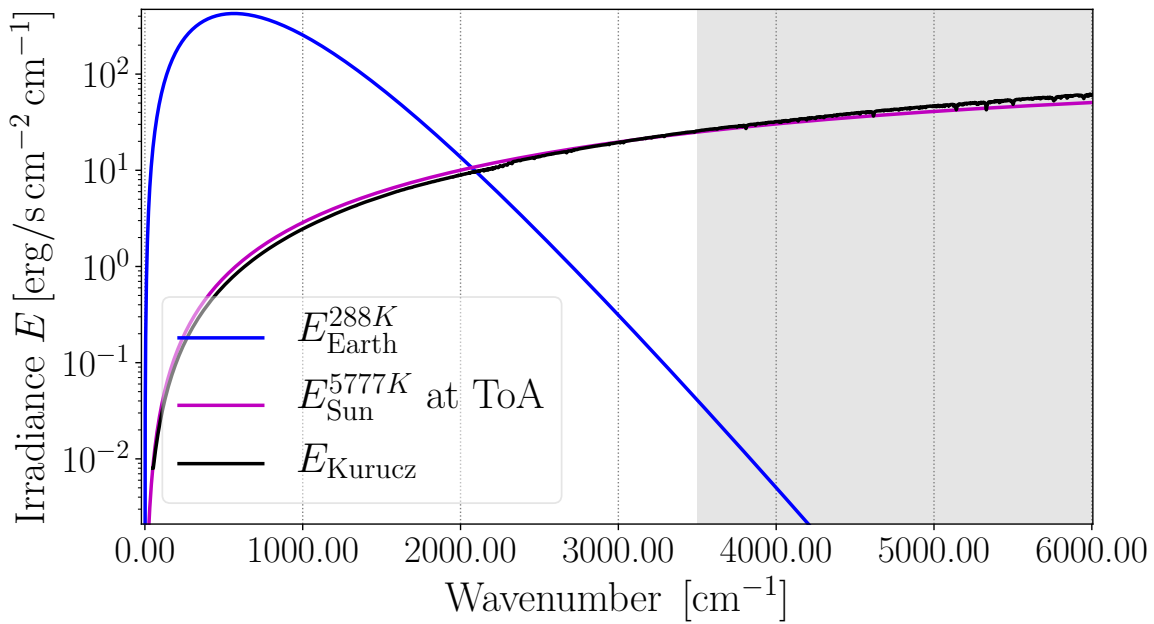
Figure 1.1: Spectral irradiance for a blackbody at $288\,\mathrm{K}$ representing the Earth's surface temperature. The solar blackbody irradiance at a temperature of $5777\,\mathrm{K}$ is distance scaled by one astronomical unit (see Sec. 1.2). For comparison the Kurucz solar reference spectrum is shown too (*Kurucz*, 2005).

by many satellite instruments (e.g. *Frankenberg et al.*, 2005; *Buchwitz et al.*, 2007a,b; *Deeter et al.*, 2009; *Veefkind et al.*, 2012). A downside, however, is that the neglect of some processes in the SWIR radiative transport limits its application, e.g., absence of scattering limits the signal-to-noise (SNR) ratio of measurements, particularly over dark surfaces such as water (*Richter, A.*, 2010).

The retrieval of trace gas concentrations from spectroscopic measurements constitutes an inverse problem. The spectral dependence of molecular absorption is used to identify atmospheric species in the measurements. In order to deduce the amounts of molecules from the measured spectrum the observed absorption is compared to modeled spectra from a radiative transfer model. The goal of the retrieval is to find the inverse solution of the forward model, i.e., parameters in the model that can not be observed directly. In general, since the relation between radiance $\boldsymbol{y}$ and the state vector $\boldsymbol{x}$ is often not linear

$$\boldsymbol{y} \;=\; \boldsymbol{F}(\boldsymbol{x})\,, \tag{1.1}$$

an iterative approach is required (details see Sec. 3.2.2).

In atmospheric remote sensing the forward problem $\boldsymbol{F}$ is described by a radiative transfer model which encompasses the parameters of interest. The so called retrieval parameters make up the state vector for the forward model. In the retrieval algorithm the set of parameters is updated/modified until agreement is found within predefined limits with the measurement, e.g., by testing the relative change of the state vector $\boldsymbol{x}$. In general, an initial guess or a priori estimate needs to be provided for the state vector. Moreover, the forward model requires input of non-fitted quantities that are assumed to be sufficiently well known such as vertical pressure and temperature profiles as well as the observation geometry, molecular line data etc.

1.3   Carbon monoxide

1.3.1   Atmospheric chemistry

The background concentrations of CO vary from $\approx 50 - 100$ ppbv, however, these levels can be increased many times in polluted urban areas or in environments with extensive biomass burning. Its atmospheric lifetime in the order of a few weeks makes it an atmospheric tracer for the transport of pollutants on the global scale (*Holloway et al.*, 2000).

Large anthropogenic sources of CO are the incomplete combustion of fossil fuels or biomass (*Burrows et al.*, 2011). Another main source of CO in the atmosphere is the oxidation of CH_4 and other non-methane hydrocarbons (NMHC) which in turn constitute a main sink for the hydroxyl radical OH in the troposphere (*Jacob*, 1999, Sec. 11.2). The net mechanism for the oxidation of CH_4 yields

$$CH_4 + 3\,OH + 2\,O_2 \longrightarrow CO_2 + 3\,H_2O + HO_2 \;\cdot \qquad\qquad (R1)$$
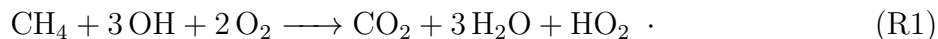
In the cascade of reactions in React. (R1) an intermediate reaction of CH_4 with the OH radical yields CO as a product so that the source of CO depends on the distribution of the hydroxyl radicals and hydrocarbons (*Jacob*, 1999, Sec. 11.3.3). It is estimated that NMHC oxidation causes 40–60 % of surface CO levels over the continents, slightly less over the oceans, and 30–60 % of CO levels in the free troposphere (*Poisson et al.*, 2000). According to Fig. 1.2 its vertical distribution is rather equal across various climatological regions in the troposphere but differs above.
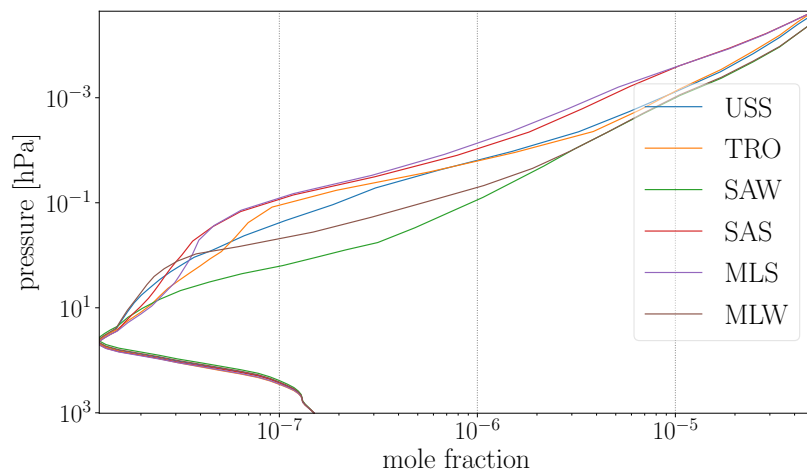
Figure 1.2: Carbon monoxide mole fraction vertical profiles for various reference atmospheres (*Anderson et al.*, 1986), i.e., US-Standard (USS), tropical (TRO), subarctic winter and summer (SAW and SAS), and midlatitude summer and winter (MLS and MLW).

A primary loss mechanism of CO in the Earth's atmosphere is its oxidation by the OH which hence constitutes another leading sink of the radical. Moreover, CO is a main determinant of tropospheric air quality, particularly in polluted areas, since CO serves as a precursor for ozone in the troposphere by affecting the catalytic production and destruction of ozone (*Jacob*, 1999, Sec. 11.3). Whether tropospheric ozone is produced or destroyed in the oxidation of CO is depending on the availability of nitrogen oxides (NO_x), which comprises nitric oxide NO and nitrogen dioxide NO_2. Both are typically created in combustion processes at high temperatures. Although NO_x is emitted mainly as NO, cycling between NO and NO_2 takes place in the troposphere on the time scale

of a minute (*Jacob*, 1999, Sec. 11.4). In case of high NO_x concentrations (≈ 100 pptv at 1 atm or 1013.25 hPa) the net reaction producing ozone is (*Holloway et al.*, 2000)

$$CO + 2\,O_2 \longrightarrow CO_2 + O_3 \;\cdot \tag{R2}$$
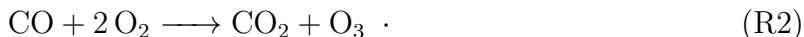
This net reaction indicates that the presence of NO_x allows the regeneration of OH so that no hydroxyl radicals are consumed in React. (R2). In a clean environment without sufficient NO_x concentrations the net reaction is a O_3 and OH destruction

$$2\,CO + O_3 + OH \longrightarrow 2\,CO_2 + HO_2 \;, \tag{R3}$$
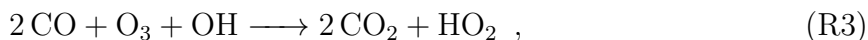
which yields a hydroperoxyl radical HO_2. In Reactions (R2) and (R3) the greenhouse gas carbon dioxide is produced.

1.3.2   Rotation-vibration bands in the SWIR

Carbon monoxide has a diatomic configuration where a carbon and oxygen atom join together (see Fig. 1.3). The asymmetric charge distribution causes a permanent dipole moment in the molecule (*Zdunkowski et al.*, 2007). The spectral characteristics of diatomic molecules can be derived from the Schrödinger equation (*Hanel et al.*, 2003).
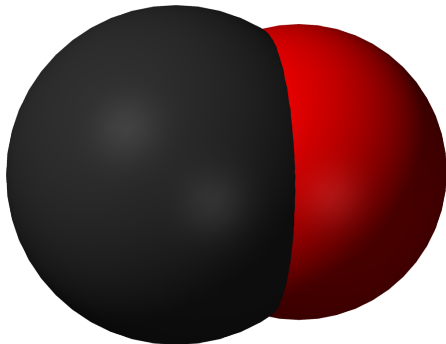
Figure 1.3: Carbon monoxide consists of one carbon atom (black) and one oxygen atom (red). Both are connected by a triple bond that consists of two π bonds and one σ bond. The computed fractional bond order is 2.6 (*Martinie et al.*, 2011).

In the SWIR and TIR spectral regions CO displays P and R branches (see Fig. 1.4 and *Zdunkowski et al.*, 2007, Fig. 8.10). These branches are caused by transitions where the vibrational transition number $\Delta v = v' - v$ is fixed (in general $v = 0$ and $v' = 1, 2, 3$) while the rotational transition number $\Delta J = J' - J$ varies. Transitions in the TIR correspond to the fundamental band ($\Delta v = 1$) while the SWIR spectrum of the molecule between $4250 - 4350\,cm^{-1}$ and $6350 - 6450\,cm^{-1}$ belong to the first overtone ($\Delta v = 2$) and second overtone ($\Delta v = 3$), respectively. The energy of states within each band are characterized by the rotational quantum number J'. Based on selection rules the transition between rotational energy states are one quantum rotation level away $\Delta J = \pm 1$. In the P branch $\Delta J = -1$ and in the R branch it is $\Delta J = +1$. Consequently, every absorption line in the CO spectrum can be assigned to a change in vibrational state and changes of $\Delta J = \pm 1$ in the rotational quantum number.

1.4   Space-based measurements of CO concentrations

In the past decades, multiple instruments were launched to quantify global CO concentrations from space. The first space-based CO measurements were made in 1981 and

1984 via the MAPS (Measurement of Air Pollution from Satellites; *Newell et al.*, 1988; *Reichle Jr. et al.*, 1999) instrument aboard the Space Shuttle. The data acquired proved the feasibility of the measurement technique and many of the gross features of global CO, e. g., biomass burning, anthropogenic pollution, and North/South hemispheric gradients where identified.

The advent of MOPITT measurements in early 2000 allowed to identify major sources of CO emissions for the first time (*Deeter et al.*, 2003; *Edwards et al.*, 2004) and revolutionized the understanding of natural and anthropogenic tropospheric pollution. The MOPITT sensor was designed to retrieve CO profiles from thermal emission and absorption as well as from reflected SWIR radiation, however, the SWIR channels suffered from instrumental issues and were not used for CO retrieval (*Clerbaux et al.*, 2008).

Carbon monoxide measurements are also available from AIRS (*McMillan et al.*, 2005). The data has been used to look at long range transport of pollution and were also compared to MOPITT and ground-based observations (*Warner et al.*, 2007). A sensor aboard NASA's Aura spacecraft named TES (Tropospheric Emissions Spectrometer; *Rinsland et al.*, 2006) provided vertical distributions of CO that were also validated against MOPITT observations (*Luo et al.*, 2007; *Kopacz et al.*, 2010). The Michelson interferometer IASI also measures TIR emission spectra and contributes to the suite of CO observations *George et al.*, 2009, 2015. The GOSAT-2 mission from JAXA is capable of recording CO absorption in the $4.200 - 4.300\,\mathrm{cm}^{-1}$ spectral range (*Suto et al.*, 2021). Yet another mission that measured CO concentration profiles in limb viewing geometry from the mid-troposphere to the thermosphere was the ACE-FTS (*Bernath et al.*, 2005; *Cler-*
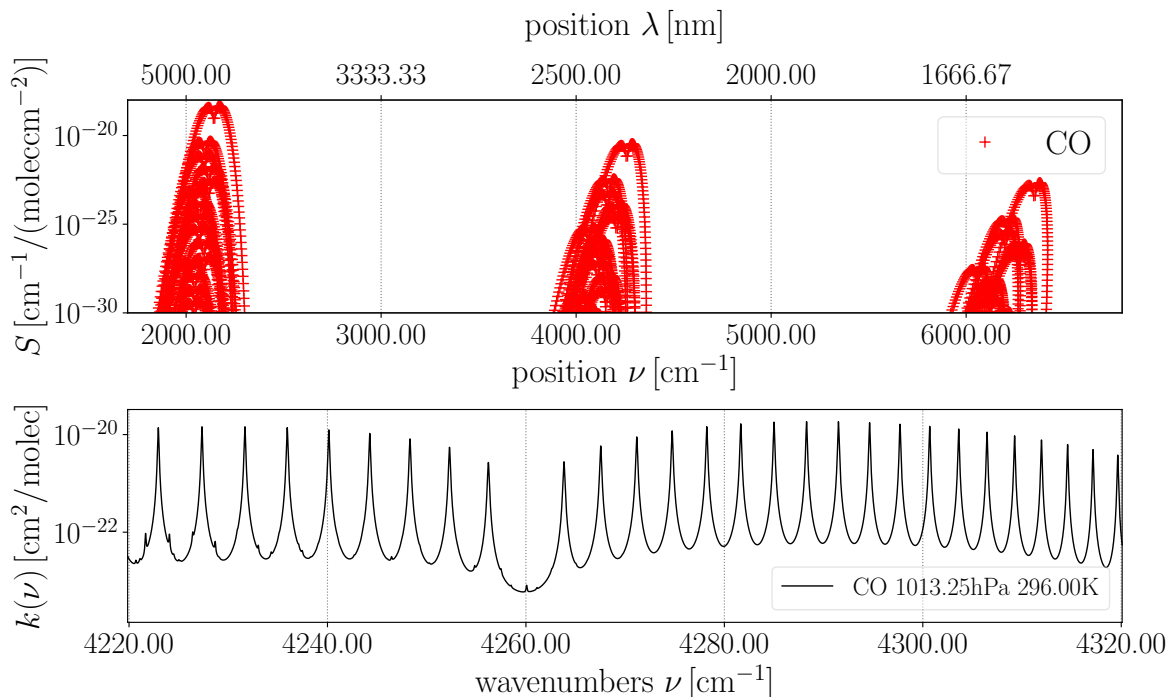
Figure 1.4: Line strengths of CO according to the GEISA 2020 line data from the TIR to NIR (top). The SWIR absorption lines of CO around $4000\,\mathrm{cm}^{-1}$ correspond to the first overtone ($v' = 2$). Absorption cross sections for standard atmospheric conditions in the SWIR are depicted below. Rotation-vibration bands $< 4260\,\mathrm{cm}^{-1}$ correspond to the P branch while transitions $> 4260\,\mathrm{cm}^{-1}$ are attributed to the R branch.

baux et al., 2005). Also the MIPAS instrument was providing limb observations of CO in the TIR around $2100\,\mathrm{cm}^{-1}$ (fundamental band $v' = 1$). Two other instruments, namely SCIAMACHY and TROPOMI, are of particular interest for this thesis. Both provide measurements in the SWIR from which CO concentrations can be inferred by fitting its transmission in the first overtone (see Fig. 1.4).

(a)   **(b)**
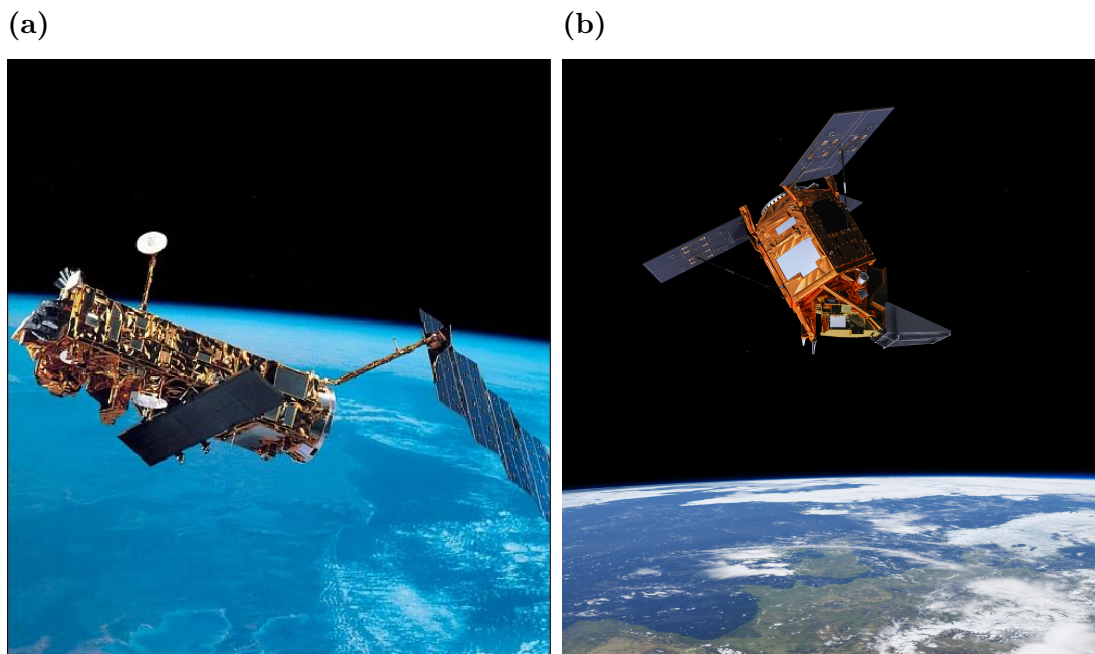
Figure 1.5: The **(a)** ENVISAT and **(b)** Sentinel-5P spacecrafts in low Earth orbit (LEO) over Europe (image credit: ESA).

1.4.1   SCIAMACHY aboard ENVISAT

The SCIAMACHY instrument was part of ENVISAT's atmospheric mission and was launched on March 1, 2002 (*Gottwald and Bovensmann*, 2011). SCIAMACHY was a passive absorption spectrometer in the UV–SWIR range and consisted of eight spectroscopic channels with the SWIR channel 8 spanning the wavelength range from 2259–2386 nm. The nominal spectral and spatial resolution in the SWIR was $0.26\,\mathrm{nm}$ and $120\,\mathrm{km}\times30\,\mathrm{km}$ in nadir viewing geometry, respectively. The mission was lost way beyond its nominal lifetime of 5 years in April 2012.

Although the instrument encountered issues with its SWIR channels (*Gloudemans et al.*, 2005), CO data has been retrieved (*Buchwitz and Burrows*, 2003; *Buchwitz et al.*, 2004; *Frankenberg et al.*, 2005; *Buchwitz et al.*, 2007a; *Gloudemans et al.*, 2009; *Gimeno García et al.*, 2011) and validated (*Dils et al.*, 2006; *de Laat et al.*, 2006, 2007; *Borsdorff et al.*, 2016; Pub. I) with ground-based FTIR (Fourier transform infrared) measurements.

1.4.2   TROPOMI aboard S5P

A rather new sensor that measures CO in the SWIR is the passive grating imaging spectrometer (a pushbroom instrument) TROPOMI (*Kramer, H. J.*, 2021). The S5P, launched

on October 13, 2017, is a single payload satellite mission of the Copernicus Space Component (CSC) and its instrument shares many common features with SCIAMACHY. However, multiple new developments and a number of stringent performance requirements make TROPOMI a rather different sensor with superior performance *Vonk* (2017); *Kleipool et al.* (2018); *van Kempen et al.* (2019). It has seven spectrometers spanning from UV/VIS to the SWIR and its absolute radiometric accuracy in the SWIR channels is specified to be within 6%. The SWIR module has a spectral range from 2305 to 2385 nm ($\approx$ 4193–4338 cm^{-1}) which constitutes a subset of SCIAMACHY's channel 8. Its nominal spectral resolution of the channel is 0.25 nm with a sampling interval < 0.1 nm. The minimum radiometric resolution (SNR) in the SWIR ranges from 100–120 and it has a spatial resolution of 7.5×7.5 km in the nadir direction.

1.5   Retrieval algorithms in the SWIR

1.5.1   Competing algorithms

As indicated in Sec. 1.2 the retrieval of molecular concentrations from radiance or transmission spectra poses an inverse problem that is typically solved by optimization algorithms. Several codes have been developed for the analysis of SWIR spectra at different European institutes, e.g., the Weighted Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS; *Buchwitz et al.*, 2004, 2005) algorithm, the Iterativ Maximum A Posteriori (IMAP-DOAS; *Frankenberg et al.*, 2005) method, the Iterative Maximum Likelihood Method (IMLM; *Gloudemans et al.*, 2005), the Shortwave Infrared CO Retrieval (SICOR; *Borsdorff et al.*, 2017, 2018) algorithm, and the Beer InfraRed Retrieval Algorithm (BIRRA; *Gimeno García et al.*, 2011; Pub. I).

The WFM-DOAS algorithm approximates the logarithm of the Sun-normalized measured intensity by linearization around an initial guess according to a first order Taylor expansion plus a low order polynomial (*Buchwitz et al.*, 2000, Eq. (1)). The difference between the radiance of the observed and initial guess spectrum is described by fitting altitude independent factors that scale weighting functions for the parameters of interest. Because of the linearization of the forward model it is possible to apply a linear least squares fit to determine the parameters from SWIR observations.

The IMAP DOAS inverse method uses optimal estimation in order to find the state vector that maximizes the a posteriori probability function(*Frankenberg et al.*, 2005). The state vector for the retrieval comprises scaling factors for the molecular concentrations at different layer heights, a climatoligical index for temperature changes in the atmsphere and polynomial coefficients accounting for low frequency absorption and scattering. Moreover, it computes the atmospheric transmission on a monochromatic grid, i.e., before convolution with the instrument's spectral response. This is particularly important in the SWIR where narrow lines that exhibit rather strong absorptions are measured at a moderate spectral resolution.

The IMLM algorithm is based on scaling a priori profiles (*Gloudemans et al.*, 2008). The retrieval fits the modeled radiance to the measurements by varying the total amounts of the trace gases using a weighted least squares fit. Beside profile scaling factors the the algorithm provides estimates for the coefficients of the low order polynomial which accounts for effective surface albedo, continuum effects or smooth variations of the surface albedo with wavelength.

The inverse method SICOR also uses a profile scaling approach which is considered a kind of regularization of the CO profile retrieval (*Landgraf et al.*, 2016). The algorithm employs an unconstrained least squares fit to adjust CO abundances by scaling a reference profile. Another regularization is introduced for the inversion of cloud and surface parameters as the latest SICOR variant accounts for light path enhancements by a two stream radiative transfer solver. However, for the retrieval of CO from clear sky observations the fit of a scattering layer is ignored.

1.5.2   Beer Infrared Retrieval Algorithm (BIRRA)

The BIRRA algorithm has been developed by the Remote Sensing Technology Institute (IMF) at the German Aerospace Centre (DLR) and its least squares solvers are provided by the PORT optimization library (*Fox et al.*, 1978). To infer trace gas concentrations the inverse method performs a nonlinear least squares fit (NLS; *Rust*, 2002; Sec. 3.2.2) of the observed radiance by scaling initial guess profiles. The BIRRA code features a separable least squares solver (SLS; *Gay*, 1990; Sec. 3.2.3) with optional bounds to specify physical constraints, exact analytical derivatives, and line-by-line computation of molecular cross sections.

The forward model in BIRRA (*Gimeno García et al.*, 2011) is based on the Generic Atmospheric Radiation Line-by-line Infrared Code (GARLIC; *Schreier et al.*, 2014). The molecular spectroscopic parameters for the line-by-line calculations are obtained from, e.g., HITRAN (HIgh-resolution TRANsmission molecular absorption database; *Rothman et al.*, 2009) or GEISA (Gestion et Etude des Informations Spectroscopiques Atmosphériques; *Jacquinet-Husson et al.*, 2008) databases. Moreover, the CKD (Clough, Kneizys, Davies continuum; *Clough et al.*, 1989) absorption model and Collision Induced Absorption (CIA; *Richard et al.*, 2012; *Karman et al.*, 2019) can be considered. As a forward model for nadir observations, GARLIC provides model spectra for up- and down-looking observation constellations in a spherical geometry.

Beside an operational version there is a scientific (prototype) variant of the BIRRA algorithm. The former designates the operational SCIAMACHY Level-1b $\rightarrow$ 2 processor for the CO and CH_4 (*Hamidouche et al.*, 2016) products. As an operational ESA processor this variant is subject to the European Cooperation for Space Standardization (ECSS) standards and is hence not used for scientific retrievals or prototyping.

The scientific BIRRA described in *Gimeno García et al.* (2011) incorporates additional enhancements over the operational variant which are particularly tailored to mitigate the adverse impact of the SCIAMACHY instrument's degrading SWIR spectra from channel 8 (*Gloudemans et al.*, 2005; *Lichtenberg et al.*, 2010). Moreover, the framework of the scientific processor allows for more flexibility in the retrieval setup. In order to enable processing of large datasets with the prototype BIRRA a suite of Python and C-Shell scripts were developed at DLR (S. Gimeno García 2016, personal communication). This framework prepares files and tasks required by the retrieval algorithm to assimilate model spectra upon measurements. The scientific BIRRA variant with the prescribed features was the basis for the developments described in this thesis.

1.6   Motivation and objectives

Validation studies of SCIAMACHY inferred CO with ground-based remote sensing observations were conducted for multiple different retrieval algorithms in the past, e.g., *Dils et al.* (2006); *Sussmann and Buchwitz* (2005); *de Laat et al.* (2010); *Kerzenmacher et al.* (2012); *Borsdorff et al.* (2016). Most of these studies found that the retrieval and validation of CO from SCIAMACHY is challenging because single SCIAMACHY CO measurements show instrument noise errors up to 100% of the total CO column value—in particular for the later years of the mission (i.e. from 2006 onward). However, a similar verification and validation is considered a crucial step in the BIRRA algorithm refinement by *Gimeno García et al.* (2011, Sec. 5). Moreover, assessing the quality of BIRRA retrieved CO concentrations by a true reference throughout its further developments becomes even more important as the algorithm should be applied to data from latest sensors. Nonetheless, reprocessing SCIAMACHY observations with state of the art retrieval algorithms remains relevant for long-term consistency and to provide a homogeneous multi-mission time series for CO.

A different issue that applies across sensors is spectroscopic uncertainty, in particular uncertainties in the line intensities and broadening parameters. Both can cause systematic errors in the retrieval product and multiple studies called attention to this issue in the past. Quite recent studies by *Galli et al.* (2012) and *Checa-García et al.* (2015) found that spectroscopic errors in the $2.3\,\mu$m band can induce errors that exceed TROPOMI's error budget and that further efforts should be directed to improve the H_2O (water) and CH_4 spectroscopy in this regime. Earlier studies such as, e.g., *Frankenberg et al.* (2005) pointed out the classical Voigt profile, i.e., the to date standard in the modeling line-by-line radiative transfer in the atmosphere, is not a fully accurate representation of the spectral line shape observed in gas mixtures and that narrowing mechanisms such as line-mixing should be taken into account in the calculation of molecular absorption cross sections. In preparation of the S5P mission, ESA commissioned the compilation of a new spectroscopic database in the SWIR according to the needs of the TROPOMI instrument. The project yielded the SEOM–IAS (Scientific Exploitation of Operational Missions – Improved Atmospheric Spectroscopy; *Birk et al.*, 2017a,b) line list.

Another important facet of spectroscopic measurements is modeling the instrument's line shape (ILS). In order to accurately model variations in the measured signal to changes in the optical depth precise knowledge of the instrument's spectral response function is indispensable (*Frankenberg et al.*, 2005, Fig. 2) (*Gloudemans et al.*, 2005). While often a Gaussian slit function is used (e.g. *Gimeno García et al.*, 2011) sensors such as TROPOMI provide tabulated response functions that significantly differ from the Gaussian shape. However, often a parameterized model is used in the Level-1b $\rightarrow$ 2 processing as it allows for more flexibility in the retrieval.

Yet another important issue in the retrieval of atmospheric constituents from space is the fact that, beside molecular absorption, scattering by aerosols (and molecules) contributes to the radiative transfer (*Burrows et al.*, 2011, Chap. 6). Although scattering is of little importance for clear sky observations in the SWIR, aerosol loaded scenes or cirrus clouds can modify the light path in a way that leads to an inaccurate estimation of the true concentration of, e.g, CO or CH_4, if not appropriately accounted for (*Landgraf et al.*, 2016).

The prescribed aspects lead to the formulation of four objectives which are addressed

in this thesis.

(1) Validation of the retrieval algorithm — in order to quantify the accuracy of the prototype BIRRA inferred CO columns from SCIAMACHY observations a thorough intercomparison against globally distributed ground-based observations should be carried out. The validation should be performed for all SCIAMACHY measurements from $2003 - 2011$, which basically encompasses the full-mission dataset.

(2) Impact of spectroscopy — line-by-line radiative transfer codes require spectroscopic line data and models to calculate molecular cross sections in order to describe absorption. Latest scientific advances in the field of SWIR nadir retrievals indicated that the accurate computation of cross sections is crucial for current missions such as, e. g., S5P/TROPOMI or the GOSAT and OCO missions but also for upcoming missions such as the CO_2 Monitoring (CO2M) mission in order to meet the specified requirements. The impact of most recent molecular line lists such as SEOM–IAS, HITRAN, or GEISA on the retrieval of CO in the $2.3\,\mu$m region should be assessed for TROPOMI measurements, but also for heritage missions such as SCIAMACHY.

(3) Instrument line shape — instrument spectral response functions (ISRFs) are required for the convolution of the monochromatic line-by-line spectrum to instrumental resolution and incorrectly modeled instrument line shapes can cause errors in the retrieved columns. In order to account for modifications of the slit function caused by, e. g., heterogeneous scenes or by changes in the instrument characteristics, tabulated spectral response functions from on-ground calibration should be replaced by an appropriate parameterized model which allows, e. g., for variations in a pixel's center frequency or its half width. Appropriate parameterizations for TROPOMI's SWIR band should be developed and their applicability should be demonstrated in the retrieval of CO from TROPOMI measurements.

(4) Aerosol extinction — aerosols contribute to optical depth in the Earth's atmosphere. In order to account for aerosol extinction in the CO retrieval an adequate parameterized model should be examined. The retrieval should be modified so that the state vector includes aerosol parameters which are then co-retrieved with CO concentrations. The model and implementation should be assessed with observations, e. g., from TROPOMI.

1.7   Outline

The subsequent chapters are organized as follows. In Chapters 2 and 3 the methodology with respect to the forward modeling and inversion is described, respectively. Chapter 4 is dedicated to the results of the thesis. More specifically, Secs. 4.1, 4.4, and 4.5 are summaries of the papers published in peer-reviewed journals. Section 4.2 provides more discussions on the averaging kernels that were relevant for Pub. I. Section 4.3 briefly discusses the results from the verification of the retrieval algorithm upgrade. Section 4.6 presents findings with respect to instrument line shape parameterizations and tabulated response functions. Section 4.7 discusses the feasibility of aerosol parameter fits in the SWIR. Finally, Chap. 5 concludes and summarizes the results and provides an outlook for future studies.

Note that some results regarding higher order line profiles—relevant for Pub. II and Pub. III—are already presented in Chap. 2 to facilitate readability. Another exemplary result, demonstrating the application of inverse methods and adequate postprocessing for a set of TROPOMI observations, is also presented in Chap. 3.

Chapter 2

Methods: Radiative transfer — forward modeling

2.1 Blackbody radiation

Electromagnetic radiation emitted by the Sun is the primary source of radiative energy on Earth. The Earth's atmosphere is heated by absorption of solar radiation, and cooled by emission of thermal infrared radiation (*Stamnes et al.*, 2017). The assumption that the local thermal emission is balanced by the local rate of heating due to absorption of radiation at all wavelengths is called local radiative equilibrium (LTE). It allows to assign a local thermodynamic temperature which is particularly important for radiative transfer modeling (details see Sec. 2.2 and *Fischer and Hase*, 2015).

Although the Sun is not an ideal blackbody (absorption $\mu_a = 1.0$ independent of frequency) described by Planck's law (*Weinberg*, 2015, Chap. 1) the physics that lead to the spectral energy distribution of a blackbody is of fundamental importance for passive remote sensing (*Pierrehumbert*, 2010, Sec. 3.2). In order to adequately explain the frequency distribution of radiation from a blackbody it is necessary to go beyond classical physics and take the quantum nature of the electromagnetic field into account (*Weinberg*, 2015, Sec. 1.1).

Max Planck (1858–1947) and Albert Einstein (1879–1955) proposed the assumption that the spectral energy could exist only in discrete quanta which are proportional to the frequency according to

$$E \ = \ h \cdot f \ = \ h\,c\,\nu \, , \tag{2.1}$$

where c designates the speed of light. The discrete energy parcels $h{\cdot}f$ of an electromagnetic field at a certain frequency are called photons with $h = 6.62607015{\cdot}10^{-27}$ (erg s) designating Planck's constant which relates a photon's energy to its frequency f (or wavenumber $\nu = f/c$). It is important to note that the energy of electromagnetic waves is quantized even in free space, i.e., when the waves are not constrained. Moreover, it is important to note that there is no finite lower limit or upper limit on the possible energy of a photon since the frequency is continuous (see Eq. (2.1)).

In order to model the spectral energy radiated from a hot cavity the solutions for electromagnetic standing waves (also known as resonant or normal modes) in a confinement according to Eq. (2.2)) are considered (*Hanel et al.*, 2003, Chap. 1).

All possible solutions for standing waves in a resonator with physical boundaries $x = \{0, L\}$ have different frequencies according to

$$f_n = \frac{n\,c}{2\,L}\,, \tag{2.2}$$

with $n \in \mathbb{N}^+$, f_1 being known as the fundamental frequency and $f_{n>1}$ the overtones. The solutions for these normal modes of the harmonic oscillator are then given by

$$E_n = 2\,E_0\,\sin\left(\frac{n\pi x}{L}\right)\cos(2\pi f_n t)\,, \tag{2.3}$$

and possible solutions for E are now limited by f_n (see Fig. 2.1).
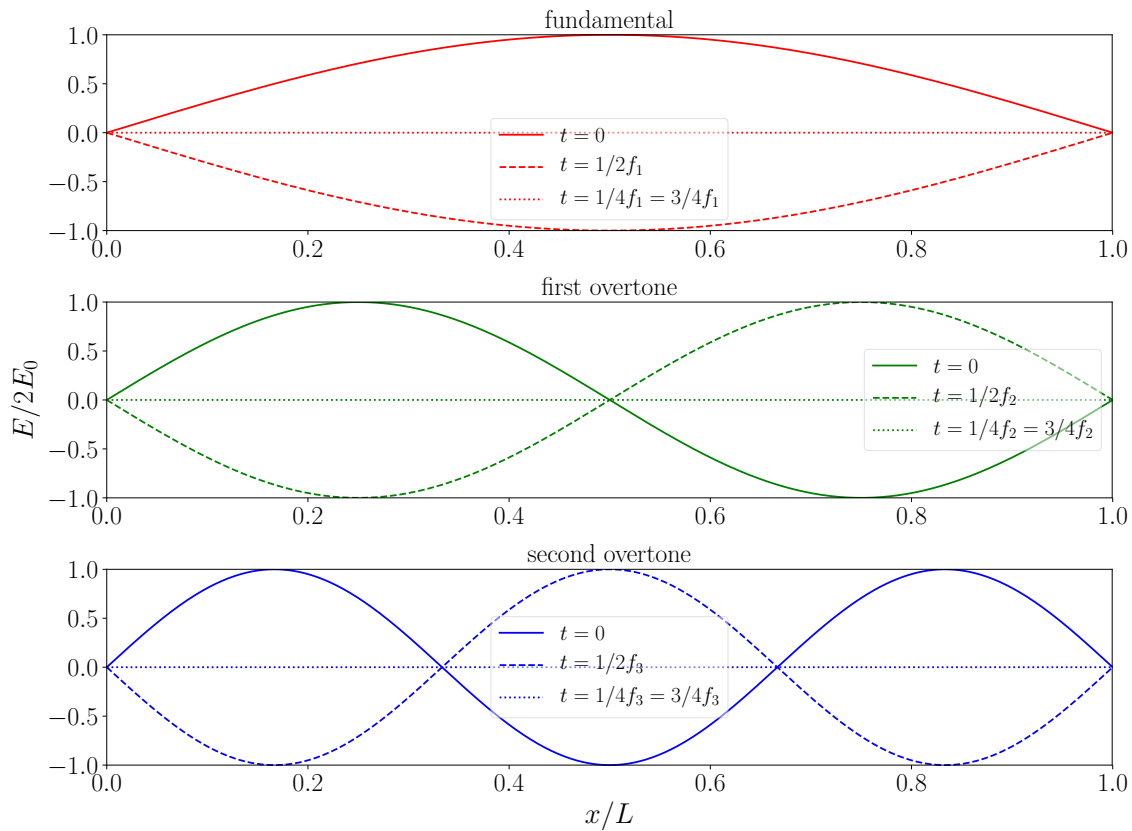
Figure 2.1: The normal modes of the harmonic oscillator for $n = 1, 2, 3$. Note that every point x behaves like a separate oscillator with amplitude $2\,E_0\,\sin\frac{n\pi x}{L}$.

In thermal equilibrium and for the classical limit (corresponds to low particle densities and high temperatures T) the probability for the occupancy of modes of energy $E_n = n\,h\,\nu$ is described by the Boltzmann distribution (P_B; *Rybicki and Lightman*, 2008, Sec. 1.5) according to

$$P_\mathrm{B}(n, T) \;=\; \frac{\exp\left(-E_n/k_\mathrm{B}T\right)}{\sum_{n=0}^{\infty}\exp\left(-E_n/k_\mathrm{B}T\right)}\,, \tag{2.4}$$

where $k_\mathrm{B} = 1.380649 \times 10^{-16}\,\mathrm{erg\,K^{-1}}$ represents the Boltzmann constant (Ludwig Boltzmann 1844–1906). The average energy of a harmonic oscillator (mode) with wavenumber

ν is then given by the energy of the mode times the probability that the mode will be occupied according to

$$E(\nu, T) \;=\; \sum_{n=1}^{\infty} E_n \, P_{\mathrm{B}}(n, T) \;=\; \frac{h \, c \, \nu}{\exp\left(\frac{hc\nu}{k_{\mathrm{B}}T}\right) - 1} \,. \tag{2.5}$$

The power emitted per volume and wavenumber interval ν to $\nu + \mathrm{d}\nu$ is then given by the product of Eq. (2.5) and the number of modes per interval in a volume L^3

$$\frac{\mathrm{d}N}{L^3} \;=\; 8 \, \pi \, \nu^2 \, \mathrm{d}\nu \,, \tag{2.6}$$

according to Planck's formula

$$B(\nu, T) \;=\; \frac{2 \, h \, c^2 \, \nu^3}{\exp\left(\frac{hc\nu}{k_{\mathrm{B}}T}\right) - 1} \,. \tag{2.7}$$

The radiated power per volume and wavenumber interval, i.e., the spectral radiance, only depending on the object's temperature (see Fig. 1.1). In the limit of small wavenumbers and large temperatures Planck's formula gives the Rayleigh-Jeans law while for large wavenumbers the formula reduces to Wien's law. Geometric considerations account for the factor $(\pi c)/4$ in Eq. (2.7) with respect to the product of Eqs. (2.5) and (2.6).

As indicated above the spectral radiance B describes the power per unit projected area $\mathrm{d}A_\perp$ into a unit solid angle $\mathrm{d}\Omega$ per wavenumber $\mathrm{d}\nu$. The solid angle $\mathrm{d}\Omega$ is defined as the section of a sphere with radius r according to

$$\mathrm{d}\Omega \;=\; \frac{r \, \mathrm{d}\varphi \, r \sin\theta \mathrm{d}\theta}{r^2} \;=\; \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi \,. \tag{2.8}$$

Passive remote sensors aboard satellites quantify the energy received per area, direction and time by looking at a surface from a specified viewing angle. This measurement constitutes an estimate of the instantaneous power $I \approx B(\nu, T)$. Irradiance E, the quantity depicted in Fig. 1.1, is a hemispheric flux since the power is received per surface area $A_\perp$ from all directions of a hemisphere (half space). Given that the emitted or received radiance is isotropic, i.e. not depending on the zenith and azimuth angles, the source is said to be Lambertian with irradiance E and radiance related by π according to

$$I \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi/2} \sin\theta \, \mathrm{d}\theta \;=\; \pi \, I \;=\; E \,. \tag{2.9}$$

2.2   The radiative transfer equation

The radiative transfer equation is a macroscopic description for the exchange of energy between the radiation field and the medium it passes through and gives the specific intensity of radiation I at wavenumber ν during its propagation by a distance s (*Goody and Yung*, 1995; *Zdunkowski et al.*, 2007; *Stamnes et al.*, 2017). Due to the conservation of energy, for any term that introduces a loss there must be a term that introduces a gain. The equation states that the radiance I is subject to losses due to extinction (absorption and scattering) and gains due to emissions and scattering according to

$$\frac{\mathrm{d}I(\nu, s)}{\mathrm{d}s} \;=\; -\, \mu(\nu, s) \, I(\nu) \;+\; \rho \, J(\nu, s) \,, \tag{2.10}$$

with

$$\mu \;=\; \mu_a + \mu_s \;=\; \mu_a^{(\mathrm{gas})} + \mu_a^{(\mathrm{aer})} + \mu_s^{(\mathrm{gas})} + \mu_s^{(\mathrm{aer})} \tag{2.11}$$

the extinction coefficient. It is defined as the sum of the absorption (attenuation) coefficient μ_a and scattering coefficients μ_s for gases and particles, respectively. In the source term the emission coefficient per unit mass is designated J and the mass density is given by ρ.

In a homogeneous gaseous medium in local thermodynamic equilibrium (LTE) with $\mu_a^{(\mathrm{aer})} = 0$, the molecular absorption coefficient $\mu_a^{(\mathrm{gas})}$ is equal to the emission coefficient according to Kirchhoff's law (Gustav Robert Kirchhoff 1824–1887). The assumption for LTE is important since it guarantees that the temperature of a gas does not vary with time (steady state) and that the exchange of energy from the radiation field of the source $B(\nu, T)$ is in equilibrium with both, the excitation temperature in the Boltzmann distribution (see Eq. (2.4)) and kinetic temperature of the Maxwell (James Clerk Maxwell 1831–1879) velocity distribution (*López-Puertas and Taylor*, 2001; *Müller-Kirsten*, 2013). The source function without scattering accounts for molecular emission with emissivity $\varepsilon = \mu_a^{(\mathrm{gas})}$ and is then equal to

$$\rho\, J(\nu) \;=\; \varepsilon\, B(\nu, T) \;=\; \mu_a^{(\mathrm{gas})}\, B(\nu, T) \tag{2.12}$$

where B is the Planck function from Eq. (2.7) at temperature T. If scattering is considered $\mu_s \neq 0$, photons deflected out of the probing path contribute to the loss of intensity while the fraction of radiance scattered into the path of propagation (multiple scattering) constitutes a source of radiation and is hence attributed to the source term of Eq. (2.14). The total contribution of radiation incident from all directions into the direction of interest Ω is then given by the source function J_{S}

$$J_{\mathrm{S}} \;=\; \frac{1}{4\pi} \int\limits_0^{4\pi} P(\Omega', \Omega)\, I(\Omega)\, \mathrm{d}\Omega\,. \tag{2.13}$$

The normalized scattering phase function P_{S}

$$P_{\mathrm{S}} \;=\; \frac{1}{4\pi} \int\limits_0^{4\pi} P(\Omega', \Omega)\mathrm{d}\Omega \;=\; 1\,, \tag{2.14}$$

gives the probability of incident radiation being scattered from direction Ω into the direction of propagation Ω'. Substitution of the above assumptions into Eq. (2.10) gives the differential equation

$$\frac{\mathrm{d}I}{\mathrm{d}s} \;=\; -\left(\mu_a^{(\mathrm{gas})} + \mu_s^{(\mathrm{gas})} + \mu_s^{(\mathrm{aer})}\right) I \;+\; \mu_a^{(\mathrm{gas})}\, B \;+\; \left(\mu_s^{(\mathrm{gas})} + \mu_s^{(\mathrm{aer})}\right) J_{\mathrm{S}}\,, \tag{2.15}$$

where $\mu = \mu_a^{(\mathrm{gas})} + \mu_s^{(\mathrm{gas})} + \mu_s^{(\mathrm{aer})}$ represents the extinction so that the ratio

$$\frac{\mu_s^{(\mathrm{gas})} + \mu_s^{(\mathrm{aer})}}{\mu} \;=\; 1 - \frac{\mu_a^{(\mathrm{gas})}}{\mu}\,, \tag{2.16}$$

is referred to as the single scattering albedo (SSA).

In the infrared spectral range scattering often plays a minor role, or can even be neglected (i. e. $\mu_s = 0$ in Eq. (2.16)), and a pure gas atmosphere can be assumed (*Zdunkowski*

et al., 2007). In case of an inhomogeneous medium, i.e. when the physical properties are not constant along the path, the solution for the radiance I received by an instrument at distance s is then given by the Schwarzschild equation

$$I(\nu, s) = I_0(\nu)\mathcal{T}(\nu, s) - \int_s^{s_0} B(\nu, T(s')) \frac{\partial \mathcal{T}(\nu, s')}{\partial s'} \, \mathrm{d}s' \quad \text{(2.17a)}$$

$$= I_0(\nu)\mathcal{T}(\nu, s) + \int_\tau^{\tau_0} B(\nu, T(\tau')) \mathcal{T}(\nu, s') \, \mathrm{d}\tau' \, , \quad \text{(2.17b)}$$

where $\mathcal{T}$ is the monochromatic transmission according to Beer's law (named after August Beer $1825 - 1863$)

$$\mathcal{T}(\nu, s) = \exp\left(-\tau(\nu, s)\right) = \exp\left[-\int_s^{s_0} \mu(\nu, s') \, \mathrm{d}s'\right] \, , \quad \text{(2.18)}$$

with the molecular optical depth τ. The volume absorption coefficient is given by

$$\mu_a^{(\text{gas})} = \sum_m k_m(\nu, p(s), T(s)) \, n_m(s) \, , \quad \text{(2.19)}$$

where k and n_m are the absorption cross section and number density of molecule m.

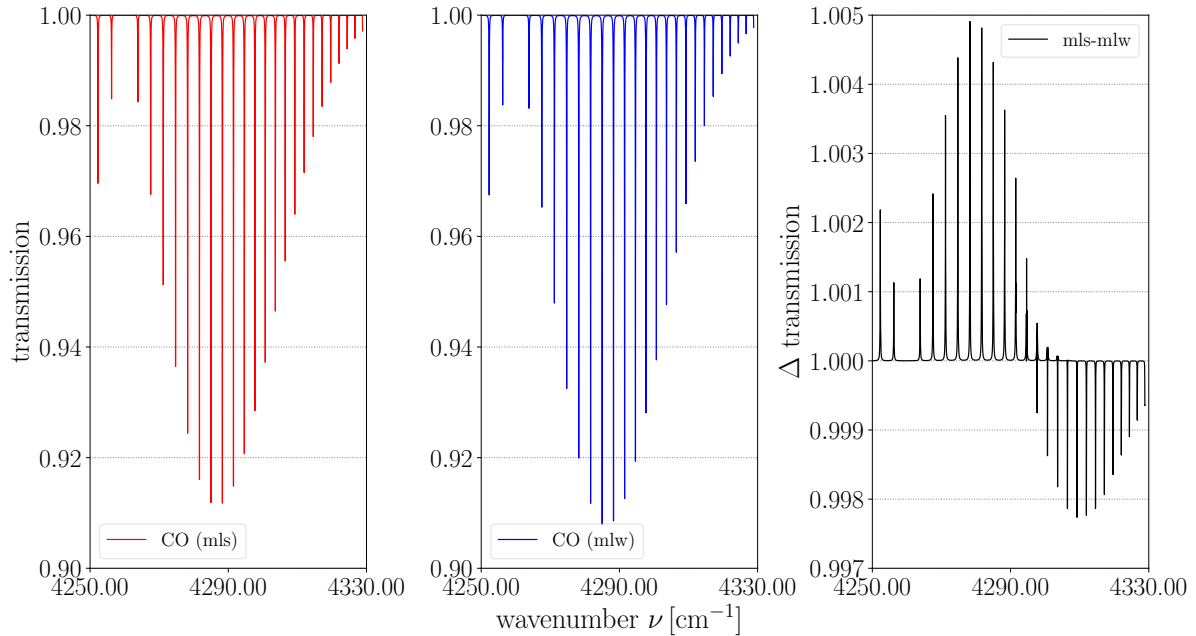**Figure 2.2:** Single path transmission of CO according to Beer's law for US-Standard mid-latitude summer and mid-latitude winter atmospheric conditions.

Particularly in the SWIR, thermal emission from Earth's atmosphere and surface is still weak compared to reflected sunlight (see Fig. 1.1). Hence the integral term in Eq. (2.17) also becomes negligible and the radiative transfer equation is equivalent to

Beer's law for a double path through the atmosphere according to

$$
I(\nu) \quad = \quad \frac{r(\nu)}{\pi} \, \cos(\theta) \, E_{\mathrm{sun}}(\nu) \, \mathcal{T}_{\downarrow}(\nu) \, \mathcal{T}_{\uparrow}(\nu)
$$

$$
= \quad \frac{r(\nu)}{\pi} \, \cos(\theta) \, E_{\mathrm{sun}}(\nu)
$$

$$
\exp\left(- \int\limits_{\mathrm{earth}}^{\mathrm{sun}} \mathrm{d}s' \sum_m n_m(s') \, k_m\Big(\nu, p(s'), T(s')\Big) \right)
$$

$$
\exp\left(- \int\limits_{\mathrm{earth}}^{\mathrm{ToA}} \mathrm{d}s'' \sum_m n_m(s'') \, k_m\Big(\nu, p(s''), T(s'')\Big) \right), \tag{2.20}
$$

where r refers to the surface reflectivity, E_{sun} the irradiance at ToA, $\mathcal{T}_{\downarrow}$ and $\mathcal{T}_{\uparrow}$ (with $\mathcal{T} = \mathcal{T}_{\downarrow}\mathcal{T}_{\uparrow}$) denote transmission between Sun and reflection point (e.g. Earth's surface) and between reflection point and ToA, respectively. For a spherical symmetric, plane-parallel atmosphere, the path length s' is related to altitude z' via $s' = z'/\cos(\pi)$ for an nadir looking observer. Likewise it holds $s'' = z''/\cos(\theta)$ for a solar zenith angle θ.

2.3   Molecular absorption

In the infrared spectral range molecular absorption is due to radiative transitions between rotational and ro–vibrational states of the molecules. The infrared radiation excites vibrations from their ground states ($v' = 0$) to excited states ($v' = 1$ the fundamental, or $v' = 2$ the first overtone) (*Zdunkowski et al.*, 2007). A criterion for infrared absorption is a net change in dipole moment in a molecule as it vibrates or rotates (*Hanel et al.*, 2003; *Goody and Yung*, 1995). So called ro-vibrational spectra occur when rotational energy states are superimposed upon vibrational transitions.

In general, a single spectral line is characterized by its position $\hat{\nu}$, line strength S, and line width Γ where the transition wavenumber $\hat{\nu}$ is determined by the energies E_{i}, E_{f} of the initial and final state, $|i\rangle$, $|f\rangle$

$$
\hat{\nu} \quad = \quad \frac{1}{hc} \, (E_{\mathrm{f}} - E_{\mathrm{i}}) \; . \tag{2.21}
$$

A molecule's absorption cross section k is defined as the product of the line strength S and a normalized line profile function g

$$
k(\nu) \quad = \quad S \; \cdot \; g(\nu - \hat{\nu}, \Gamma) \qquad \text{with} \quad \int\limits_{-\infty}^{+\infty} g \, \mathrm{d}\nu \;\; = \;\; 1 \; , \tag{2.22}
$$

and describes the exchange of energy between the radiation field and energy levels of a molecule. The argument Γ designates the broadening coefficient (details see Sec. 2.4.1.1 and Sec. 2.4.1.2).

2.3.1   Line strength and partition functions

The line strength for a transition is proportional to the Einstein coefficient for absorption B_{if} ($|i\rangle < |f\rangle \longrightarrow E_i < E_f$) (*Rothman et al.*, 1998, A. 1; *Bernath*, 2016, Sec. 1.3)

according to

$$S(T) = \frac{h\,\hat{\nu}\,n_i}{c\,n}\left(1 - \frac{g_i\,n_f}{g_f\,n_i}\right)B_{if}\,, \tag{2.23}$$

where n is the molecule's total number density, n_i and n_f are the number densities of the upper and lower states, and g_i and g_f are statistical weights for the electronic, vibrational, and rotational transitions according to the degeneracy of the lower and upper energy states of the transition, respectively (*Edwards*, 1988). The Einstein coefficients for absorption B_{12} and stimulated emission B_{21} are related to the Einstein A_{21} coefficient by the Einstein relations (*Seager*, 2010, Eq. (8.34)).

In LTE the population partition of states is given by the Boltzmann-distribution according to Eq. (2.4), so that

$$\frac{n_i}{N} = \frac{g_i\,\exp\left(-\frac{E_i}{kT}\right)}{Q(T)}\,, \tag{2.24}$$

with the total internal partition sum Q(T) given by

$$Q(T) = \sum_j g_j \exp\left(-\frac{E_j}{k\,T}\right)\,. \tag{2.25}$$

Accordingly, the ratio in Eq. (2.23) is equal to

$$\frac{g_i\,n_f}{g_f\,n_i} = \exp\left(-\frac{hc\hat{\nu}}{kT}\right)\,. \tag{2.26}$$

In the infrared spectral region the total partition sum is the product of the rotational (*Norton and Rinsland*, 1991) and vibrational (*Eriksson*, 1999) partition functions, $Q = Q_{\text{rot}} \cdot Q_{\text{vib}}$, whose temperature dependence are calculated from

$$Q_{\text{rot}}(T) = Q_{\text{rot}}(T_0)\left(\frac{T}{T_0}\right)^{\beta}\,, \tag{2.27}$$

$$Q_{\text{vib}}(T) = \prod_{j=1}^{q} [1 - \exp(-hc\omega_j/kT)]^{-d_j}\,, \tag{2.28}$$

where β is the temperature coefficient of the rotational partition function, and q the total number of vibrational modes with wavenumbers ω_j and (vibrational) degeneracies d_j.

The square of the temperature dependent matrix element $\boldsymbol{D}$ of the electric dipole moment operator $R_{if} = |\langle f|\boldsymbol{D}|i\rangle|^2$ describes the probability of the transition and is hence related to the Einstein coefficient B_{if} according to

$$R_{if} = \frac{3\,h^2}{8\,\pi^3}\,B_{if}\cdot 10^{36}\,. \tag{2.29}$$

Further factors accounting for the partition function, Boltzmann-distribution, and stimulated emission finally determine the line strength for electric dipole transitions according to

$$S(T) = \frac{8\pi^3}{3hc}\,\frac{g_i I_a}{Q(T)}\,\hat{\nu}\,e^{-E_i/kT}\left[1 - e^{-hc\hat{\nu}/kT}\right]\,R_{if}\cdot 10^{-36} \tag{2.30}$$

where I_a is the relative abundance of the isotope.

In spectroscopic line lists such as HITRAN, GEISA or SEOM–IAS the line strength is provided for a reference temperature (e. g. $T_0 = 296\,\text{K}$). The line intensity at any other temperature T is given by the ratio of line strengths at the two temperatures according to

$$S(T) \;=\; S(T_0) \;\cdot\; \frac{Q(T_0)}{Q(T)} \;\; \frac{\exp\left(-E_i/kT\right)}{\exp\left(-E_i/kT_0\right)} \;\; \frac{1 \;-\; \exp\left(-hc\hat{\nu}/kT\right)}{1 \;-\; \exp\left(-hc\hat{\nu}/kT_0\right)} \;. \tag{2.31}$$
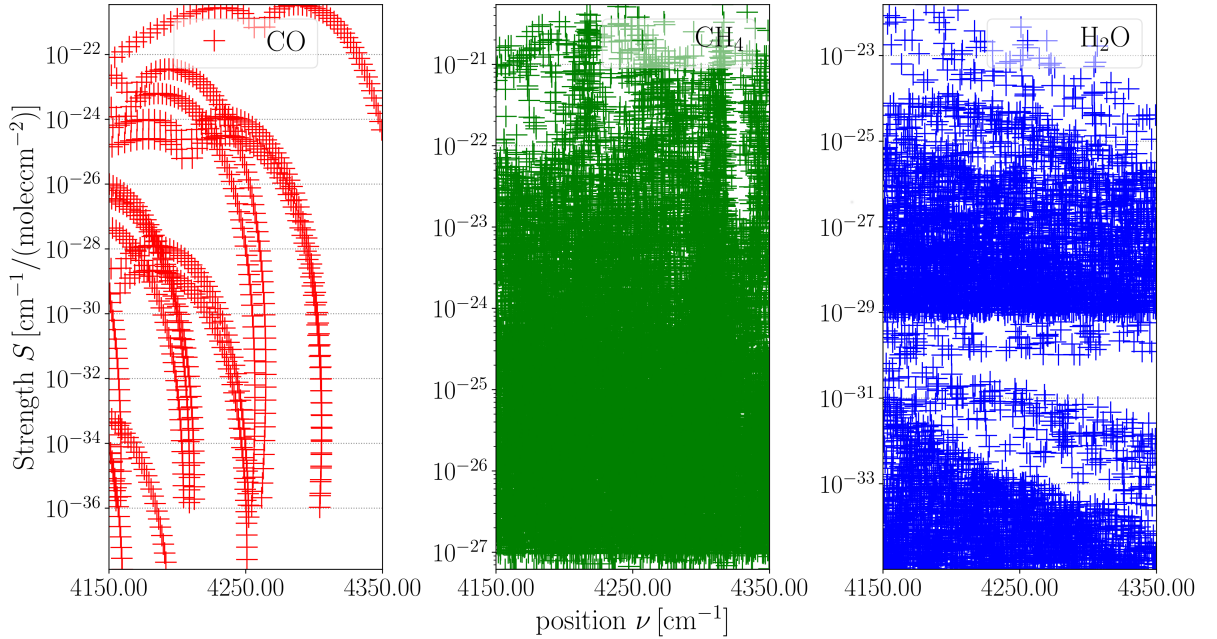
Figure 2.3: Line strengths of CO, CH$_4$, and H$_2$O according to the GEISA 2020 line list.

2.4   Line profile functions

In high-resolution spectroscopy of vibrational-rotational and pure rotational transitions, pressure and temperature greatly affect the shape of an absorption line. The effects are caused by various physical phenomena such as the finite natural lifetime (relaxation rate) of the upper energy state (natural broadening), collisions between molecules (pressure or collision broadening), and the thermal motion of molecules (*Struve*, 1989; *Tennyson*, 2005; *Hartmann et al.*, 2008; *Brooks*, 2014; *Bernath*, 2016). Narrowing mechanisms arise from collision-induced velocity changes and the speed-dependence of the relaxation rates (*Varghese and Hanson*, 1984; *Boone et al.*, 2007). Moreover, effects such line-mixing can cause asymmetrical line shapes (*Rosenkranz*, 1975).

2.4.1   The Voigt profile

Various models have been developed to characterize the physical processes caused by the interaction of the absorbing molecule and the surrounding perturbers. The Voigt profile (*Armstrong*, 1967) is the simplest line shape model accounting for pressure and velocity effects. The pressure induced modifications are described by the Lorentz profile

Table 2.1: Overview of the absorption line parameters (in units cm^{-1}) required for the Voigt profile.

Parameter	Symbol
Line position	$\hat{\nu}_0$
Pressure-induced line shift	Δ_L
Air- and self broadening	Γ_L
Doppler broadening	Γ_G

while effects caused by the velocity distribution of molecules are modeled by the Doppler profile.

New releases of HITRAN and GEISA databases are made available every few years (*Gordon et al.*, 2017; *Jacquinet-Husson et al.*, 2016) and include Voigt parameters for every transition. Line shape parameters in spectroscopic databases (e.g. HITRAN, GEISA, GOSAT-2009/2014 *Nikitin et al.*, 2010, 2015) are provided for a given reference pressure and temperature. In those line lists, the Voigt parameters are tabulated for a reference pressure $p_0 = 1013.25\,\mathrm{hPa}$ and temperature $T_0 = 296\,\mathrm{K}$.

In analogy to the temperature conversion in Eq. (2.31) each line shape parameter needs to be converted to the actual p, T values. Table 2.1 summarizes the converted parameters for the Voigt profile. The Doppler broadening parameter is not included in spectroscopic line lists as it can be computed according to Eq. (2.41).

2.4.1.1 Pressure (collisional) broadening — Lorentz profile

An ideal spectrometer with full spectral resolution would not observe an absorption line that corresponds to an infinitely narrow delta function $g(\nu - \hat{\nu}) \neq \delta(\nu - \hat{\nu})$ due to natural line broadening (*Bernath*, 2016, Sec. 1.3). The effect is caused by the finite lifetime of the excited state τ as the Heisenberg uncertainty principle (Werner Heisenberg 1901–1976) states that uncertainty in lifetime σ_t causes uncertainty in energy σ_E

$$\sigma_E\,\sigma_t \geq \frac{h}{4\pi} \;=\; \frac{\hbar}{2} \tag{2.32}$$

and so in a line's position (see Eq. (2.21)). Although this effect can be neglected in atmospheric spectroscopy, the impact of uncertainty on line broadening is briefly described subsequently since both, the natural upper state lifetime and pressure induced mechanisms have the same effect on the line shape.

The corresponding line profile can be derived from the decay of the excited state according to a damped oscillating dipole moment $\boldsymbol{M}$ at the Bohr (Niels Bohr 1885–1962) angular frequency $\omega_{10} = 2\pi c\hat{\nu}$

$$\boldsymbol{M}(t) \;=\; \boldsymbol{M}_0 \exp\left(-\frac{t}{2\,\tau}\right)\cos(\omega_{10}\,t)\,, \tag{2.33}$$

where τ is the mean natural lifetime in the upper state which is inverse proportional to $\sigma_E/h \propto 1/\tau$ and so inverse proportional to the Einstein A_{10} coefficient for spontaneous emission (*Seager*, 2010).

The collision of molecules causes the upper state lifetimes to be shortened and is referred to as phase-changing or dephasing collisions (*Bernath*, 2016; *Hartmann et al.*, 2008). Given that τ_d is the average time between two collisions, the Fourier transform

of Eq. (2.33) leads to the frequency content of the dipole according to *Bernath* (2016, Eq. (1.71)). With the normalization condition from Eq. (2.22), the Lorentzian line profile normalized to 1 is then given by

$$g_\mathrm{L}(\nu - \hat{\nu}, \Gamma_\mathrm{L}) \;=\; \frac{\Gamma_\mathrm{L}/\pi}{(\nu - \hat{\nu})^2 + \Gamma_\mathrm{L}^2}\,, \tag{2.34}$$

where $\Gamma_\mathrm{L} = 1/(2\pi c \tau_\mathrm{d})$ specifies the half width at half maximum (HWHM) of the profile (also see *Bernath* (2016, Eq. (1.75))). Since the number of collisions depend on the molecular number density and velocities, the Lorentz half width is a function of pressure and temperature (*Zdunkowski et al.*, 2007). The Lorentzian profile therefore constitutes an adequate model to describe pressure broadening in molecular absorption.

The pressure broadening effect is small at low pressure but dominates the line's shape at pressure levels found in the lower atmosphere (*Schreier*, 2011, Fig. 2). More specifically, the Lorentz width Γ_L is inversely proportional to τ_d which is inverse proportional to pressure so that $\Gamma_\mathrm{L} \propto p$ and hence decays approximately exponentially with altitude. Moreover, it decreases with increasing temperature T. Therefore, in a gas mixture with total pressure p and partial pressure p_s of the absorber molecule the total width is given by the sum of a self broadening contribution due to collisions between the absorber molecules and a broadening contribution due to collisions with other molecules,

$$\Gamma_\mathrm{L} \;=\; \left(\gamma_\mathrm{L}^{(0,\mathrm{air})}\,(p - p_\mathrm{s}) \;+\; \gamma_\mathrm{L}^{(0,\mathrm{self})}\,p_\mathrm{s}\right) \;\cdot\; \left(\frac{T_0}{T}\right)^n\,. \tag{2.35}$$

The exponent n quantifying the dependence of temperature is known for many transitions of the most important molecules. In case it is not specified, the kinetic theory of gases (collision of hard spheres) yields the classical value $n = 1/2$. The self broadening coefficient $\gamma_\mathrm{L}^{(\mathrm{self})}$ is known for many (strong) transitions, however, if not specified, the coefficient is set to $\gamma_\mathrm{L}^{(\mathrm{self})} = \gamma_\mathrm{L}^{(\mathrm{air})}$, i.e. the broadening coefficient with respect to air (*Rothman et al.*, 1987), so that

$$\Gamma_\mathrm{L} \;=\; \gamma_\mathrm{L}^{(\mathrm{air})}\,p\,\left(\frac{T_0}{T}\right)^n\,. \tag{2.36}$$

Moreover, the contribution of the self broadening coefficient is weak for molecules with low atmospheric concentrations. In the terrestrial atmosphere, for example, only N_2, O_2 and H_2O in certain climatological regions have a significant share p_s in the total pressure. On other planets such as Mars or Venus with atmospheres mainly composed of CO_2, self broadening becomes crucial for a different set of molecules (*Hanel et al.*, 2003).

Another pressure related effect changes the energy levels of the absorber leading to a shift in the spectral line's position. This pressure-induced line shift Δ_L

$$\Delta_\mathrm{L} \;=\; \hat{\nu} - \hat{\nu}_0\,, \tag{2.37}$$

designates the difference of a position $\hat{\nu}$ from its wavenumber at p_0 and T_0, which is linearly dependent on pressure. Similar to above, the self-pressure induced line shift coefficients for temperature and pressure are largely unknown so that the shift is modeled with respect to air $\delta_p^{(\mathrm{air})}$ and $\delta_T^{(\mathrm{air})}$ according to

$$\Delta_\mathrm{L} \;=\; \delta_p^{(\mathrm{air})}\,(p - p_\mathrm{s}) \;+\; \delta_T^{(\mathrm{air})}\,(T - T_0)\,(p - p_\mathrm{s})\,. \tag{2.38}$$
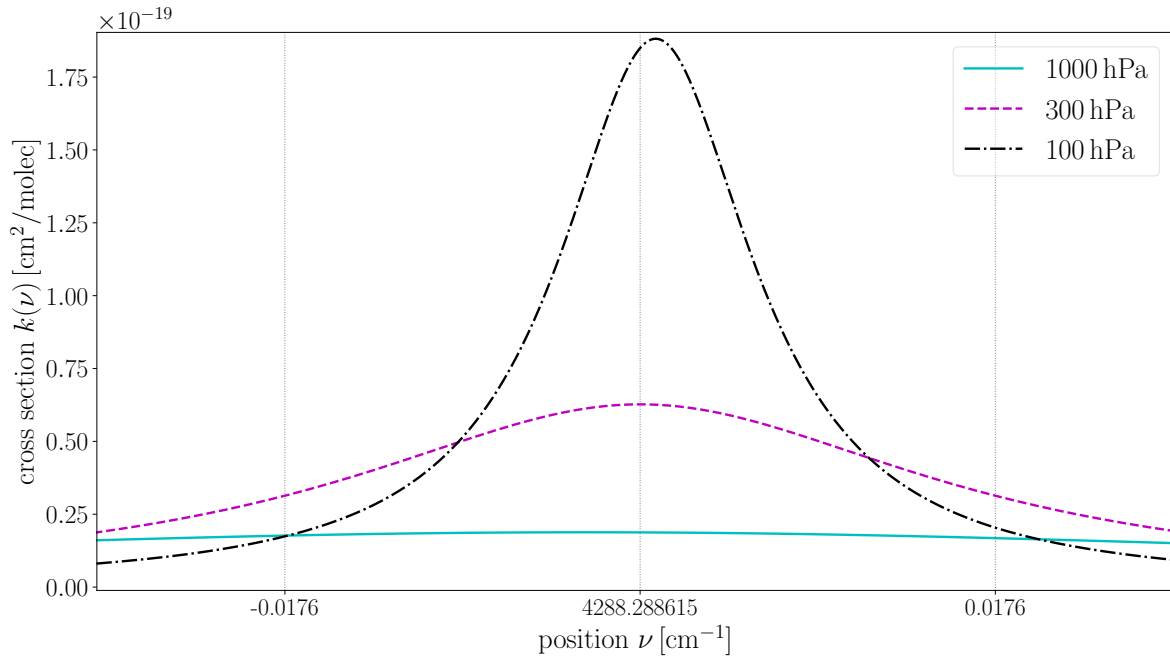
Figure 2.4: Lorentz cross sections of CO according to HITRAN 2016 for three different pressures at 296 K. The vertical grid lines to the left and right of the lines center position designate the half width of the line profile. The lines for 1000 hPa and 100 hPa are pressure shifted since the figure is centered on the 300 hPa line position.

2.4.1.2   Doppler broadening — Doppler profile

In thermal equilibrium the velocity distribution of molecules is given by the Maxwell-Boltzmann distribution

$$q(v) \;=\; \left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} \exp\left(-(v/v_0)\right) , \tag{2.39}$$

where $v_0 = \sqrt{2kT/m}$ is the most probable speed of an individual molecule of mass m (*Bransden et al.*, 2003; *Zdunkowski et al.*, 2007). As a result, the thermal motion leads to the broadening of spectral lines caused by an ensemble of Doppler shifts (Doppler effect). The resulting line shape is described by a Gaussian distribution

$$g_{\mathrm{D}}(\nu - \hat{\nu}, \Gamma_{\mathrm{G}}) \;=\; \frac{1}{\Gamma_{\mathrm{G}}}\left(\frac{\ln 2}{\pi}\right)^{1/2} \cdot \exp\left[-\ln 2\left(\frac{\nu - \hat{\nu}}{\Gamma_{\mathrm{G}}}\right)^2\right] . \tag{2.40}$$

The HWHM is essentially determined by the line position $\hat{\nu}$, the temperature T, and the molecular mass m,

$$\Gamma_{\mathrm{G}} \;=\; \hat{\nu}\left(\frac{2\ln 2\; kT}{mc^2}\right)^{\frac{1}{2}} \;=\; \frac{\hat{\nu}\sqrt{\ln 2}\, v_0}{c} , \tag{2.41}$$

and hence not considered as a free parameter in laboratory spectroscopy. For the Earth's atmosphere one finds

$$\Gamma_{\mathrm{G}} \approx 6\cdot 10^{-8}\,\hat{\nu}\,\sqrt{T} \qquad \text{for} \quad m \approx 36\,\mathrm{amu} ,$$
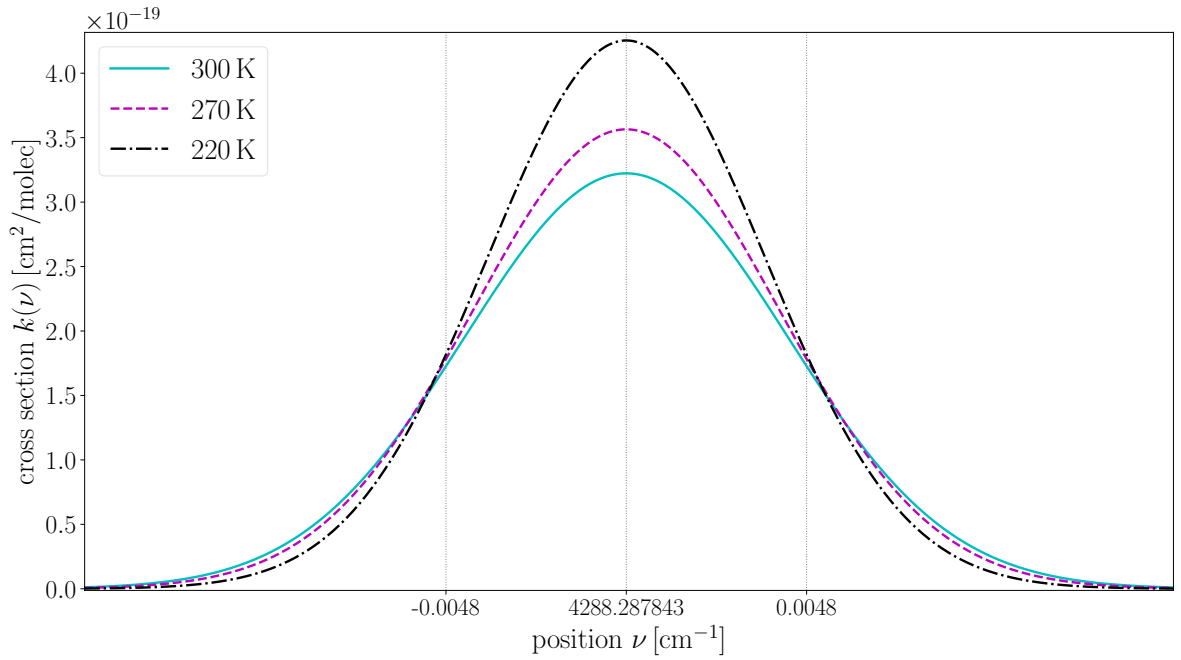
Figure 2.5: Doppler cross sections of CO according to GEISA 2020 for three different temperatures at 500 hPa. The vertical grid lines to the left and right of the lines center position designate the half width of the Gaussian profile.

where 36 amu designates the atomic mass unit for the most important infrared active molecules in the Earth's atmosphere. In contrast to Γ_L, the Doppler width Γ_G does hardly vary with altitude. As a result lines are generally pressure broadened in a high pressure regime such as the troposphere. Beside temperature, the transition to the Doppler regime is depending on the spectral region (*Schreier*, 2011, Fig. 1). Moreover, the uncertainty principle from Eq. (2.32) with $\sigma_E \sigma_t = \sigma_x \sigma_p$, where σ_x and $\sigma_p = \frac{h}{\lambda}$ designate the standard deviation of position and momentum, respectively, confines the wavenumber shift for a single molecule to be $\sigma_x \geq \frac{\lambda}{2\pi}$ in the direction of the observer (*Bernath*, 2016).

2.4.1.3 Combined Pressure and Doppler broadening — Voigt profile

The Voigt line profile combines the effects of both broadening mechanisms by convolution of the Lorentz and the Gauss profiles according to

$$
\begin{aligned}
g_V(\nu - \hat{\nu}, \Gamma_L, \Gamma_G) &\equiv g_L \otimes g_D \\
&= \int_{-\infty}^{\infty} d\tilde{\nu} \, g_L(\tilde{\nu}, \Gamma_L) \cdot g_D(\nu - \hat{\nu} - \tilde{\nu}, \Gamma_G) \, .
\end{aligned}
\tag{2.42}
$$

Every line profile needs to be normalized to unit area according to Eq. (2.22) so that the Voigt profile gives

$$
g_V(\nu - \hat{\nu}, \Gamma_L, \Gamma_G) = \frac{\sqrt{\ln 2/\pi}}{\Gamma_G} K(x, y) \, ,
\tag{2.43}
$$

$$
K(x, y) = \frac{y}{\pi} \int_{-\infty}^{\infty} \frac{e^{-t^2}}{(x - t)^2 + y^2} \, dt \, ,
\tag{2.44}
$$

with $K(x,y)$ representing the Voigt function which is normalized to $\sqrt{\pi}$. The dimension-less variables x, y are defined in terms of the distance from the center position, $\nu - \hat{\nu}$, and the Lorentzian and Gaussian half–widths Γ_L, Γ_G according to

$$x \;=\; \sqrt{\ln 2}\,\frac{\nu - \hat{\nu}}{\Gamma_G} \qquad \text{and} \qquad y \;=\; \sqrt{\ln 2}\,\frac{\Gamma_L}{\Gamma_G}\;. \tag{2.45}$$
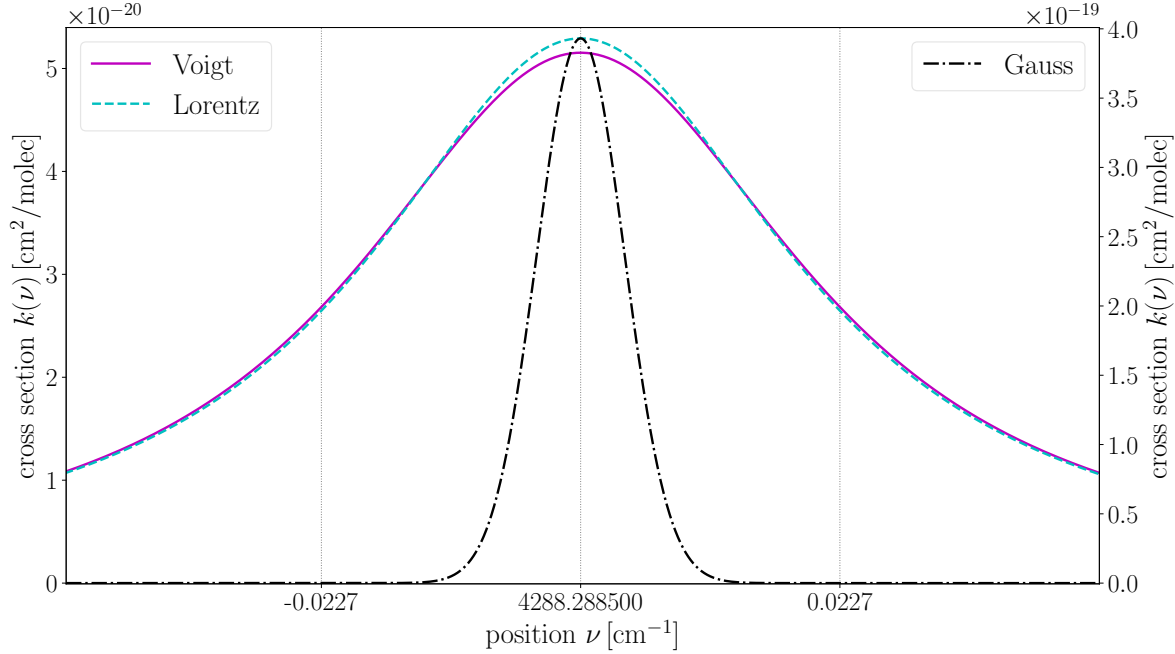
Figure 2.6: Voigt, Lorentz, and Doppler cross sections of CO according to GEISA 2020 at 330 hPa and 242 K. The absorption cross sections of the former two profiles are depicted on the left y-axis while the latter is shown on the right axis. The vertical grid lines to the left and right of the lines center position designate the half width of the Voigt profile which is similar to the half width of the Lorentz profile but much greater than the Gaussian half width describing the Doppler broadening.

The Voigt function represents the real part of the complex function

$$W(z) \;\equiv\; K(x,y) + \mathrm{i}L(x,y) \;=\; \frac{\mathrm{i}}{\pi}\int_{-\infty}^{\infty}\frac{e^{-t^2}}{z - t}\,\mathrm{d}t \qquad \text{with} \qquad z = x + \mathrm{i}y. \tag{2.46}$$

Given that $y > 0$, $W(z)$ is identical to the complex error function (probability function, *Abramowitz and Stegun*, 1964) defined by

$$w(z) \;=\; e^{-z^2}\left(1 \;+\; \frac{2\mathrm{i}}{\sqrt{\pi}}\int_0^z e^{t^2}\,\mathrm{d}t\right) \;=\; e^{-z^2}\left(1 \;-\; \mathrm{erf}(-\mathrm{i}z)\right). \tag{2.47}$$

with the normalization condition for the error function $\mathrm{erf}(\pm\infty) = \pm 1$. It satisfies the differential equation

$$w'(z) \;=\; -2z \cdot w(z) + \frac{2\mathrm{i}}{\sqrt{\pi}} \tag{2.48}$$

and the series and asymptotic expansions (where Γ is the gamma function)

$$w(z) = \sum_{n=0}^{\infty} \frac{(\mathrm{i}z)^n}{\Gamma\left(\frac{n}{2}+1\right)} \tag{2.49}$$

$$w(z) = \frac{\mathrm{i}}{\pi} \sum_{k=0}^{\infty} \frac{\Gamma\left(k+\frac{1}{2}\right)}{z^{2k+1}} \;\; = \;\; \frac{\mathrm{i}}{\pi}\left(\frac{\sqrt{\pi}}{z}+\dots\right). \tag{2.50}$$

According to Eq. (2.47) for vanishing arguments x or y one has

$$
\begin{aligned}
K(0,y) &= e^{y^2}\left(1-\mathrm{erf}(y)\right), \\
K(x,0) &= e^{-x^2}
\end{aligned}
$$

$$K(0,y) \;\; = \;\; e^{y^2}\left(1-\mathrm{erf}(y)\right), \tag{2.51}$$

$$K(x,0) \;\; = \;\; e^{-x^2} \tag{2.52}$$

respectively. Furthermore, truncating the asymptotic expansion of the complex error function from Eq. (2.50) readily shows that the wing of the Voigt profile is approximated by a Lorentzian (*Armstrong*, 1967; *Abramowitz and Stegun*, 1964; *DLMF*). The fact that the function values vary rapidly only near the line center but decrease slowly with increasing distance is exploited by many optimization schemes in state of the art line-by-line models (e. g. *Schreier*, 2011).

The convolution integral of a Lorentzian and Gaussian function can not be evaluated in closed analytical form (*Schreier*, 2016) hence most modern algorithms for the Voigt function employ approximations for the complex error function. Rational approximations, i. e. the quotient of two polynomials of degree M and N according to $R_{M,N} \equiv P_M/Q_N$ (*Ralston and Rabinowitz*, 1978), have been proven to be an efficient and accurate approach to evaluate the complex error function (*Hui et al.*, 1978; *Humlíček*, 1979, 1982; *Weideman*, 1994). Because of the asymptotic behaviour of the complex error function $w \sim 1/z$ (see Eq. (2.50)), the degree of the nominator and denominator polynomials are constrained by $N = M + 1$.

A variety of rational approximations for the complex error function were examined in (*Schreier*, 2011, 2018). *Schreier* (2011) proposes a combination of the Humlíček $R_{1,2}$ rational approximation (*Humlíček*, 1982) and the Weideman approximation (*Weideman*, 1994) according to

$$w(z) \;\; = \;\; \begin{cases} \dfrac{\mathrm{i}z/\sqrt{\pi}}{z^2-\frac{1}{2}} & \text{for } |x|+y \geq 15 \\[2ex] \dfrac{\pi^{-1/2}}{L-\mathrm{i}z} \;\; + \;\; \dfrac{2}{(L-\mathrm{i}z)^2}\sum_{n=0}^{N-1} a_{n+1}\left(\dfrac{L+\mathrm{i}z}{L-\mathrm{i}z}\right)^n & \text{else (with } L = 2^{-1/4}N^{1/2}) \end{cases} \tag{2.53}$$

for line-by-line computations where speed is an issue since the time consuming Weideman approximation is only used near the line center while the significantly faster asymptotic rational approximation is evaluated in the line wings. The real-values polynomial coefficients $a_1, \dots, a_N$ can be computed once and for all by a single fast Fourier transform (*Weideman*, 1994). The combination of both approximations has demonstrated to be efficient and accurate for all x and y given the accuracy required for atmospheric spectroscopy applications. More specifically, $N = 24$ is considered sufficient for Voigt line profile modeling, whereas $N = 32$ should be applied for derivative as well as for line shape computations 'beyond Voigt' (*Schreier*, 2011, Fig. 8).

2.4.2 Beyond Voigt profiles

Many studies found systematic discrepancies between molecular laboratory spectroscopy measurements and modeled spectra using the Voigt profile (*Lisak et al.*, 2004; *Hartmann et al.*, 2008; *Schneider et al.*, 2011; *Kochanov*, 2012; *Ngo et al.*, 2012; *Birk and Wagner*, 2016). This indicates that the classical profile is not a fully accurate representation of the spectral line shape observed in gas mixtures and that the differences are caused by physical processes that are not considered in the Voigt model (*Varghese and Hanson*, 1984; *Tennyson et al.*, 2014). For that reason, several more refined line profiles have been developed (*Berman*, 1972; *Rautian*, 1999).

The complex error function in Eq. (2.47) can be used to compute those more sophisticated profiles that take additional effects such as collisional narrowing, the speed-dependence of collisional broadening (*Boone et al.*, 2007; *Ngo et al.*, 2013, 2014; *Tran et al.*, 2013, 2014), or line coupling (asymmetry) into account (*Tran et al.*, 2011; *Boone et al.*, 2011). The symmetry relations (*Olver et al.*, 2010)

$$w(-z) \;=\; 2e^{-z^2} - w(z) \;=\; \overline{w(z)}, \tag{2.54}$$

and particularly

$$K(-x, y) \;=\; + K(x, y) \tag{2.55}$$
$$L(-x, y) \;=\; - L(x, y), \tag{2.56}$$

make the imaginary part of $w(z)$ a useful tool in modeling those higher order effects.

A report by *Tennyson et al.* (2014) summarizes the results from a IUPAC (International Union of Pure and Applied Chemistry) Task Group on line profiles. It advocates the partially Correlated quadratic-Speed-Dependent Hard-Collision profile (pCqSD-HCP) as the appropriate model for high-resolution spectroscopy. This model is also known as the Hartmann-Tran (HT) profile and accounts for various additional collisional contributions to an isolated absorption line (*Ngo et al.*, 2013, 2014; *Tran et al.*, 2013, 2014). Consequently, the refined model requires more free parameters to characterize the transition.

In the line lists from Sec. 2.4.1, the line parameters were fitted upon the Voigt profile. However, with the recommendation of the IUPAC to standardize on the HT profile line lists such as the HITRAN version of 2016 started to provide beyond Voigt parameters for the Speed-Dependent Voigt, Galatry, and Hartmann-Tran line shapes for some transitions (*Gordon et al.*, 2017). Nonetheless, the classical Voigt parameters are still given for every transition in HITRAN. The latest version of GEISA (2020) also includes updated parameters for the Voigt profile (Fig. 2.7) but does not yet specify any higher-order parameters.

Beside these general spectroscopic databases the SEOM–IAS is an improved line parameter database of H_2O, CH_4, and CO compiled within the framework of an ESA project according to the needs of the TROPOMI instrument (*Birk et al.*, 2017a,b). The absorption lines were fitted using the HT profile without considering partial correlation η (*Loos et al.*, 2017). In order to account for line-mixing, the profile was extended using the approximation by Rosenkranz (Y; *Rosenkranz*, 1975; *Boone et al.*, 2011) and Smith (*Smith*, 1981). Table 2.2 summarizes the set of spectroscopic line parameters required to describe molecular absorption according to the HT model.

While Γ_L and Δ_0 (see Table 2.1) describe the collisional width and shift for the averaged speeds of the molecules, Γ_2 and Δ_2 specify their dependence on the absorbing
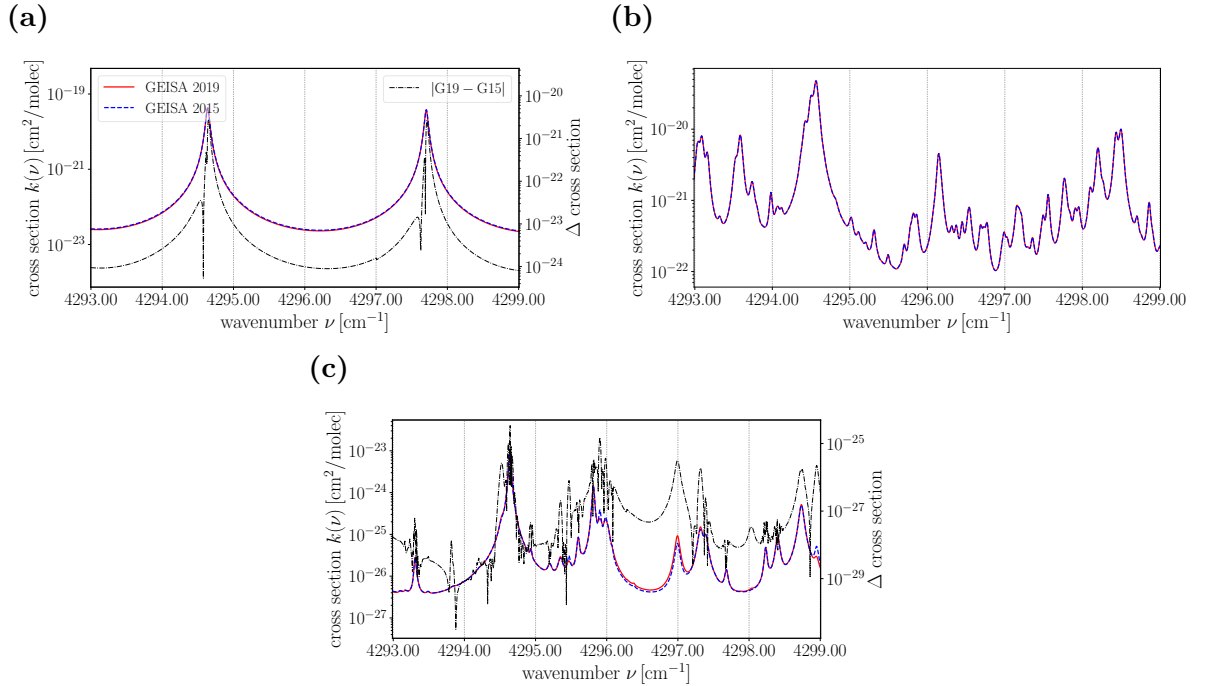
(a)

(b)

(c)

Figure 2.7: **(a)** Cross sections for CO from GEISA 2020 compared to GEISA 2015. **(b)** Comparison for H_2O. No difference was observed in the examined interval for CH_4.

Table 2.2: Overview of the various 'beyond Voigt' line parameters.

Parameter	Symbol	Units
Speed-dependence of air broadening	Γ_2	cm^{-1}
Speed-dependence of pressure-induced line shift	Δ_2	cm^{-1}
Frequency of velocity-changing collisions (Dicke effect)	Γ_N	cm^{-1}
Correlation of velocity and rotational state changes	η	1
Coupling coefficient for Rosenkranz line-mixing	Y	1

molecule's speed (*Varghese and Hanson*, 1984; *Rohart et al.*, 2008; *Pine*, 1994). In the HT profile as well as the Speed-Dependent Voigt profile described by *Boone et al.* (2007) a quadratic-speed-dependence of these parameters is assumed.

The parameter $\Gamma_N = \nu_{vc}$ (see *Tennyson et al.*, 2014) describes modifications of the spectral line shape caused by collision-induced velocity changes that influence the Doppler broadening (*Hartmann et al.*, 2008). The effect is known as collisional (Dicke) narrowing (*Dicke*, 1953).

The parameter η quantifies the partial correlation between velocity and rotational state changes since velocity-changing and phase-changing collisions are correlated. However, η has so far only been fitted for very few transitions and hence the parameter was not available for the studies related to this thesis (Pub. II; Pub. III; *Schreier and Hochstaffl*, 2021).

Finally, the coupling coefficient Y represents the Rosenkranz parameter which approximates line-mixing in the first order. Line-mixing arises for lines which are close together and is dependent upon molecular collisions (*Lévy et al.*, 1992), hence, also modifies the collisional broadening (Lorentz) portion of the complex error function (*Letchworth and Benner*, 2007).

Despite their importance for the applicability of the HT profile the temperature dependence of the new parameters is still an open question (*Ngo et al.*, 2013, 2014; *Tennyson et al.*, 2014). To mitigate the problem, a separate reference temperature is provided for each temperature interval in HITRAN 2016 (*Gordon et al.*, 2017). Within an interval, pressure dependence follows a linear relation and temperature dependence is in general approximated with the conventional power law (*Birk et al.*, 2017b). In case of the SEOM–IAS line list the pressure and temperature dependence of the beyond Voigt parameters is modelled according to

$$\Gamma_2 = \gamma_2^{(\text{air})} \, p \, \left(\frac{T_0}{T}\right)^n \tag{2.57}$$

$$\Delta_2 = \delta_2^{(\text{air})} \, p \tag{2.58}$$

$$\Gamma_\text{N} = \gamma_\text{N}^{(\text{air})} \, p \, \frac{T_0}{T} \tag{2.59}$$

$$Y = Y_p^{(\text{air})} \, p \, \left(1 + \frac{Y_T^{(\text{air})}}{60} \, (T - T_0)\right) , \tag{2.60}$$

hence the zeroth and second order broadening parameters are assumed to have the same p, T dependence, i.e. $\Gamma_{L,2} \propto p T^{-n}$, whereas the narrowing parameter is inversely proportional to temperature, $\Gamma_\text{N} \propto p/T$. As pointed out by *Tennyson et al.* (2014) collisional parameters for the more complex parameterization are, strictly speaking, no longer linear combinations of the various molecule-perturber pairs in the gas mixture (*Ngo et al.*, 2014), but rather separate profiles for perturbations by e.g., O_2 or N_2 need to be added. However, so far, databases such as SEOM–IAS treat collisional parameters as linear combinations. Since partial correlation between speed-dependence and velocity changes is not considered in SEOM–IAS, i.e., it is zero for all lines, the Speed-Dependent Rautian with line-mixing (SDRM) profile is sufficient for this database (Pub. II; *Schreier and Hochstaffl*, 2021). Table 2.3 provides an overview of the various higher-order profiles and their set of parameters, respectively (also see *Tennyson et al.*, 2014). The HT profile is introduced first as the simpler models are its limiting cases.

2.4.2.1 The Hartmann-Tran profile

The HT profile refines the description for a single isolated absorption line by accounting for collisional effects not considered in the classical Voigt model. It is based on seven

Table 2.3: Line profiles with corresponding line parameters.

Acronym	Profile	Parameters
-	Voigt	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}$
VM	Voigt with line-mixing	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}, Y$
-	Rautian	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}, \Gamma_\text{N}$
SDV	Speed-Dependent Voigt	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}, \Gamma_2, \Delta_2$
SDVM	Speed-Dependent Voigt with line-mixing	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}, \Gamma_2, \Delta_2, Y$
SDR	Speed-Dependent Rautian	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}, \Gamma_2, \Delta_2, \Gamma_\text{N}$
SDRM	Speed-Dependent Rautian with line-mixing	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}, \Gamma_2, \Delta_2, \Gamma_\text{N}, Y$
HT	Hartmann-Tran (pCqSD-HCP)	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}, \Gamma_2, \Delta_2, \Gamma_\text{N}, \eta$
HTM	Hartmann-Tran with line-mixing	$\Gamma_\text{G}, \Gamma_\text{L}, \Delta_\text{L}, \Gamma_2, \Delta_2, \Gamma_\text{N}, \eta, Y$

parameters, of which six are collisional parameters (see Table 2.3) that determine the shape for each spectral line and perturber. According to *Ngo et al.* (2013); *Tennyson et al.* (2014) the normalized (to unit area) HT profile is given by

$$
\begin{aligned}
g_{\mathrm{HT}}(\nu - \hat{\nu}, \Gamma_{\mathrm{L}}, \Gamma_2, \Delta_L, \Delta_2, \Gamma_{\mathrm{G}}, \Gamma_{\mathrm{N}}, \eta) \;\; &= \\
&= \;\; \frac{1}{\pi} \operatorname{Re}\left\{ \frac{A(\nu)}{1 - (\Gamma_{\mathrm{N}} - \eta\,(C_{\mathrm{L}} - 3C_2/2))\,A(\nu) + (\eta\,C_2/v_0^2)\,B(\nu)} \right\}, \quad (2.61)
\end{aligned}
$$

where A and B are combinations of the complex error functions from Eq. (2.47) according to

$$
\begin{aligned}
A(\nu) \;\; &= \;\; \frac{\sqrt{\ln 2/\pi}}{\Gamma_{\mathrm{G}}} \left[w(\mathrm{i}z_-) - w(\mathrm{i}z_+) \right], & (2.62) \\
B(\nu) \;\; &= \;\; \frac{v_0^2}{(1-\eta)\,C_2} \left(\frac{\sqrt{\pi}}{2\sqrt{Y}}\,(1 - z_-^2)\,w(\mathrm{i}z_-) - \frac{\sqrt{\pi}}{2\sqrt{Y}}\,(1 - z_+^2)\,w(\mathrm{i}z_+) - 1 \right), & (2.63)
\end{aligned}
$$

v_0 designates the most probable speed of the molecules (see Eq. (2.39) and (2.41)) and $\hat{\nu}_0$ specifies the line's position (see Eq. (2.37)). The complex arguments $\mathrm{i}z_\pm$ are given by

$$
z_\pm \;\; = \;\; \sqrt{X + Y} \; \pm \; \sqrt{Y}, \quad (2.64)
$$

where

$$
\begin{aligned}
X \;\; &\equiv \;\; \frac{\Gamma_{\mathrm{N}} + \mathrm{i}(\nu - \hat{\nu}_0)}{(1-\eta)\,C_2} + \frac{C_{\mathrm{L}}}{C_2} - \frac{3}{2}, & (2.65) \\
Y \;\; &\equiv \;\; \left(\frac{\hat{\nu}_0\,v_0}{2c\,(1-\eta)\,C_2} \right)^2 & (2.66)
\end{aligned}
$$

and

$$
\begin{aligned}
C_{\mathrm{L}} \;\; &= \;\; \Gamma_{\mathrm{L}} + \mathrm{i}\Delta_{\mathrm{L}} & (2.67) \\
C_2 \;\; &= \;\; \Gamma_2 + \mathrm{i}\Delta_2 & (2.68)
\end{aligned}
$$

according to (*Tennyson et al.*, 2014; *Tran et al.*, 2013). The quadratic-speed-dependence in Eq. (2.68) is modeled as a modification of the collisional broadening (Lorentz) portion of the complex error function while collision-induced velocity changes reduce the Doppler broadening (*Boone et al.*, 2007; *Varghese and Hanson*, 1984).

The HT profile can be calculated readily from the complex error function as it involves the difference of two complex error functions (*Schreier*, 2017; *Schreier and Hochstaffl*, 2021). In view of the calculation of differences, highly accurate numerical algorithms are required for the computation of the convolution integral defining the complex error functions. For this thesis the rational approximation according to Eq. (2.53) with $N = 32$ was hence used to compute the integral for all combinations of x and y (*Schreier*, 2017). Beside the model's high accuracy for various combinations of absorbers and perturbers (*Ngo et al.*, 2014) the computational time is in the order of simpler models ($\approx$ a factor 2–3 slower) and in addition beyond Voigt profiles only need to be calculated for the (strongest) lines that hold additional parameters (see Pub. II, Table 1). These are crucial considerations for the performance of line-by-line radiative transfer codes (*Edwards*, 1988).

The computational aspects of the complex square roots in the calculation of the complex argument $iz_\pm$ are discussed in (*Schreier and Hochstaffl*, 2021, Sec. 3.1 and 3.2). Particularly for small y, standard floating point precision fails to evaluate the difference in Eq. (2.64). A reliable way to avoid the subtraction of two similar numbers $\sqrt{X+Y}$ and $\sqrt{Y}$ is

$$z_- \ = \ \sqrt{X+Y} \ - \ \sqrt{Y} \ = \ \frac{X}{\sqrt{X+Y} \ + \ \sqrt{Y}} \ = \ \frac{X}{z_+}. \tag{2.69}$$

The HT model is compatible with current implementations of line-mixing (see Sec. 2.4.2.5 and *Rosenkranz*, 1975; *Boone et al.*, 2011). Moreover, accordingly to the limiting cases of the Lorentzian and Doppler shapes in case of the Voigt profile, simpler (lower-order) models emerge from limiting cases of the HT profile where not all the parameters have been determined.
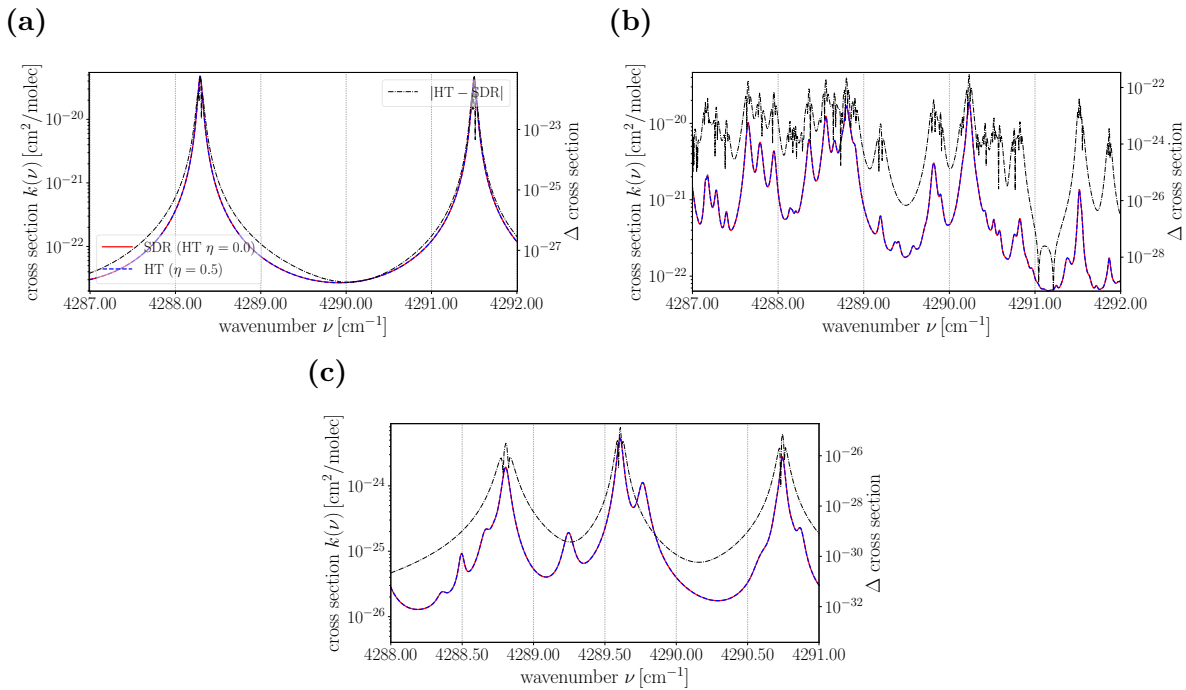
(a)

(b)

(c)

Figure 2.8: Absorption cross sections of **(a)** CO, **(b)** CH_4, and **(c)** H_2O for the HT profile and SEOM–IAS line data, computed for $356\,\text{hPa}$ and $236\,\text{K}$ with partial correlation set to $\eta = 0.5$ and $\eta = 0.0$, respectively.

2.4.2.2 Rautian profile

The effect of line narrowing described by the Dicke parameter can be modeled using the hard- or soft-collision models. The hard-collision model, referred to as the Rautian profile (*Rautian*, 1999), is most appropriate when the perturbers are considerably more massive than the absorbing molecule since after a single collision the velocity is completely uncorrelated to the velocity prior to it. In the soft-collision model, known as the Galatry profile, a single collision is highly correlated with the velocity prior to the collision and many collisions are required before the velocity becomes randomly distributed. Both models introduce one extra parameter to quantify the frequency of the velocity-changing

collisions. The Rautian profile for the hard-collision model with the narrowing parameter Γ_N is given by

$$g_{\mathrm{RTN}}(\nu - \hat{\nu}, \Gamma_L, \Delta_L, \Gamma_G, \Gamma_N) \;=\; \frac{\sqrt{\ln(2)/\pi}}{\Gamma_G} \mathrm{Re}\left\{ \frac{w(x, y + \zeta)}{1 - \sqrt{\pi}\,\zeta\,w(x, y + \mathrm{i}\zeta)} \right\}. \qquad (2.70)$$

where

$$\zeta \;=\; \frac{\sqrt{\ln 2}\,\Gamma_N}{\Gamma_G} \qquad (2.71)$$

is the frequency of collision changes (Dicke narrowing) normalized by the Gaussian width. According to *Dicke* (1953) collision-induced velocity changes narrow the spectral line shape and it becomes noticeable for small values of the parameter y, i.e. when the line profile is dominated by Doppler broadening.
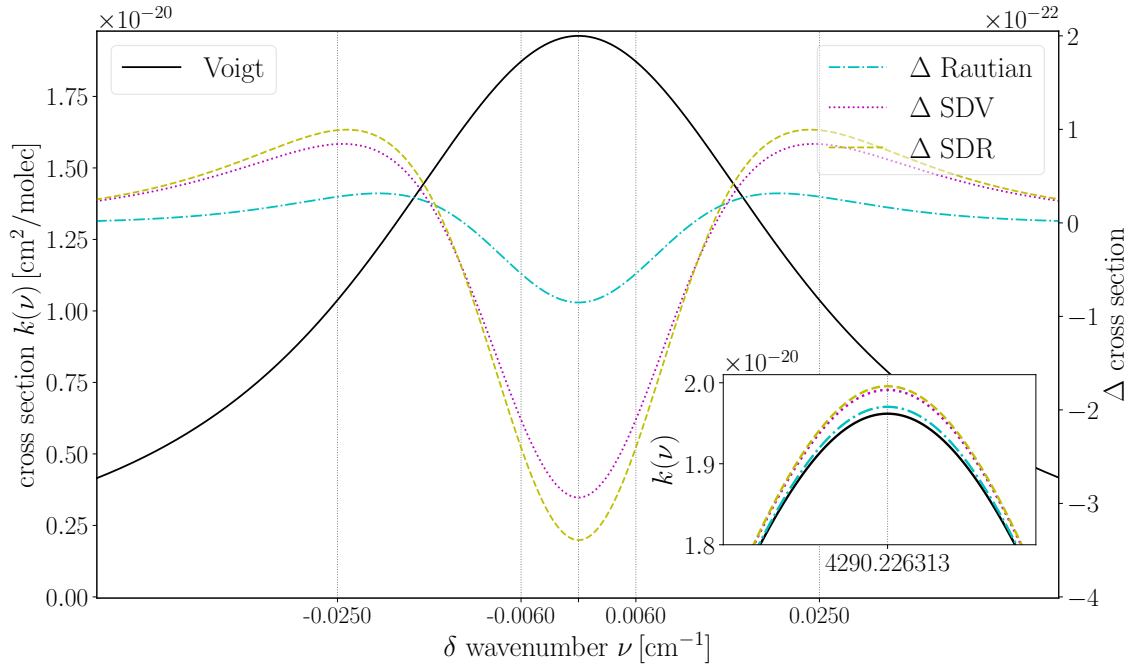

Figure 2.9: Absorption cross sections of CH_4 for various line profiles beyond Voigt at 330 hPa and 243 K computed with SEOM–IAS line data. The difference is calculated with respect to the Voigt profile. The inset shows the center region of the individual line shapes.

2.4.2.3   Speed-Dependent Voigt profile

For vanishing $\eta = 0$ and $\Gamma_N \propto \zeta = 0$ the limit of the HT profile is the Speed-Dependent Voigt (SDV) profile (*Boone et al.*, 2007; *Schreier*, 2017) given by

$$g_{\mathrm{SDV}}(\nu - \hat{\nu}, \Gamma_L, \Gamma_2, \Delta_L, \Delta_2, \Gamma_G) \;=\; \frac{\sqrt{\ln 2/\pi}}{\Gamma_G} \mathrm{Re}\left\{ w(\mathrm{i}z_-) - w(\mathrm{i}z_+) \right\}. \qquad (2.72)$$

2.4.2.4 Speed-Dependent Rautian profile

With vanishing correlation ($\eta = 0$) the HT model reduces to the Speed-Dependent Rautian (SDR) profile (*Varghese and Hanson*, 1984; *Tennyson et al.*, 2014; *Schreier and Hochstaffl*, 2021). The SDR function is essentially the quotient of the difference of two complex error functions. The profile is given by

$$g_{\text{SDR}}(\nu - \hat{\nu}, \Gamma_{\text{L}}, \Gamma_2, \Delta_{\text{L}}, \Delta_2, \Gamma_{\text{G}}, \Gamma_{\text{N}}) =$$

$$= \frac{\sqrt{\ln 2/\pi}}{\Gamma_{\text{G}}} \text{Re} \left\{ \frac{w(\text{i}z_-) - w(\text{i}z_+)}{1 - \sqrt{\pi}\,\zeta\,(w(\text{i}z_-) - w(\text{i}z_+))} \right\}, \quad (2.73)$$
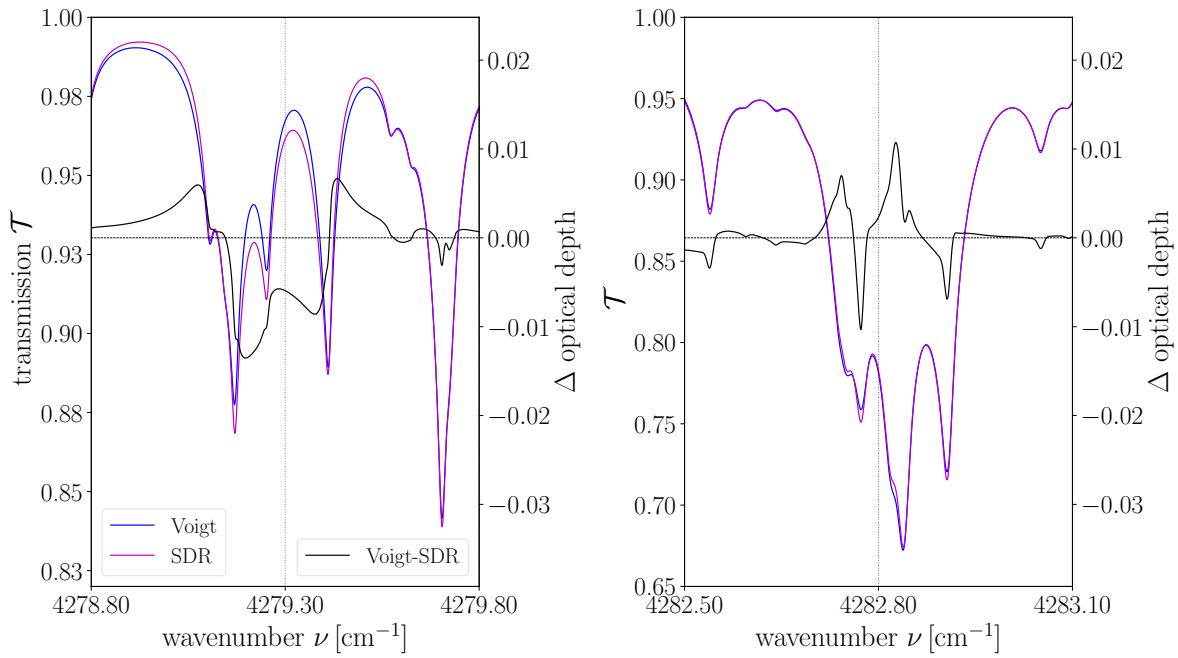

Figure 2.10: Single path nadir CH_4 transmission for SEOM–IAS based Voigt and SDR cross sections and two spectral intervals. US-Standard atmospheric conditions apply.

2.4.2.5 Rosenkranz Line-Mixing

The line profiles considered above have assumed that the spectral line being calculated is sufficiently isolated. If two strong lines have transitions $|f_1\rangle \leftarrow |i_1\rangle$ and $|f_2\rangle \leftarrow |i_2\rangle$ with transition wavenumbers $\hat{\nu}_1$ and $\hat{\nu}_2$ within e.g., their half width Γ, the molecule has several possible paths to make the transition from state $|i_1\rangle$ or $|i_2\rangle$ to $|f_1\rangle$ or $|f_2\rangle$ (*Lévy et al.*, 1992; *Loos et al.*, 2015; *Hartmann et al.*, 2009; *Tran et al.*, 2010). Line-mixing can only occur for lines of the same molecule (and isotopologue) and the effect becomes more important with increasing pressure (i.e. for large values of the parameter y *Loos et al.*, 2017). Figure 2.11 shows the effect of line-mixing for 330 hPa. The line shapes drop off more gradually on the side where another strong line is present and more quickly on the side where there is no coupling.

The effect of first order line-mixing on the line profile can be modeled by the Rosenkranz approximation (*Rosenkranz*, 1975; *Boone et al.*, 2011; *Strow et al.*, 1994; *Pine and Gabard*,

2000). In case of the Voigt profile the formalism utilizes the imaginary component $L(x, y)$ according to

$$
\begin{aligned}
g_{\mathrm{VM}}(\nu - \hat{\nu}, \Gamma_{\mathrm{L}}, \Gamma_{\mathrm{G}}, Y) &= \frac{\sqrt{\ln(2)/\pi}}{\Gamma_{\mathrm{G}}} \mathrm{Re}\left((1 - \mathrm{i}\, Y)\, w(z)\right) \\
&= \frac{\sqrt{\ln(2)/\pi}}{\Gamma_{\mathrm{G}}} \left(K(x, y) + Y\, L(x, y)\right) .
\end{aligned}
\tag{2.74}
$$

For higher order line models such as the HT profile from Eq. (2.61) Rosenkranz line-mixing is considered by introducing Y according to

$$
\begin{aligned}
g_{\mathrm{HTM}}(\nu - \hat{\nu}, \Gamma_{\mathrm{L}}, \Gamma_2, \Delta_{\mathrm{L}}, \Delta_2, \Gamma_{\mathrm{G}}, \Gamma_{\mathrm{N}}, \eta, Y) &= \\
&= \frac{1}{\pi} \mathrm{Re}\left\{ \frac{(1 - \mathrm{i}\, Y)\, A(\nu)}{1 - (\Gamma_{\mathrm{N}} - \eta(C_{\mathrm{L}} - 3C_2/2))\, A(\nu) + (\eta C_2/v_0^2)\, B(\nu)} \right\} .
\end{aligned}
\tag{2.75}
$$

The Hartmann-Tran with line-mixing (HTM) profile reduces to the HT profile for $Y = 0$ and a comparison of both models is shown in Fig. 2.12. Other profiles such as the SDRM or Speed-Dependent Voigt with line-mixing (SDVM) are treated accordingly (*Boone et al.*, 2011, Sec. 4).
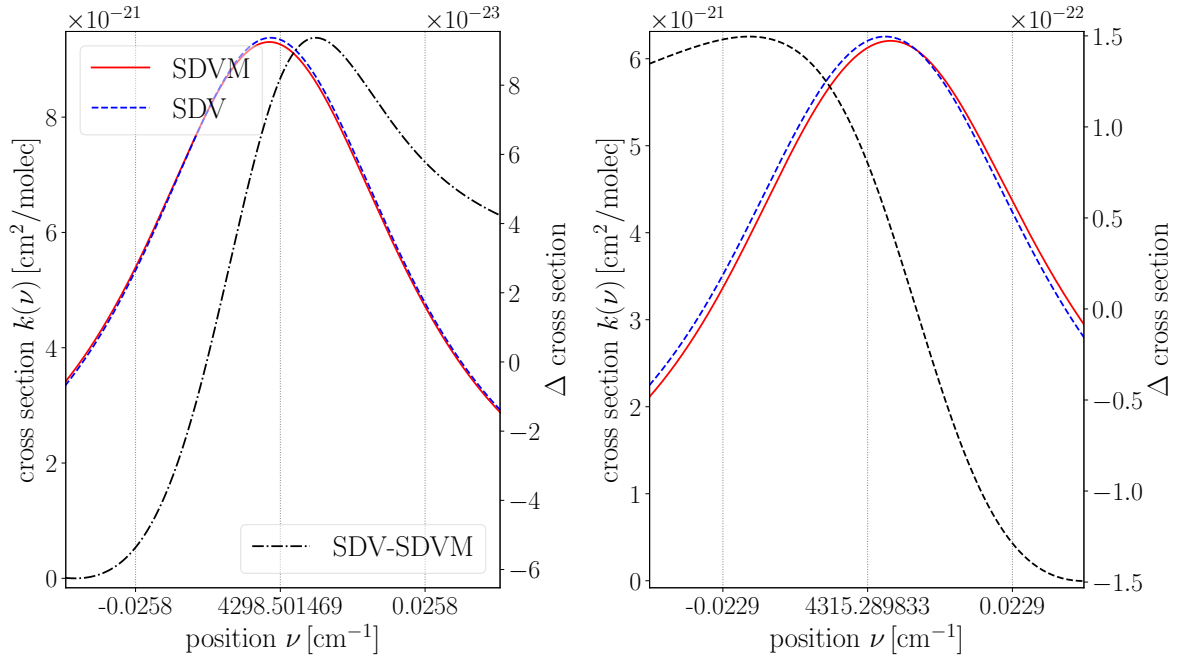

Figure 2.11: The effect of line-mixing on molecular cross sections for the SEOM–IAS based Speed-Dependent Voigt profile at 330 hPa and 243 K. While in the left figure the line strengths of the two neighboring lines differ by a factor ≈ 2 the line strengths in the figure on the right are almost equal causing the mixing effect to be more pronounced.
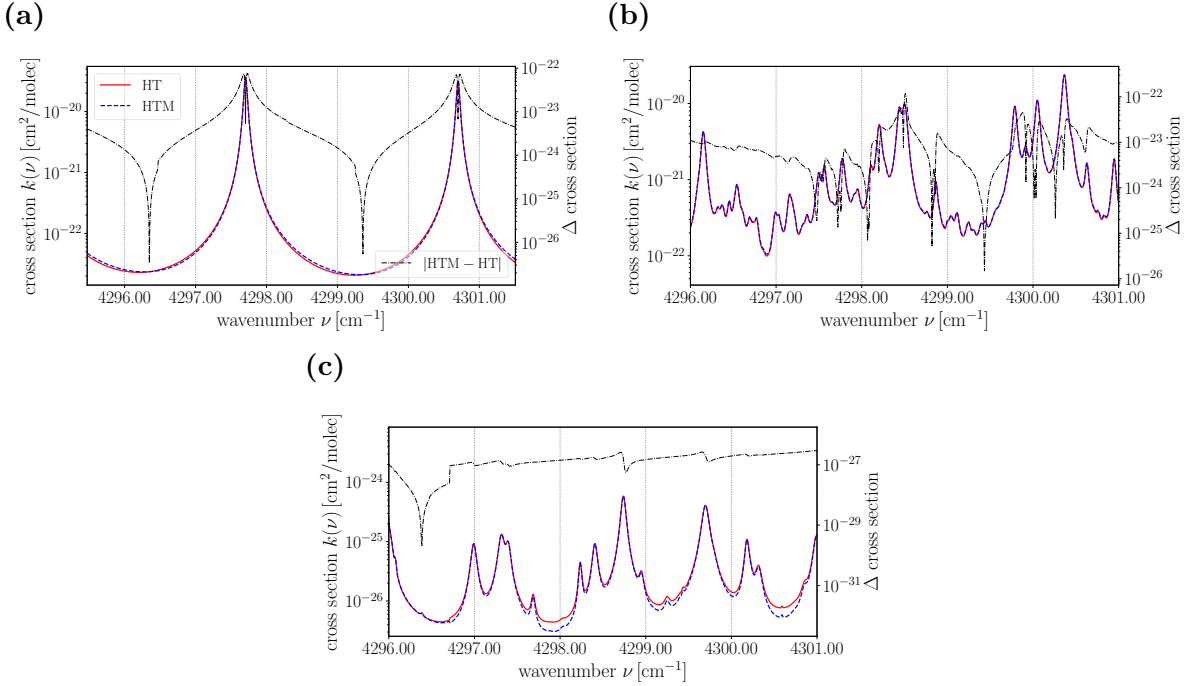
(a)

(b)

(c)

Figure 2.12: Line-mixing effect demonstrated for the HT absorption cross sections of **(a)** CO, **(b)** CH_4 and **(c)** H_2O at 356 hPa and 236 K, respectively, with $\eta = 0.5$ and $Y = 10^{-2}$ (for the HTM profile).

2.5 Extinction by particles and molecular scattering

In addition to molecular absorption scattering by molecules also contributes to attenuation as photons scattered out of the probing beam do not reach the detector. Furthermore, particles (aerosols) attenuate by scattering and absorption which is known as extinction (*Burrows et al.*, 2011, Chap. 6).

The loss term on the right hand side of Eq. (2.16) indicated that both, absorption and scattering contribute to the loss of radiative intensity. In contrast to molecular absorption where transitions from lower to higher energy states determine absorption, scattering by molecules (Rayleigh scattering) and aerosols redirect the incident energy according to their single scattering phase function (see Eq. (2.14) and *Seager*, 2010). Aerosols can also absorb some of the incident intensity as a gray body with emissivity $\varepsilon = \mu_a^{(\mathrm{aer})}$ according to Kirchhoff's law (*Stamnes et al.*, 2017). Note that this aspect was not considered in the radiative transfer equation in Eq. (2.16) since it would violate LTE which, strictly speaking, compromise predictions from Eq. (2.17) and Eq. (2.18), both requiring a thermodynamic temperature according to LTE conditions (*Fischer and Hase*, 2015).

The Ångstrom exponent β (Anders Jonas Ångstrom 1814–1874) is a measure for the aerosol optical depth τ_{aer} at wavelength λ according to

$$\frac{\tau_{\mathrm{aer}}(\lambda)}{\tau_{\mathrm{aer}}(\lambda_0)} = \left(\frac{\lambda}{\lambda_0}\right)^{-\beta}, \qquad (2.76)$$

where $\tau_{\mathrm{aer}}(\lambda_0)$ is the optical thickness at a reference wavelength (e.g. $\lambda_0 = 1~\mu$m). The exponent β basically depending on the scatterer's size distribution $n(r)$ which in turn is

primarily depending on the aerosol's origin (e.g. rural-, urban-, polar-, marine- aerosols). For clear sky conditions with weak scattering by haze or dust a β could be set to 1.3 (*Liou*, 2002; *Yan et al.*, 2015) while for hazy conditions it is assumed to increase to ≈ 2.5 (*Wunderlich et al.*, 2021).

As indicated in Eq. (2.19) the exponential attenuation of intensity when passing through a medium is composed of several components according to their optical depths. In order to account for the additional attenuation caused by Rayleigh and aerosol scattering, the exponent can be complemented accordingly

$$\tau_{\text{tot}}(\nu, s) \; = \; \tau(\nu, s) \; + \; \tau_{\text{ray}}(\nu, s) \; + \; \tau_{\text{aer}}(\nu, s) \; , \tag{2.77}$$

where τ_{ray} and τ_{aer} are the Rayleigh and aerosol optical depth at a particular wavenumber ν, respectively, and τ the molecular optical depth from Eq. (2.18) (*Bodhaine et al.*, 1999).

2.5.1   Aerosol extinction

In analogy to molecular absorption coefficient from Eq. (2.19) a aerosol cross section k_{aer} can be formulated by using Eq. (2.76). As mentioned in the previous section an aerosol optical depth that is appropriate for clear conditions is proportional to $\lambda^{-1.3}$ so that

$$k_{\text{aer}}(\lambda) \; = \; k_{\text{aer}}^1/\lambda^{1.3} \qquad \text{with} \quad k_{\text{aer}}(\lambda_0 = 1\,\mu\text{m}) \equiv \text{k}_{\text{aer}}^1 = 1.4 \cdot 10^{-27} \tag{2.78}$$
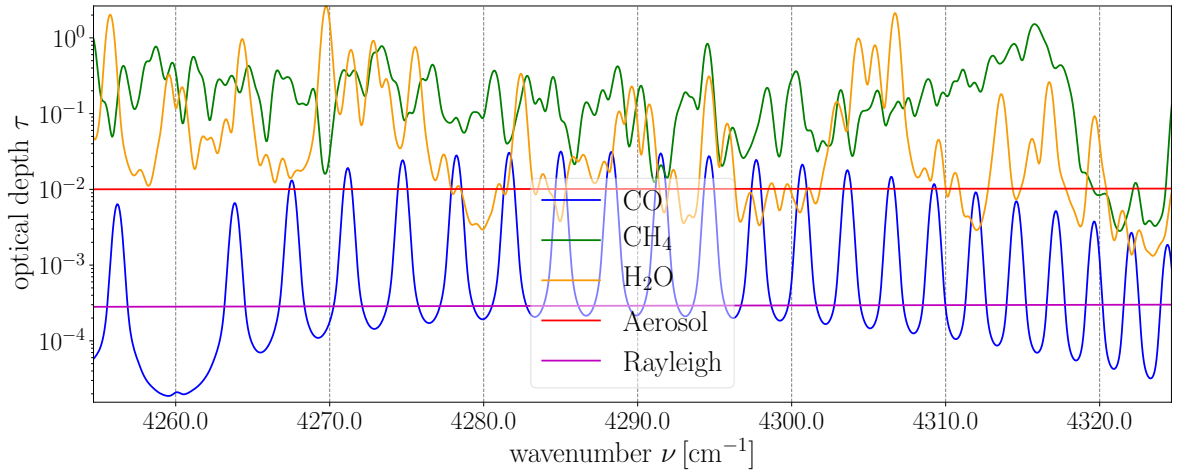

Figure 2.13: Molecular and aerosol optical depths for a double path through the atmosphere. The aerosol optical thickness (red) for $\beta = 1.3$ is of the same magnitude as the CO optical depth. However, it does not show any high frequent spectral variations. The Rayleigh optical depth (magenta) is almost two orders of magnitude smaller.

The aerosol optical depth for a double path in nadir viewing geometry is then given by (*Liou*, 2002; *Kaltenegger and Traub*, 2009)

$$
\begin{aligned}
\tau_{\text{aer}}(\nu) \; &= \; \int\limits_{\text{earth}}^{\text{sun}} k_{\text{aer}}(\nu)\, n_{\text{air}}(z')/\cos(\theta)\; \text{d}s' \; + \; \int\limits_{\text{earth}}^{\text{ToA}} k_{\text{aer}}(\nu)\, n_{\text{air}}(z'')/\cos(\pi)\; \text{d}s' \\
&= \; \left(N_{\text{air}}^{\downarrow} \; + \; N_{\text{air}}^{\uparrow}\right)\, k_{\text{aer}}^1\, \left(10^4/\nu\right)^{-1.3} \\
&= \; \left(N_{\text{air}}^{\downarrow} \; + \; N_{\text{air}}^{\uparrow}\right)\, 8.833\,10^{-33}\, \nu^{1.3} \; . \tag{2.79}
\end{aligned}
$$

The total column density for air N_air is simply the sum of the individual molecular columns

$$N_m(z_0) \equiv \int_{z_0}^{\infty} n_m(z)\,\mathrm{d}z \;=\; \int_{z_0}^{\infty} n_\text{air}(z)\,q_m(z)\,\mathrm{d}z \;=\; \int_{z_0}^{\infty} \frac{p_s(z)\,N_\text{A}\,q_m(z)}{g(z)\,m_\text{mol}^\text{air}}\,\mathrm{d}z \;, \qquad (2.80)$$

where $q_m(z)$ is the mole fraction of molecule m at level z, p_s is partial pressure, z_0 is the ground elevation (surface altitude) and g the gravitational acceleration of the Earth and m_mol^air the molar mass of air that is approximately constant in the so called homosphere ($z < 80\,\text{km}$). Figure 2.13 shows the optical depth for the CO retrieval interval according to Eq. (2.79).

2.5.2 Rayleigh extinction

Elastic scattering by air molecules is referred to as Rayleigh scattering and was originally formulated by Lord Rayleigh ($1842-1919$). It describes scattering of radiation by particles with a size parameter $x = 2\pi\,r\,\nu$ that is $2\cdot10^{-3} < x < 2\cdot10^{-1}$ with r designating the radius of a spherical particle and ν the wavenumber of the incident radiation. Similar to the aerosol cross section, the loss in intensity due to Rayleigh scattering is calculated by the scattering cross section k_ray per molecule (*Bodhaine et al.*, 1999). It is essentially independent of n_air so that the optical depth τ_ray at a particular wavenumber for a double path through the atmosphere is given by the product

$$\tau_\text{ray}(\nu) \;=\; \int\limits_\text{earth}^\text{sun} k_\text{ray}(\nu)n_\text{air}(z')/\cos(\theta)\,\mathrm{d}s' + \int\limits_\text{earth}^\text{ToA} k_\text{ray}(\nu)n_\text{air}(z'')/\cos(\pi)\,\mathrm{d}s''$$

$$\;=\; \left(N_\text{air}^{\downarrow} \;+\; N_\text{air}^{\uparrow}\right)\,k_\text{ray}(\nu)\;. \qquad (2.81)$$

Rayleigh extinction by scattering is roughly proportional to the fourth power of a wave's frequency $k_\text{ray} \propto \nu^4$ (*Seager*, 2010). As a consequence scattering decreases rapidly with increasing wavelength so that, according to Eq. (2.76), scattering is approximately 600 times less effective in the SWIR ($\approx 3\cdot10^{-4}$ at $2.3\,\mu\text{m}$) compared to the intensity scattered at 550 nm (*Bodhaine et al.*, 1999, Tab. 3). The Rayleigh optical thickness in Fig. 2.13 was calculated according to *Zdunkowski et al.* (2007, Eq. (11.12)).

Different authors proposed various models to approximate k_ray in order to compute the Rayleigh optical depth (*Nicolet*, 1984; *Bucholtz*, 1995; *Bodhaine et al.*, 1999) and also datasets for the scattering cross sections at different wavelength have been compiled, e. g., by *Penndorf* (1957); *Sneep and Ubachs* (2005); *Thalman et al.* (2014) or the DLR Institute of Planetary Research (DLR-IPF) in Berlin (F. Schreier, personal communication).

2.6 Continuum absorption

The continuous spectral absorption by a gas without an apparent line structure (resonance lines) is known as continuum absorption (*Shine et al.*, 2012; *Mlawer et al.*, 2012). This non-resonant absorption is defined empirically as the difference between and the experimentally observed total absorption and the calculated contribution of molecular absorption, aerosol extinction, and Rayleigh scattering (*Serov et al.*, 2017; *Elsey et al.*, 2020)

$$\mu^{(\text{con})}(\nu, s) \;=\; \mu - \mu_a^{(\text{gas})} - \mu_a^{(\text{aer})} - \mu_s^{(\text{gas})} - \mu_s^{(\text{aer})}\;. \qquad (2.82)$$

Although the causes of the continuum remain a subject of controversy it is postulated to have its origin in far-wing broadening, e.g., by collisional effects, and absorption due to water dimers (*Shine et al.*, 2012). In general the continuum is broken down into two components, i.e. the foreign- and self continuum. In case of e.g., water, the foreign continuum arises from interactions between H_2O molecules with other (abundant) atmospheric molecules such as N_2 or O_2 while the self continuum is due to the interaction of two H_2O molecules (*Shine et al.*, 2012). In the terrestrial atmosphere where molecular densities are sufficiently high, even molecules that have no intrinsic dipole moment such as nitrogen N_2 or oxygen O_2 absorb radiation (*Karman et al.*, 2019).

In the empirical CKD model the definition of the water vapor continuum coefficient is defined as the sum of the contributions from all lines beyond $25\,\mathrm{cm}^{-1}$ from its center (*Shine et al.*, 2012; *Mlawer et al.*, 2012). It constitutes a slowly varying function which is tabulated at $10\,\mathrm{cm}^{-1}$ intervals in the CKD (and MT-CKD) models. The CKD continuum used a wavenumber dependent 'χ function' modified line profile (super- and sub-Lorentzians, depending whether $\chi \gtrless 1.0$, respectively) in order to account for far-wing mechanisms causing the continuum. The updated MT-CKD model includes inelastic collisions of molecules (collision-induced absorption) in addition to far-wing contributions and so called 'fudge factors' to account for discrepancies with respect to latest measurements. However, there is strong evidence that the MT-CKD still underestimates the H_2O continuum absorption in the SWIR windows (*Shine et al.*, 2012).

Beside the two CKD models, collision-induced absorption accounts for the contribution of binary molecular complexes to absorption cross sections (*Karman et al.*, 2019). More specifically, in a gas mixture of two molecules A and B the continuum contribution from molecular pairs would be

$$\mu^{(\mathrm{con})}(\nu, s) \;\;=\;\; k_{(A-A)}\,\hat{n}_{(A)}^2 \;+\; k_{(A-B)}\,\hat{n}_{(A)}\,\hat{n}_{(B)} \;+\; k_{(B-B)}\,\hat{n}_{(B)}^2 \,, \qquad (2.83)$$

so that the volume absorption coefficient is given by

$$\mu_a^{(\mathrm{gas})}(\nu, s) \;\;=\;\; k_A(\nu)\,n_A(s) \;+\; k_B(\nu)\,n_B(s) \;+\; \mu^{(\mathrm{con})}(\nu, s) \,. \qquad (2.84)$$

It is composed of the (monomer) contributions according to Eq. (2.19) and the CIA contribution caused by collisions-induced absorption of molecular pairs $A-B$ with number densities $\hat{n}_{(A)}$ and $\hat{n}_{(B)}$ of corresponding units. In GARLIC/BIRRA inelastic collisions of pairs including H_2O, N_2, O_2, and CO_2 can be considered (also see *Karman et al.*, 2019, Tab. 1).
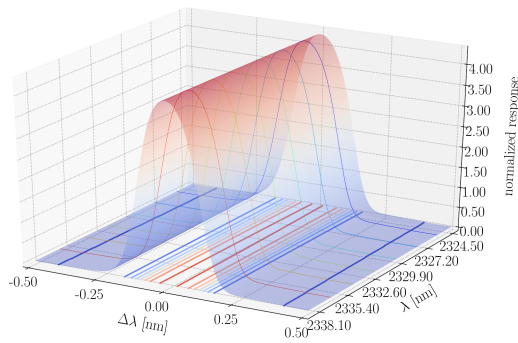
2.7   Instrument — spectral response

The monochromatic spectrum is subject to smoothing because of an instrument's finite spectral resolution of $\lambda/\Delta\lambda = \Delta\nu/\nu$, where $\Delta\lambda$ designates the smallest spectral separation between two peaks that the instrument can resolve at wavelength λ. The ISRF represents the instrument's response $\mathcal{S}$ to a monochromatic stimulus and is required for the convolution of the high-resolution spectrum to instrumental resolution (*Beirle et al.*, 2017). Given that $\mathcal{S}$ is not depending on λ but only on the difference $\lambda - \lambda'$ the signal recorded by the instrument can be described by

$$\hat{I}(\lambda) \;=\; I(\lambda) \otimes \mathcal{S}(\lambda) \;\;\longrightarrow\;\; (I \otimes \mathcal{S})(\lambda) \;=\; \int I(\lambda')\,\mathcal{S}(\lambda - \lambda')\,\mathrm{d}\lambda' \,. \qquad (2.85)$$

The retrieval of number densities of atmospheric constituents from spectrometers requires that instruments are calibrated and well characterized with respect to known radiometric sources (*Kleipool et al.*, 2018). An importrant part of the on-ground calibration is the measurement of the ISRF since it is directly linked to the radiative transfer model by convolution (*van Hees et al.*, 2018). A measured and tabulated ISRF can either be directly applied to the convolution of the monochromatic signal (Pub. III) or it can be parameterized by an appropriate function $\mathcal{S}(\lambda, \zeta)$ where the parameter ζ accounts for instrument characteristics (Sec. 4.6). In case of diffraction grating spectrometers the ISRF can often be approximated by a Gaussian function (*Gimeno García et al.*, 2011; *Munro et al.*, 2016), however, also more complex parameterizations with several parameters have been developed (e. g. *Veefkind et al.*, 2012).

In case of TROPOMI the results of the calibration measurements are stored in the calibration key data (CKD). The main spectral characteristics of the TROPOMI's SWIR spectrometer is provided in (*Kleipool et al.*, 2018, Tab. 1). In the CKD the response data for a given detector pixel along spectral dimension is provided as a normalized function of wavelength. Figure 2.14 shows the response function along the spectral and spatial axes. In course of the instrument's on-ground calibration the pixel's response was determined up to 4.5 pixels away from the center (i.e. within a range of sufficient SNR) and the ISRF is set to zero outside that range. The authors in *Kleipool et al.* (2018); *van Hees et al.* (2018) conclude that the ISRF determined for the SWIR spectrometer meets the requirements (i.e. $< 1\%$ of its maximum) and should thus be sufficient for trace gas retrieval over the full operational lifetime.

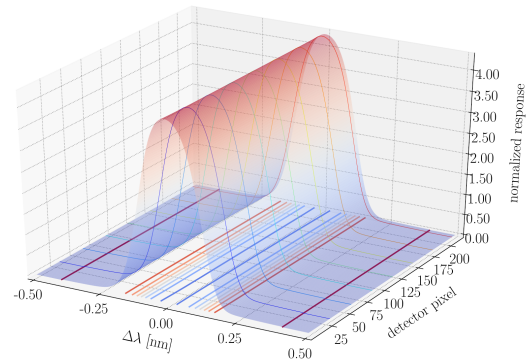**(a)**   **(b)**

Figure 2.14: (a) The tabulated SWIR spectral response function across the spectral axis for the detector pixel 128 (center, nadir direction) of the TROPOMI instrument. **(b)** The spectral response for a single spectral pixel (at $\approx 4289\,\mathrm{cm}^{-1}$) across the detector's spatial axis.

Often, however, an appropriate function is used to parameterize the ISRF upon one or multiple parameters (*Beirle et al.*, 2017). For example, various analytical functions were proposed for the spectral response in the channels of the SCIAMACHY instrument. Other studies by e. g., *Gimeno García et al.* (2011) or Pub. I used a Gaussian to parameterize the instrument's spectral response which allowed to include an estimation for the shift and width of the (unknown) ISRF in the Level 2 retrieval. In case of the SWIR spectrometer aboard TROPOMI the measured ISRF is slightly flat-topped (*van Hees et al.*, 2018; *Beirle et al.*, 2017). In the effort to find an appropriate analytical ISRF for the TROPOMI

SWIR of the CO retrieval window ($4277.2 - 4302\,\mathrm{cm^{-1}}$) a skewed Gaussian model with parameters for amplitude, shift, half width, and skewness was examined. An advantage of using a parameterized model for the ISRF is that it allows to account for apparent modifications of the tabulated response values specified in the CKD. In particular, a heterogeneous reflectivity in the observed scene can cause inhomogeneous illumination of the entrance slit of the instrument. Scenes with strong variations in the reflectivity such as observations along coastlines do alter the instrument's spectral response (a.k.a. slit function) for which Level 2 processing should account for (*Noël et al.*, 2012; *Hummel et al.*, 2021).

A function that is often used for the parameterization of the ISRF is the Gaussian function

$$G(\nu) \;=\; \frac{1}{\sqrt{2\pi}\sigma}\,\exp\left(-\frac{(\nu - \hat{\nu})^2}{2\sigma^2}\right), \tag{2.86}$$

with the property that the integral over all wavenumbers gives a probability of 1 according to

$$\int_{-\infty}^{\infty} G(\nu)\,\mathrm{d}\nu \;=\; 1\,. \tag{2.87}$$

It is important to note that, strictly speaking, the subsequent parameterizations are normalized on an infinite interval, however, the ISRF is only calculated on a finite interval.

The half width at half maximum (HWHM) of this symmetric distribution is given by

$$\gamma \;=\; \pm\sqrt{2\ln 2}\,\sigma\,, \tag{2.88}$$

where σ represents the standard deviation of Eq. (2.86) from $\hat{\nu}$, and 2γ corresponds to the full width (FWHM). Since a spectrometer's resolution is often specified in terms of FWHM, Eq. (2.86) is often expressed in terms of γ according to

$$\mathcal{S}_2(\nu) \;=\; \frac{\sqrt{\ln 2}}{\gamma\sqrt{\pi}}\,\exp\left(-\frac{\ln 2\,\nu^2}{\gamma^2}\right). \tag{2.89}$$

In order to model the flat-topped distribution of TROPOMI's spectral response a generalized normal distribution (*Nadarajah*, 2005; *Beirle et al.*, 2017) is required, i. e., a class of functions

$$\mathcal{S}_k(\nu) \;=\; A(\omega, k)\,\exp\left(-\left|\frac{\nu}{\omega}\right|^k\right), \tag{2.90}$$

with the parameter $k > 2$ and the normalization of the integral to 1 via

$$A(\omega, k) \;=\; \frac{k}{2\,\omega\,\Gamma\left(\frac{1}{k}\right)}\,, \tag{2.91}$$

where Γ is the gamma function. If γ is used to describe the width then

$$\omega_k \;=\; \frac{\gamma}{\sqrt[k]{\ln 2}} \tag{2.92}$$

is depending on k while if the half width at $1/e$ maximum (HWEM=ζ) is used then $\omega = \zeta$ and hence not dependent on k (see *Beirle et al.*, 2017, Fig. 1). The super-Gaussian with half width γ is then given by

$$\mathcal{S}_4(\nu) \;=\; \frac{2\sqrt[4]{\ln 2}}{\gamma\,\Gamma\left(\frac{1}{4}\right)}\,\exp\left(-\frac{\ln 2\,\nu^4}{\gamma^4}\right). \tag{2.93}$$

The application of the above introduced ISRFs as well as an assessment of more sophisticated instrument line shapes with respect to TROPOMI is presented in Sec. 4.6.

Chapter 3

Methods: Retrieval — inverse problem

3.1 Inversion

The extraction of atmospheric properties and constituents from a spectrum constitutes an ill-posed inverse problem (*Rodgers*, 2000; *Neto and da Silva Neto*, 2012). The retrieval generally consists of a forward model $F(\boldsymbol{x})$ and an inverse method in order to solve for the quantity of interest (see Fig. 3.1). The most important mathematical aspects that are encountered in the solution of ill-posed inverse problems such as the retrieval of trace gas concentrations from an observed spectrum are described subsequently (*Hansen et al.*, 2013; *Aster et al.*, 2018; *Gill et al.*, 2019). This is important in order to be able to interpret the results in Sec. 4 and apprehend the methods limitations.

3.1.1 Ill-posed problems

In atmospheric remote sensing the property of interest $X \in \mathbb{R}$ is usually a function of altitude (e.g. molecular concentrations or temperature profile) that is not observed
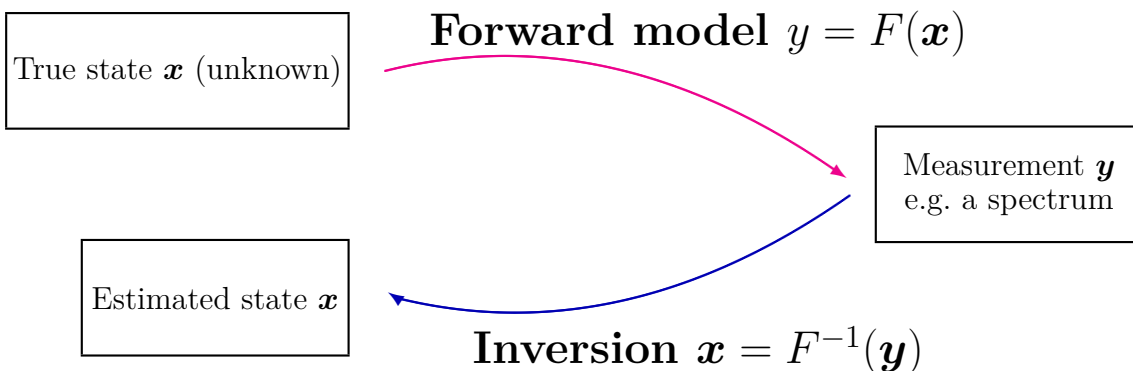

Figure 3.1: Schematic depiction of the forward model and inversion methodology. The forward model provides a spectrum for a given state vector $\boldsymbol{x}$. The retrieval does the inverse and yields an estimate for $\boldsymbol{x}$ for a given spectrum $\boldsymbol{y}$.

directly but represented by a model function $f(\nu)$, generally represented by the radiative transfer equation (see Sec. 2.2). In atmospheric inversion problems the Fredholm integral of the first kind appears (*Hanson and Phillips*, 1975)

$$f(\nu) \;=\; \mathcal{I}[X(z)] \;\equiv\; \int_{z_0}^{z_1} k(\nu, z)\, X(z)\, \mathrm{d}z \,, \tag{3.1}$$

with $f(\nu)$ the image of $X(z)$ under the integral transformation $\mathcal{I}$. The function $k(\nu, z)$ is referred to as the integral kernel of the transform. It is considered to be given, and encodes the physics relating an unknown function $X(z)$ to corresponding data $f(\nu)$. The goal is to find the inverse transform In general, the smoothness of the integral kernel limits the ability to recover higher frequency components in $X(z)$ from Eq. (3.1) (*Margerum*, 1983). This causes ambiguity that is often described as inherent instability or ill-posedness.

According to Jacques Salomon Hadamard $(1865 - 1963)$ a problem is called well-posed if (a) there exists a solution to the problem, (b) there is at most one solution to the problem, and (c) the solution depends continuously on the data. In the case the problem is not well-posed it is called ill-posed. The evaluation of a function at a given point can be well- or ill-posed (*Neto and da Silva Neto*, 2012). It is an intrinsic property of the function being evaluated and it does not depend on approximations.

In cases where the mathematical model for a system's physics is approximate or the data is noisy, no exact solution might exist (a). The reason for (b) is that in many physical phenomenon (spectroscopy, gravitational fields, ...) the result can be the same for different models. For case (c), given that the model is continuous and a unique solution exists, a more practical definition could be used, e.g. that a model $F(X, \nu)$ is well-posed at X if $|(F(X + \varepsilon), \nu) - F(X, \nu))/\varepsilon| \leq 1$ holds, and ill-posed otherwise (*Neto and da Silva Neto*, 2012).

3.1.2 Discrete inverse problems

In practice, even if the function $f(\nu)$ might in theory be known, function values are usually provided by measurements of finite resolution. In atmospheric remote sensing a spectrometer measures at discrete wavenumbers ν_j to provide a set of intensity values over some spectral range. A consequence of the discrete representation of the spectrum is that the object of interest $X(z)$ also needs to be defined on a discrete grid. Series expansion of order N with basis functions ψ_i^N give

$$X(z) \;\approx\; X_N(z) \;=\; \sum_{i=1}^{N} a_i\, \psi_i^N(z) \,, \tag{3.2}$$

so that the continuous Eq. (3.1) can be represented in the discrete form

$$f(\nu_j) \;=\; \sum_{i=1}^{N} a_i \int_{z_0}^{z_1} k(\nu_j, z)\, \psi_i^N(z)\, \mathrm{d}z \,, \tag{3.3}$$

where the solution consists of determining the constants a_i. It can be formulated as

$$f_j \;=\; \sum_{i=1}^{N} K_{ji}\, x_i \,, \tag{3.4}$$

with the operator

$$K_{ji} \; = \; \int_{z_0}^{z_1} k_j(z)\, \psi_i^N(z)\, \mathrm{d}z \,, \tag{3.5}$$

and x_i corresponds to the coefficient a_i. A model $\mathsf{K}\,\boldsymbol{x}$ of a discrete spectrum $\boldsymbol{y} \in \mathbb{R}^M$ with parameters $x_i \in \boldsymbol{x}$ is corresponding to a linear transformation from $\mathbb{R}^N \to \mathbb{R}^M$ according to

$$\boldsymbol{y} \; = \; \mathsf{K}\,\boldsymbol{x} \,, \tag{3.6}$$

where $\boldsymbol{x}$ is a column vector with N elements (model parameters) and K is an $M \times N$ matrix (*Zhang*, 2011; *Golub and Van Loan*, 2013). Equation (3.6) constitutes a linear discrete inverse problem with respect to $\boldsymbol{x}$, hence methods of linear algebra can be applied in order to solve the problem (*Gentle*, 2017). In principle, if K is a square matrix ($M = N$) of independent columns, the solution for the inverse problem reduces to a simple matrix inversion. However, from the set of linear equations that aim to fit the model vector $\boldsymbol{x}$ to measurement data $\boldsymbol{y}$ issues such as solution existence, solution uniqueness, and instability of the solution remain.

3.1.3   Discrete ill-posed inverse problems

For the subsequent sections it is useful to recall that given a matrix $\mathsf{K}^{M \times N}$ the dimension (rank) of the row space is equal to the dimension of the column space $\mathrm{rk}(\mathsf{K}) = \mathrm{rk}(\mathsf{K}^T) \in \mathbb{R}$ (*Zhang*, 2011). The span of the columns $\boldsymbol{v}$ in K define the range of the matrix $\mathcal{R}$ according to

$$\mathcal{R}(\mathsf{K}) \; = \; \mathrm{span}(\boldsymbol{v}_l)_{l=1,\dots,N} \; = \; \left\{ \mathsf{K}\,\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{R}^N \right\}, \tag{3.7}$$

that are all vectors $\boldsymbol{y} \in \mathbb{R}^M$ for which $\mathsf{K}\,\boldsymbol{x} = \boldsymbol{y}$ is consistent. The column space is a subset of $\mathbb{R}^M$ while the row space with range $\mathcal{R}(\mathsf{K}^T)$ is a subset of $\mathbb{R}^N$. Furthermore, the (column) null space $\mathcal{N}$, defined as

$$\mathcal{N}(\mathsf{K}) \; = \; \left\{ \mathsf{K}\,\boldsymbol{x} = \boldsymbol{0} \mid \boldsymbol{x} \neq \boldsymbol{0} \right\}, \tag{3.8}$$

with $\boldsymbol{x} \in \mathbb{R}^N$ is a subset of $\mathbb{R}^N$ with the dimension $\dim(\mathcal{N}(\mathsf{K})) = N - \mathrm{rk}(\mathsf{K})$ according to the rank-nullity theorem (*Zhang*, 2011, Thm. 1.3).

In practice, no model might fit the data exactly, so that the measured data $\boldsymbol{y}$ is not in the range of matrix K, hence

$$\boldsymbol{y} \notin \mathcal{R}(\mathsf{K}): \; \mathsf{K}\boldsymbol{x} \neq \boldsymbol{y} \mid \boldsymbol{y} \in \mathbb{R}^M, \, \forall \boldsymbol{x} \in \mathbb{R}^N \,, \tag{3.9}$$

and no exact solution exists. Non-uniqueness, on the other hand, is a characteristic of rank deficiency meaning the matrix K has a non-empty null space $\mathcal{N}(\mathsf{K}) \neq \{0\}$ so that any linear combination of vectors $\boldsymbol{x}$ can be added to Eq. (3.6) without affecting $\boldsymbol{y}$.

As radiance measurements at different wavenumbers do not contain independent information on the individual parameters of $\boldsymbol{x}$, the columns in K do not form an orthogonal basis. As a result, the inverse solution for $\boldsymbol{x}$ can be very unstable with respect to small changes $\boldsymbol{\varepsilon}$ in the data vector

$$\boldsymbol{y} + \boldsymbol{\varepsilon} \; = \; \boldsymbol{y}^{(\varepsilon)} \; = \; \mathsf{K}\,\boldsymbol{x}^{(\varepsilon)} \,. \tag{3.10}$$

Consequently, small errors in the measurement, caused by noise, can induce substantial change in the solution $\boldsymbol{x} - \boldsymbol{x}^{(\varepsilon)} \gg \boldsymbol{\varepsilon}$ so that $\boldsymbol{x}^{(\varepsilon)} \to \boldsymbol{x}$ whenever $\boldsymbol{y}^{(\varepsilon)} \to \boldsymbol{y}$ is violated (c) (*Gill et al.*, 2019).

The ill-conditioned nature of the problem brings that K is not automatically over-determined if there are more equations than unknowns $M > N$. Some of the components in the solution $\boldsymbol{x}$ might be over-determined, while other components are under-determined (*Aster et al.*, 2018). Whether a component is over- or under-determined depends on how the parameter is mapped from the vector space $\boldsymbol{x} \in \mathbb{R}^N$ to the measurement space $\boldsymbol{y} \in \mathbb{R}^M$ and also on the magnitude of $\boldsymbol{\varepsilon}$ which will be further examined in the subsequent sections. The ill-posedness is hence caused by both, inconsistent measurements and model parameters that cannot be determined from the measurement vector $\boldsymbol{y}$.

3.1.4   Solving inverse problems

Given the matrix $\mathsf{K}^{M \times N}$ with $M \geq N$ it is said to have full rank if $\mathrm{rk}(\mathsf{K}) = N$ so that $\mathcal{R}(\mathsf{K}) \in \mathbb{R}^N$. Consequently, the matrix has a trivial null space of dimension $\dim(\mathcal{N}(\mathsf{K})) = 0$. A so called left inverse K^+ can be formulated using the transposed matrix of K given by K^T. The product $\mathsf{K}^T \mathsf{K}$ is a positive definite symmetric $N \times N$ matrix that is injective if it is of full rank. Moreover, every positive (semi-)definite matrix is convex which guarantees that Eq. (3.10) has a unique closed-form solution (*Gentle*, 2017). Therefore, it is not singular and has non-negative real eigenvalues. The inverse of a squared matrix is given by

$$\left(\mathsf{K}^T \mathsf{K}\right)^{-1} \;=\; \mathsf{K}^+ \left(\mathsf{K}^T\right)^{-1} , \tag{3.11}$$

so that

$$\mathsf{K}^+ \left(\mathsf{K}^T\right)^{-1} \left(\mathsf{K}^T \mathsf{K}\right) \;=\; \mathsf{I}, \tag{3.12}$$

and the left inverse (generalized/pseudo inverse) can be given by

$$\mathsf{K}^+ \;=\; \left(\mathsf{K}^T \mathsf{K}\right)^{-1} \mathsf{K}^T . \tag{3.13}$$

The product $\mathsf{K}^+ \mathsf{K}$ only equals the identity matrix I if $\boldsymbol{y} \in \mathbb{R}^N$ and K is positive definite (*Zhang*, 2011, Sec. 7.1). Accordingly, $\boldsymbol{y} \in \mathcal{R}(\mathsf{K})$, so that the generalized inverse K^{-1} does provide the exact solution

$$\mathsf{K}^{-1} \boldsymbol{y} - \boldsymbol{x} \;=\; \mathbf{0} . \tag{3.14}$$

In the over-determined case, however, $\mathsf{K}^+ \mathsf{K} = \mathsf{P}$ gives a matrix of size $N \times N$ that is the projection of $\boldsymbol{y} \in \mathbb{R}^M$ onto the column space $\mathcal{R}(\mathsf{K}) \in \mathbb{R}^N$ so that

$$\mathsf{K}^{-1} \boldsymbol{y} - \mathsf{P} \boldsymbol{x} \;=\; \boldsymbol{y}_\perp , \tag{3.15}$$

with $\boldsymbol{y}_\perp \in \mathsf{N}(\mathsf{K})$. As a result, $\mathsf{K}^{-1} \boldsymbol{y}$ is not an exact solution (a) for $\boldsymbol{x}$ but the one with the minimum 2-norm

$$\min_{\boldsymbol{x}} \|\boldsymbol{y} - \mathsf{K} \boldsymbol{x}\|_2 \;\;\Longleftrightarrow\;\; \min \|\boldsymbol{y}_\perp\|_2^2 . \tag{3.16}$$

In cases where $\mathrm{rk}(\mathsf{K}) < N$ the system of equations will in general have either no exact solution (a) or infinitely many solutions (b). Such rank deficient and ill-conditioned problems are over- and under-determined at the same time! The reason is that in this case K has a non-trivial null space $\mathcal{N}(\mathsf{K}) \neq \{0\}$ and data vectors $\boldsymbol{y}$ are outside the operators range $\mathcal{R}(\mathsf{K}) \notin \mathbb{R}^N$ even if $\boldsymbol{y} \in \mathbb{R}^N$. Most inverse problems in atmospheric science, including

the retrieval of atmospheric parameters from an observed spectrum, deal with the latter kind of problem.

The characteristics of rank deficient and ill-conditioned linear systems can be more thoroughly examined with the singular value decomposition (SVD, *Golub and Van Loan*, 2013, Sec. 2.4). The SVD exists for any matrix and is a factorization scheme that allows to examine various properties of a matrix K and compute a matrix's pseudo-inverse K^+. The SVD of the matrix $\mathsf{K} \in \mathbb{R}^{M \times N}$ is defined as

$$\mathsf{K} = \mathsf{U} \Sigma \mathsf{V}^T \,, \tag{3.17}$$

with unitary matrices $\mathsf{U} \in \mathbb{R}^{M \times M}$, $\mathsf{V} \in \mathbb{R}^{N \times N}$ and a diagonal matrix $\Sigma \in \mathbb{R}^{M \times N}$.

The unitary matrices U and V are orthogonal, so they are composed of column vectors (of length one) $\boldsymbol{u}_i \in \mathsf{U}$, $\boldsymbol{v}_i \in \mathsf{V}$ which form an orthogonal basis $\langle \boldsymbol{u}_k^T, \boldsymbol{u}_l \rangle = \delta_{kl}$ and $\langle \boldsymbol{v}_k^T, \boldsymbol{v}_l \rangle = \delta_{kl}$, respectively. The range of U and V is given by the span of the set of their vectors

$$\text{span}(\boldsymbol{u}_l)_{l=1,\dots,M} = \mathcal{R}(\mathsf{U}) \in \mathbb{R}^M \,, \tag{3.18}$$

$$\text{span}(\boldsymbol{v}_l)_{l=1,\dots,N} = \mathcal{R}(\mathsf{V}) \in \mathbb{R}^N \,. \tag{3.19}$$

The diagonal elements of Σ are σ_l with $l = 1, \dots, \min(M, N)$ represent the singular values that correspond to the vectors of both, U and V. The product $\mathsf{K}^T\mathsf{K}$, by which the left inverse K^+ was defined in Sec. 3.1.4, can be given in terms of the unitary matrix V

$$\mathsf{K}^T\mathsf{K} = \mathsf{V} \Sigma^2 \mathsf{V}^T \implies \mathsf{K}^T\mathsf{K}\mathsf{V} = \mathsf{V} \Sigma^2 \,, \tag{3.20}$$

so that the square roots of the N eigenvalues of $\mathsf{K}^T\mathsf{K}$ are the real-valued (right) singular values $\sigma_i \geq 0$ of K. The singular values are ordered $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_R$ where R is the smallest non-zero singular value which determines the rank of the matrix $\text{rk}(\mathsf{K}) = R$. In numerical applications, for example, the criterion that defines the rank can be formulated as

$$\sigma_R : \min(\Sigma_{ll}) \geq \varepsilon \mid \varepsilon \in \mathbb{R}_{>0},\, R \leq N \,. \tag{3.21}$$

The condition number $\text{cnd}(\mathsf{K}) \in \mathbb{R}^+$ is then given by

$$\text{cnd}(\mathsf{K}) = \frac{\sigma_1}{\sigma_R} \,. \tag{3.22}$$

It is a measure for the sensitivity of the solution to small perturbations (e.g. noise, rounding errors, ...) in the input data (c) and can become very large as $\text{cnd}(\mathsf{K}) \to \infty$ whenever $\sigma_R \to 0$.

Provided that $R = N$ and the input data is given by $\boldsymbol{y} \in \mathbb{R}^M$, the diagonal matrix Σ contains non-zero singular values up to

$$\mathsf{K}\boldsymbol{v}_l = \sigma_l \boldsymbol{u}_l \qquad \text{for} \quad l = 1, \dots, N \,, \tag{3.23}$$

so that the range and null space of the ill-conditioned system is given by

$$\text{span}(\boldsymbol{u}_l)_{l=1,\dots,N} = \mathcal{R}(\mathsf{K}) \in \mathbb{R}^N \,, \tag{3.24}$$

$$\text{span}(\boldsymbol{u}_l)_{l=N+1,\dots,M} = \mathcal{N}(\mathsf{K}) \,. \tag{3.25}$$

In analogy to Eq. (3.15), the measurement vector has two components

$$\boldsymbol{y} \quad = \quad \sum_{l=1}^{N} \langle \boldsymbol{y}, \boldsymbol{u}_l \rangle \boldsymbol{u}_l \,, \tag{3.26}$$

$$\boldsymbol{y}_\perp \quad = \quad \sum_{l=N+1}^{M} \langle \boldsymbol{y}, \boldsymbol{u}_l \rangle \boldsymbol{u}_l \,, \tag{3.27}$$

so that no exact solution exists (see Eq. (3.9)).

In the case of rank deficient ($R < N$) ill-conditioned systems, only the singular values up to σ_R are non-zero (see Eq. (3.21)) so that the matrix has a non-trivial null space, and hence

$$\mathsf{K}^T \mathsf{K} \, \boldsymbol{v}_l \ = \ \sigma_l \boldsymbol{v}_l \qquad \text{for} \quad l = 1, \ldots, R \tag{3.28}$$

$$\mathsf{K}^T \mathsf{K} \, \boldsymbol{v}_l \ = \ \boldsymbol{0} \qquad \text{for} \quad l = R+1, \ldots, N \,. \tag{3.29}$$

The decomposition of K into unitary and diagonal matrices $\mathsf{U} \, \Sigma \, \mathsf{V}^T$ allows to formulate the Moore-Penrose pseudo-inverse

$$\mathsf{K}^+ \ = \ V \, \Sigma^{-1} \, U^T \ = \ \sum_{l=1}^{N} \frac{1}{\sigma_l} \, \boldsymbol{v}_l \, \boldsymbol{u}_l^T \,, \tag{3.30}$$

which is equal to the inverse in Eq. (3.13) and exists for every matrix. Equivalent to Eq. (3.15), it is the solution with the minimum 2-norm. In the under-determined case, K^+ provides the solution which is minimized with respect to its 2-norm (see Eq. (3.16)).

3.2   Least squares

3.2.1   Linear least squares

The linear least squares solution for $\boldsymbol{x} \in \mathbb{R}^N$ corresponds to the vector that minimizes the 2-norm of the squared residual between the measurements $\boldsymbol{y} \in \mathbb{R}^M$ and the model

$$\min_x \|\boldsymbol{y} - \mathsf{K} \, \boldsymbol{x}\|_2^2 \,, \tag{3.31}$$

so that the residuum (objective) function to be minimized is given by

$$r(\boldsymbol{x}) \ = \ \left[\boldsymbol{y} - \mathsf{K} \, \boldsymbol{x} \right]^T \left[\boldsymbol{y} - \mathsf{K} \, \boldsymbol{x} \right] . \tag{3.32}$$

Provided that the columns of K comprise a set of linearly independent (but in general not orthogonal) vectors of size $M \geq N$, its second order derivative

$$\frac{\partial^2 r}{\partial \boldsymbol{x} \partial \boldsymbol{x}^T} \ = \ 2 \, \mathsf{K}^T \mathsf{K} \,, \tag{3.33}$$

is positive definite so that Eq. (3.32) constitutes a convex function with a unique global minimum (*Hansen et al.*, 2013; *Rust*, 2002) given by

$$\frac{\partial r}{\partial \boldsymbol{x}^T} \ = \ 2 \, \mathsf{K}^T \mathsf{K} \boldsymbol{x} - 2 \mathsf{K}^T \boldsymbol{y} \ = \ \boldsymbol{0} \,. \tag{3.34}$$

The appearance of the factor two explains the factor $1/2$ that is often introduced with the residuum function. The solution can be represented in form of the so called normal equation

$$\mathsf{K}^T\mathsf{K}\,\boldsymbol{x} \;=\; \mathsf{K}^T\boldsymbol{y}\,. \tag{3.35}$$

In analogy to Eqs. (3.15) and (3.23), with the columns of K assumed to be linearly independent, the matrix product on the left hand side is non-singular (*Rust*, 2001a). The least squares problem is then formally solved by

$$\boldsymbol{x} \;=\; \mathsf{K}^+\,\boldsymbol{y} \;=\; \left(\mathsf{K}^T\mathsf{K}\right)^{-1}\mathsf{K}^T\boldsymbol{y}\,. \tag{3.36}$$

In order to guarantee that Eq. (3.36) gives the solution with the minimum variance, known as the best linear unbiased estimate for $\boldsymbol{x}$ (*Rust*, 2001b), the (usually unknown) errors $\boldsymbol{\varepsilon}\in\mathbb{R}^M$ of the imperfect measurements

$$\boldsymbol{y} \;=\; \mathsf{K}\boldsymbol{x} + \boldsymbol{\varepsilon} \tag{3.37}$$

need to be unbiased and normally distributed with the mean

$$E(\boldsymbol{\varepsilon}) \;=\; \boldsymbol{0}\,, \tag{3.38}$$

and a symmetric, positive definite variance matrix with a common variance ς^2 (white noise)

$$E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T) \;=\; \varsigma^2\,\mathsf{I}\,. \tag{3.39}$$

If the imperfect measurements contain independent random errors of various magnitudes the noise (co-)variance matrix C of size $M\times M$ is given by

$$\mathsf{C} \;=\; \boldsymbol{\varsigma}^T\mathsf{I}\boldsymbol{\varsigma} \;=\; \mathbf{diag}(\varsigma_1^2,\ldots,\varsigma_M^2)\,. \tag{3.40}$$

In order to account for the statistical characteristics of the observations in the solution the least squares need to be modified. With the assumptions on errors the distribution of noise-contaminated measurements can be described by the joint probability function P_M (Gaussian distribution)

$$P_M(\boldsymbol{y}|\boldsymbol{x}) \;=\; \frac{1}{(2\pi)^{M/2}\,|\mathsf{C}|^{1/2}}\exp\left(-\frac{1}{2}\left[\boldsymbol{y}-\mathsf{K}\boldsymbol{x}\right]^T\mathsf{C}^{-1}\left[\boldsymbol{y}-\mathsf{K}\boldsymbol{x}\right]\right)\,, \tag{3.41}$$

which is formed by the product of the M individual probability functions. The likelihood L of the observation is then given by the joint probability density functions according to

$$L(\boldsymbol{x}|\boldsymbol{y}) \;=\; P(\boldsymbol{y}|\boldsymbol{x})\,. \tag{3.42}$$

The best estimate for $\boldsymbol{x}$ is found by maximizing the likelihood function for the given observations

$$\max_x\left\{L(\boldsymbol{x}|\boldsymbol{y})\right\}\,. \tag{3.43}$$

The residuum function that needs to be minimized in order to fulfill Eq. (3.43) is then given by

$$r(\boldsymbol{x}) \;=\; \left[\boldsymbol{y}-\mathsf{K}\boldsymbol{x}\right]^T\mathsf{C}^{-1}\left[\boldsymbol{y}-\mathsf{K}\boldsymbol{x}\right]\,. \tag{3.44}$$

which represents the absolute value $|\cdot|$ of the exponent without constant factors as they do not affect the maximization. The minimization problem for this weighted residuum function is

$$\min_x \left\{ r(\boldsymbol{x}) \right\} \;=\; \min_x \left\| \mathsf{C}^{-1/2} \left(\boldsymbol{y} - \mathsf{K}\boldsymbol{x} \right) \right\|_2^2 , \tag{3.45}$$

with the normal equations solution given by

$$\boldsymbol{x} \;=\; \left(\mathsf{K}^T \mathsf{C}^{-1} \mathsf{K} \right)^{-1} \mathsf{K} \mathsf{C}^{-1} \boldsymbol{y} . \tag{3.46}$$

In case of errors according to Eq. (3.39) the factors in the noise variance matrix cancel out, hence the value of ς^2 does not need to be known for the classical least squares (see Eq. (3.36) and *Rust*, 2001a).

3.2.2 Nonlinear least squares

A function f is nonlinear in α if the derivative $\partial f / \partial \alpha$ is a function of α (*Hansen et al.*, 2013). A fitting model F is nonlinear if it depends nonlinearly on one or more parameters $\boldsymbol{x} \in \mathbb{R}^N$. Examples for such models are the exponential decay

$$F(x_1, x_2, t) = x_1 \exp\left(-x_2 \, t \right) , \tag{3.47}$$

the solutions given by the wave equation

$$F(x_1, x_2, t) = x_1 \exp\left[\mathrm{i}(k x_2 - \omega t) \right] , \tag{3.48}$$

or, more specifically to atmospheric remote sensing, the convolution of a monochromatic spectrum with the instrument's spectral response function $S(\nu)$

$$F(x_1, x_2, \nu) = \int_{-\infty}^{\infty} S(\boldsymbol{x}, \nu - \tau) \, f(\tau) \, \mathrm{d}\tau . \tag{3.49}$$

A nonlinear least squares fit is required in order to find an estimate for $\boldsymbol{x}$ that minimizes the residuum function $r(\boldsymbol{x})$ with the measurements $\boldsymbol{y}$ and the model functions $\boldsymbol{F} \in \mathbb{R}^M$ in the 2-norm

$$\min_x \left\{ r(\boldsymbol{x}) \right\} \;=\; \min_x \left\| \boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}) \right\|_2^2 . \tag{3.50}$$

In principle, analogous to the linear case, the best estimate for $\boldsymbol{x}$ can be found by computing the gradient according to Eq. (3.34) and solve the equations for $\boldsymbol{x}$ (*Lawson and Hanson*, 1995). In the nonlinear case, however, this can be very challenging and often there is no closed-from solution for the estimate of $\boldsymbol{x}$ so that iterative algorithms are required to solve the optimization problem. And even if the problem does have a closed form solution, it may still be much more computationally efficient to solve the problem with iterative algorithms (*Chen and Surmont*, 1976). Moreover, ill-posed inverse problems are often not convex, so that parameters can be optimal among nearby sets of parameters without being globally optimal (*O'Leary and Rust*, 2013).

In general, iterative algorithms (see Fig. 3.2) start with an initial estimate $\boldsymbol{x}_0$ and proceed by a series of corrections

$$\boldsymbol{x}_{i+1} \;=\; \boldsymbol{x}_i + \delta \boldsymbol{x}_i , \tag{3.51}$$

that obtain each $\delta \boldsymbol{x}_i$ by solving a linear minimization problem according to Eq. (3.36) (*Rust*, 2002, 2003). For a given iteration step the minimization problem from above can then be written as

$$r_i(\delta \boldsymbol{x}) \;=\; \|\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_i + \delta \boldsymbol{x})\|_2^2 \,, \tag{3.52}$$

which is equivalent to

$$r_i(\delta \boldsymbol{x}) \;=\; \left[\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_i + \delta \boldsymbol{x})\right]^T \left[\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_i + \delta \boldsymbol{x})\right]. \tag{3.53}$$

The first order Taylor series expansion is used to linearize the forward model with respect to the quantities of interest

$$\boldsymbol{F}(\boldsymbol{x}_i + \delta \boldsymbol{x}) \;\approx\; \boldsymbol{F}(\boldsymbol{x}_i) + \mathsf{J}(\boldsymbol{x}_i)\,\delta \boldsymbol{x} \,, \tag{3.54}$$

where J is the Jacobian matrix for the state vector $\boldsymbol{x}$ which is a fundamental quantity for any nonlinear optimization method and is given by

$$\mathsf{J}(\boldsymbol{x}_i) \;\equiv\; \nabla \boldsymbol{F}(\boldsymbol{x}_i) \;=\; \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_M}{\partial x_1} & \frac{\partial F_M}{\partial x_2} & \cdots & \frac{\partial F_M}{\partial x_N} \end{pmatrix} \tag{3.55}$$

It comprises the partial derivatives with respect to the parameters in $\boldsymbol{x}$. With the Taylor series approximation from above, Eq. (3.53) becomes

$$r_i(\delta \boldsymbol{x}) \;=\; \left[(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_i)) + \mathsf{J}(\boldsymbol{x}_i)\,\delta \boldsymbol{x}\right]^T \left[(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_i)) + \mathsf{J}(\boldsymbol{x}_i)\,\delta \boldsymbol{x}\right], \tag{3.56}$$

which is similar to the linear least squares residuum function in Eq. (3.32). It can therefore be solved in the same way, i.e.

$$\delta \boldsymbol{x}_i \;=\; \left[\mathsf{J}_i^T \mathsf{J}_i\right]^{-1} \mathsf{J}_i^T \left(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_i)\right) \,, \tag{3.57}$$

In contrast to Eq. (3.36), however, it is not guaranteed that the corrected estimate $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \delta \boldsymbol{x}_i$ is closer to the 'true' value than $\boldsymbol{x}_i$. The iteration might converge to a local minimum which might not be a good fit for $\boldsymbol{y}$ and so the least squares need to be started again with a new $\boldsymbol{x}_0$ (*Rust*, 2002).

The Gauss–Newton method for the NLS (*Hansen et al.*, 2013) approximates the Hessian matrix $\mathsf{H} = \nabla^2 \boldsymbol{F} = \nabla \mathsf{J}$ by the product of the Jacobians $\mathsf{J}^T \mathsf{J}$, while the Levenberg–Marquardt adds a positive definite matrix D and a coefficient Λ as a adjustable damping constant

$$\delta \boldsymbol{x}_i \;=\; \left[\mathsf{J}_i^T \mathsf{J}_i + \Lambda \, \mathsf{D}\right]^{-1} \mathsf{J}_i^T \left(\boldsymbol{y} - \boldsymbol{F}(\boldsymbol{x}_i)\right) \,, \tag{3.58}$$

in order to counter the effects of ill-conditioning.

3.2.3 Separable least squares

In problems, where the model function F is a linear combination of nonlinear functions such as

$$F(x_1, x_2, x_3, x_4, \nu) \;=\; x_1 \exp\left(x_2 \, \nu\right) \,+\, x_3 \exp\left(x_4 \, \nu\right) \,+\, \ldots \,, \tag{3.59}$$

the SLS, also known as the Variable Projection method (*Golub and Pereyra*, 2003), can be applied to find a solution for the problem. The method eliminates the linear variables for
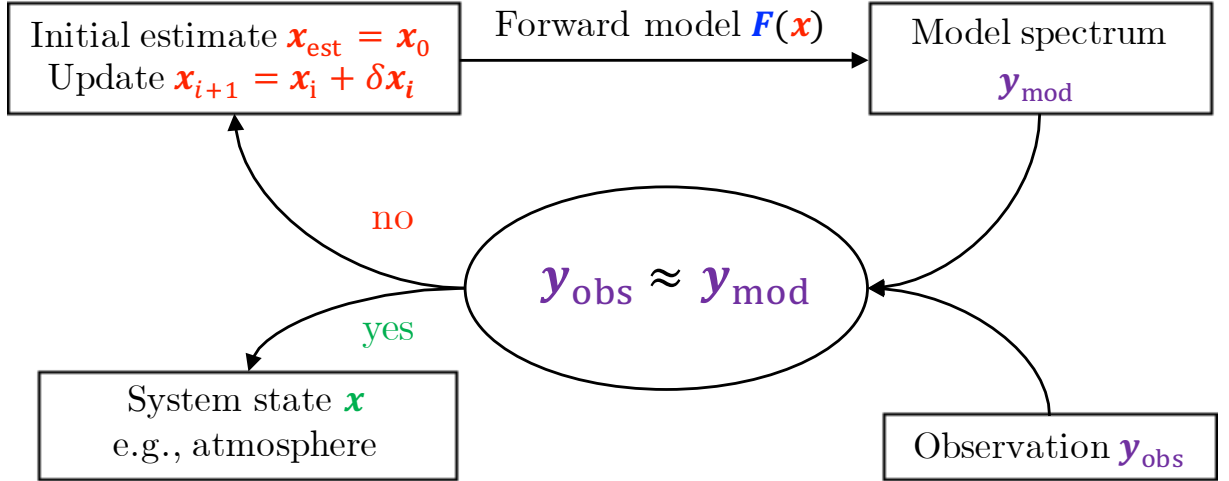
Figure 3.2: Iterative optimization algorithm. The number of iterations is influenced by the tolerance criteria $|\Delta \boldsymbol{x}| \, / \, |\boldsymbol{x}| < \varepsilon_x$ and $|\boldsymbol{y}_{\text{obs}} - \boldsymbol{y}_{\text{mod}}| < \varepsilon_y$. The PORT optimization library used in BIRRA stops the iteration process if one or both criteria are met.

the NLS fit in order to reduce the dimension of the parameter space that results in a better conditioned problem. The reduced number of parameters in the nonlinear minimization problem improves efficiency and reduces the number of local minimizers which makes convergence to the globally optimal solution more likely (*O'Leary and Rust*, 2013).

Fitting models such as Eq. (3.59) allow the state vector $\boldsymbol{x}$ to be separated in a vector of nonlinear $\boldsymbol{\eta}$ parameters and linear $\boldsymbol{\beta}$ parameters to be estimated

$$\boldsymbol{x} \longrightarrow (\boldsymbol{\eta}, \boldsymbol{\beta}) \qquad \text{with} \quad \boldsymbol{x} \in \mathbb{R}^N, \ \boldsymbol{\eta} \in \mathbb{R}^p, \ \boldsymbol{\beta} \in \mathbb{R}^q \ \text{and} \ N = p + q \,, \tag{3.60}$$

so that the minimization problem is given by

$$\min_x \left\{ r(\boldsymbol{x}) \right\} \; = \; \min_{\eta, \beta} \| \boldsymbol{y} - \boldsymbol{F}(\boldsymbol{\eta}, \boldsymbol{\beta}) \|_2^2 \,. \tag{3.61}$$

with the forward model $\boldsymbol{F} \in \mathbb{R}^M$ comprising the set of all parameters. The forward model can then be separated and formulated according to

$$\boldsymbol{F}(\boldsymbol{\eta}, \boldsymbol{\beta}) \longrightarrow \sum_{i=1}^q \beta_i \, \boldsymbol{f}_i(\boldsymbol{\eta}) \,. \tag{3.62}$$

with $\boldsymbol{f}_i \in \mathbb{R}^M$ for $i = 1, \ldots, q$. These vector valued model functions can be combined in a matrix

$$\mathsf{K}(\boldsymbol{\eta}) \equiv \Big(\boldsymbol{f}_1(\boldsymbol{\eta}), \boldsymbol{f}_2(\boldsymbol{\eta}), \ldots, \boldsymbol{f}_q(\boldsymbol{\eta}) \Big) \qquad \text{with} \quad \mathsf{K} \in \mathbb{R}^{M \times q} \tag{3.63}$$

where each model function is depending on nonlinear parameters. So for given $\boldsymbol{\eta}$, Eq. (3.61) represents a linear least squares problem for the state vector $\boldsymbol{\beta}$,

$$\min_\beta \| \boldsymbol{y} - \mathsf{K}\boldsymbol{\beta} \|_2^2 \,, \tag{3.64}$$

which is formally solved by

$$\boldsymbol{\beta} \; = \; \left(\mathsf{K}^{\mathrm{T}} \mathsf{K} \right)^{-1} \mathsf{K}^{\mathrm{T}} \boldsymbol{y} \,. \tag{3.65}$$

By inserting solution Eq. (3.65) into Eq. (3.61), the original least squares problem Eq. (3.61) can be represented only in terms of the nonlinear parameters $\boldsymbol{\eta}$ and becomes

$$\min_{\eta} \left\| \boldsymbol{y} - \sum_i \left(\left(\mathsf{K}^{\mathrm{T}} \mathsf{K} \right)^{-1} \mathsf{K}^{\mathrm{T}} \boldsymbol{y} \right)_i \boldsymbol{f}_i(\boldsymbol{\eta}) \right\|^2 . \tag{3.66}$$

This NLS problem for $\boldsymbol{\eta}$ is independent of $\boldsymbol{\beta}$ and can be solved in the usual way by means of Gauss–Newton Eq. (3.57) or Levenberg–Marquardt algorithms Eq. (3.58). Although the Jacobian matrix for Eq. (3.66) is reduced in size, it still needs to be calculated for $\nabla \boldsymbol{F}(\boldsymbol{\eta})$ (*O'Leary and Rust*, 2013). Once the optimum $\boldsymbol{\eta}$ is found, the unique solution for the linear parameter vector $\boldsymbol{\beta}$ is obtained from Eq. (3.65).

3.3 Retrieval of CO from nadir measurements

In the SWIR spectral region the forward model $F(\boldsymbol{x}, \nu)$ for the upwelling monochromatic radiance for a double path through the atmosphere can be formulated as

$$F(\boldsymbol{x}, \nu) = \frac{r(\nu)}{\pi} \cos(\theta) \, E_{\mathrm{sun}}(\nu) \exp\left(-\sum_m \alpha_m \tau_m(\nu) \right) \otimes \mathcal{S}(\nu, \gamma) + b(\nu) , \tag{3.67}$$

where the state vector $\boldsymbol{x}$ includes the retrieval parameters (also see Eq. (2.20)). In this region of the electromagnetic spectrum the nadir viewing geometry does not allow to retrieve information on the vertical distribution of trace gases. The reason is that the information of the vertical profile is well under-determined in the observed spectrum because derivatives of the top of atmosphere radiance with respect to profile changes at different altitudes (weighting functions) do not peak at various altitudes for different wavenumbers (*Gimeno García et al.*, 2011, Fig. 1) (*Buchwitz et al.*, 2000, Sec. 3). Still, instruments observing in the TIR around 2100 cm^{-1} such as AIRS, IASI or TES (see Sec. 1.4) are able to provide vertical distributions of CO. In the SWIR, however, it is customary to retrieve total column densities N_{air}. Often, instead of total column densities according to Eq. (2.80) the total column averaged dry-air mole fractions q_m are given

$$q_m \equiv \frac{N_m(z_0)}{N_{\mathrm{air}}(z_0) - N_{\mathrm{H_2O}}(z_0)} \tag{3.68}$$

which represent the abundance of m relative to that of all other components $n_m = q_m \, N_{\mathrm{air}}$ and are less sensitive to variations in z_0 (Pub. I).

The subspace of the state vector $\boldsymbol{x}$ in Eq. (3.67) is defined by the set of retrieval parameters it holds

$$\boldsymbol{x} \in \mathbf{S} \subseteq \{ \boldsymbol{r}, \boldsymbol{b}, \boldsymbol{\alpha}, \gamma, \delta, \xi \dots \} . \tag{3.69}$$

The linear parameters $r_i \in \boldsymbol{r}$ and $b_j \in \boldsymbol{b}$ represent the coefficients for polynomials of optional order $i, j \leq 2$. The polynomials model the surface reflectivity and the baseline correction (optional), respectively. The scaling factors of the individual molecules m (CO, CH$_4$, H$_2$O, CO$_2$, ...) are represented by $\alpha_m \in \boldsymbol{\alpha}$. In general, molecules with absorption lines in the observed spectral interval need to be considered in the state vector. The instrumental slit function S, approximated by functions according to Sec. 2.7, includes the optional parameters γ for the half width, δ for the wavelength shift and ξ for the skewness of the spectral response.

In the retrieval of CO from TROPOMI's band 7, an exemplary state vector for the SLS fit could be chosen to include seven parameters in total, e. g.

$$\boldsymbol{x} \longrightarrow \left(\boldsymbol{\eta}, \boldsymbol{\beta}\right) \longleftrightarrow \left((\alpha_{\mathrm{CO}}, \alpha_{\mathrm{CH_4}}, \alpha_{\mathrm{H_2O}}, \delta)^{\mathrm{T}}, (r_0, r_1, r_2)^{\mathrm{T}}\right). \tag{3.70}$$

The wavenumber shift δ is often used to account for modifications in the spectral response due to heterogeneous illumination of the entrance slit of the instrument (details see Sec. 4.6). In comparison, the state vector for the SCIAMACHY CO retrievals in Pub. I and Pub. II include an additional parameter for the half width γ of the Gaussian response function because an ice layer on the detector modified the instrument's slit function (*Lichtenberg et al.*, 2010; *Gimeno García et al.*, 2011).

The forward model at wavenumber ν_j (with $j = 1, \ldots, M$) for the SLS is then given by

$$
\begin{aligned}
F_j(\boldsymbol{\eta}, \boldsymbol{\beta}) \;=\;& r_0 f_0(\boldsymbol{\eta}) \;+\; r_1 f_1(\boldsymbol{\eta}) \;+\; r_2 f_2(\boldsymbol{\eta}) \;=\; \sum_{i=0}^{2} \beta_i f_i(\boldsymbol{\eta}) \\
=\;& \frac{r_0}{\pi} \cos(\theta)\, I_{\mathrm{sun}}(\nu_j) \exp\left(-\sum_m \alpha_m \tau_m(\nu_j)\right) \otimes \mathcal{S}(\nu_j, \delta) \\
+\;& \frac{r_1}{\pi} \nu_j \cos(\theta)\, I_{\mathrm{sun}}(\nu_j) \exp\left(-\sum_m \alpha_m \tau_m(\nu_j)\right) \otimes \mathcal{S}(\nu_j, \delta) \\
+\;& \frac{r_2}{\pi} \nu_j^2 \cos(\theta)\, I_{\mathrm{sun}}(\nu_j) \exp\left(-\sum_m \alpha_m \tau_m(\nu_j)\right) \otimes \mathcal{S}(\nu_j, \delta)\,, \tag{3.71}
\end{aligned}
$$

and the corresponding row in the Jacobi matrix by

$$\boldsymbol{J}(\boldsymbol{\eta}) \;=\; \nabla_{\boldsymbol{\eta}} F_j(\boldsymbol{\eta}, \boldsymbol{\beta})\,. \tag{3.72}$$

3.4   Pre- and Postprocessing

A crucial aspect for the successful retrieval of state parameters (elements of the state vector) is pre- and postprocessing. This included the thorough preparation of Level-1b data, including the application of quality flags such as the bad and dead pixel mask (BDPM), rigorous cloud filtering, the removal of non-converged retrievals, and the disposal of measurements with very small SNRs (e. g., observations above large bodies of water, such as lakes, rivers, etc.).

Retrieved CO mole fractions for a subset of TROPOMI observations from various orbits around the globe are depicted in Fig. 3.3. The subset includes every tenth TROPOMI scanline and one scanline consists of 215 measurements. Figure 3.3a shows all converged retrievals while only retrievals with a cloud fraction of $< 10\,\%$ according to the S5P–NPPC (S5P–National Polar-orbiting Partnership Cloud) product (see Pub. III, Sec. 2.2.4) are shown in Fig. 3.3b. The majority of CO is distributed within $20 - 100\,\mathrm{ppbv}$. However, the figure clearly shows that values below $40\,\mathrm{ppbv}$ are primarily caused by retrievals over optically thick clouds since the bimodal distributions in the non-filtered cases vanish after cloud filtering. The vertical distribution of CO has a relative maximum in the troposphere (see Pub. III, Fig. 6), hence clouds that obscure most of the lower atmosphere cause the column averaged mole fraction to decrease.
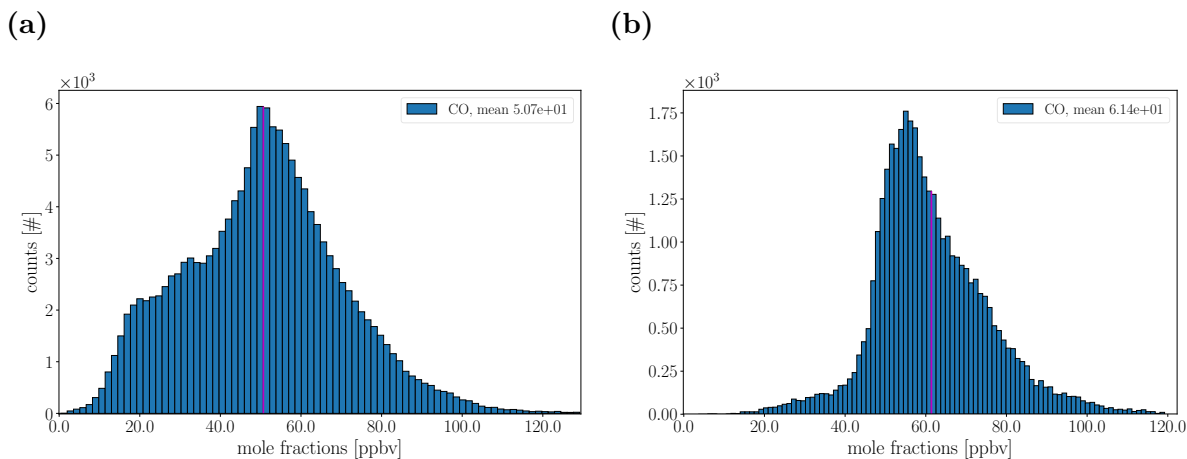
Figure 3.3: Inferred CO concentrations from TROPOMI measurements over Africa and Europe on September 21, 2019 (orbits 10045 − 10047). All converged observations are depicted in **(a)** while in **(b)** cloud filtered retrievals are shown.

3.5   Altitude sensitivity—averaging kernels

An aspect that is inherent to all remote sensed measurements but varies across different observing systems is their distinct altitude sensitivity. This feature is characterized by averaging kernels which in general need to be considered in a comparison between different observing systems (*Rodgers and Connor*, 2003).
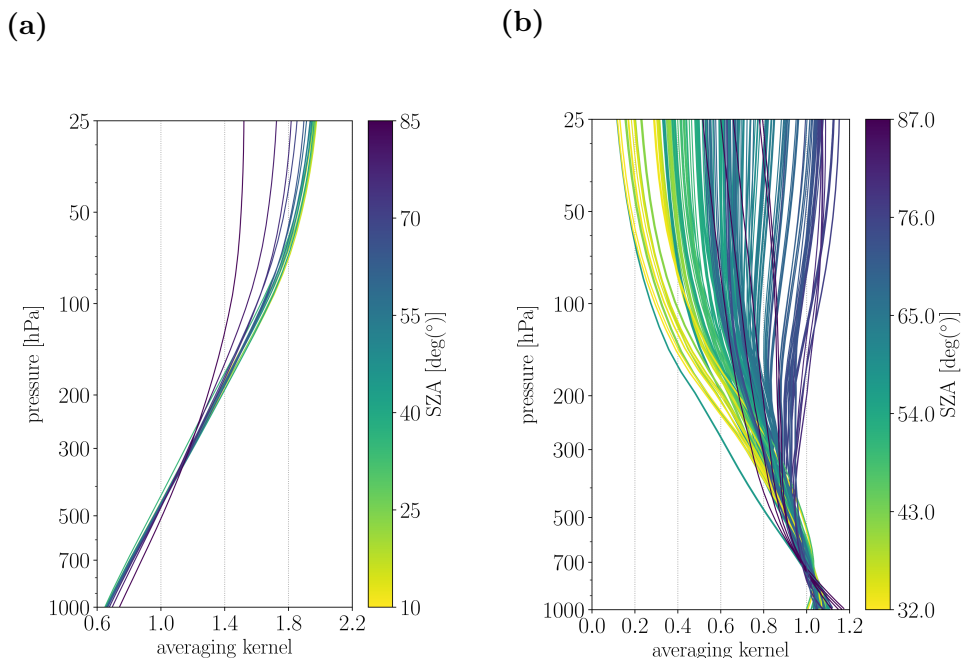
Figure 3.4: Averaging kernels for ground-based observations in Bremen. **(a)** The TCCON column averaging kernels for various solar zenith angles (SZA). **(b)** The averaging kernels for the NDACC site in Bremen.

The averaging kernel matrix $\mathsf{A} = \partial\boldsymbol{x}/\partial\hat{\boldsymbol{x}}$ relates the retrieved quantity $\boldsymbol{x}$ to the 'true' quantity $\hat{\boldsymbol{x}}$ and to any initial guess $\boldsymbol{x}_\mathrm{a}$ used in the retrieval by

$$\boldsymbol{x} - \boldsymbol{x}_\mathrm{a} \;=\; \mathsf{A}\left(\hat{\boldsymbol{x}} - \boldsymbol{x}_\mathrm{a}\right) + \boldsymbol{\epsilon_x} \,. \tag{3.73}$$

The error $\boldsymbol{\epsilon_x}$ includes both random and systematic errors in the measured signal and in the instrument's forward model. Since $\boldsymbol{x}$ describes an altitude profile of, e.g. CO, a row of A can be regarded as a smoothing function for a single altitude level. Hence, in order to avoid biases caused by different altitude resolutions and (likely different) linearization points $\boldsymbol{x}_\mathrm{a}$ averaging kernels need to be taken into account when comparing vertical profiles from different instruments or retrieval algorithms.

However, in case of column density retrievals the state vector is composed of profile scaling factors $\alpha_m \in \boldsymbol{x}$ (i.e. one scalar per molecule m) but not molecular concentrations at different altitude levels as in the profile retrieval above. Therefore, a so called column averaging kernel matrix ($\mathbf{C}$, *Buchwitz et al.*, 2004, Sec. 5) is used to describe the vertical sensitivity of the retrieval. In contrast to a profile retrieval, $\mathbf{C}$ is composed of vectors $\mathbf{c}$ each representing the column averaging kernel for one element α_m in the state vector. The column averaging kernel $\mathbf{c}$ describes the retrieval's response to a perturbation from the initial guess profile and deviations from unity can be interpreted as null space (Eq. (3.8)) or smoothing errors according to (*Rodgers and Connor*, 2003, Sec. 2 and 3). Therefore, in a comparison of total columns the equation

$$t_\mathrm{col} \;=\; \alpha N + \mathbf{c}^\mathrm{T}\left(\hat{\boldsymbol{d}} - \alpha\boldsymbol{d}_\mathrm{a}\right) \tag{3.74}$$

takes different altitude sensitivities from different observing systems into account. More specifically, t_col represents the vertically integrated $\boldsymbol{x}$ accounting for the sensitivity of the observing system. N represents the total column of the prior profile according to Eq. (2.80) and $\mathbf{c}$ is the vector containing the column averaging kernel for molecule m (also see *Wunch et al.*, 2010).
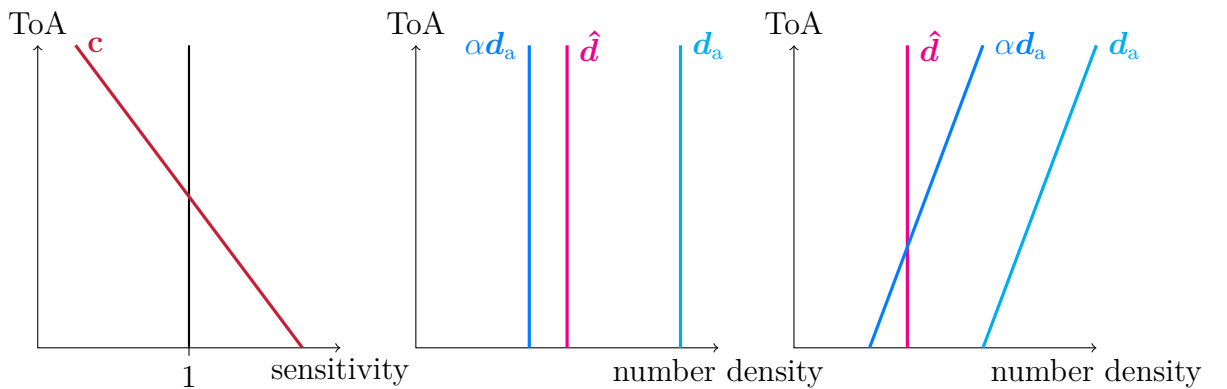
Figure 3.5: Schematic depiction of a column averaging kernel for an observing system with increased sensitivity towards the lower atmosphere (see *Buchwitz et al.*, 2004, Sec. 5) **(left)**. Exemplary retrieval outcome given that the initial guess concentration profile $\boldsymbol{d}_\mathrm{a}$ has the shape of the true profile **(center)** — the retrieved profile is a scaled version of the initial guess (prior profile). The initial guess' shape is different from the true profile **(right)**.

In Fig. 3.5 the effect of averaging kernels on total columns from different observing systems is depicted. If the initial guess profile x_a has the correct shape, proper application of the total column averaging kernel to the scaled profile $\alpha \boldsymbol{d}_a$ eliminates the null space error as the difference between $\alpha \boldsymbol{d}_a$ and $\hat{\boldsymbol{d}}$ is solely caused by the deviation of $\mathbf{c}$ from unity. For cases where shape of the initial guess is different from the true a null space error is introduced (see Eq. (3.8)). The difference of the scaled profile from the true profile will introduce another null space error which can not be corrected since the shape of the true atmospheric profile is in general not known. In the example the offset (bias) is only attributed to the systems gradient in vertical sensitivity and vanishes when $\mathbf{c}$ is taken into account. So in general t_{col} should be used for the comparison of, e.g., space-based and ground-based in order to account for these differences and correctly quantify the (possible) offset.

Chapter 4

Results

4.1 Publication I: Validation of CO total columns from SCIAMACHY

In Pub. I CO mole fractions inferred from SCIAMACHY $2.3\,\mu$m nadir observations from 2003–2011 using the BIRRA retrieval algorithm were validated against 18 stations of the ground-based networks TCCON (Total Carbon Column Observing Network) and NDACC (Network for the Detection of Atmospheric Composition Change). Weighted averages of SCIAMACHY CO observations within a circle around the g-b observing system were utilized to minimize effects due to spatial mismatch of space-based (s-b) and g-b observations. The global bias was determined to be in the order of -10 parts per billion in volume (ppbv) depending on the reference network and validation strategy used. The largest negative bias was found to occur in the northern mid-latitudes in Europe and North America. It was found that after postprocessing of the BIRRA retrieval output, the individual CO mole fractions inferred from SCIAMACHY still vary significantly between sites, ranging from around 100 ppbv up to 200 ppbv. The study also found that differences in vertical sensitivity between the BIRRA retrievals from SCIAMACHY and the ground-based retrievals from FTIR measurements are small allowing for a direct comparison.

Figure 4.1 shows the results for the comparison of the SCIAMACHY full-mission dataset with TCCON and NDACC ground-based observations. The rather large co-location radius of 500 km was required in order average over a sufficiently large ensemble of SCIAMACHY measurements, thereby reducing the noise. To mitigate representation errors however, the averages were weighted with increasing distance from the ground station. As shown in Pub. I this had a positive impact on the mean bias, i.e., reduced the representation error induced offset. The overall bias $\bar{b} = -12.1$ ppbv was determined as the average of all monthly-mean station biases weighted by their respective standard deviation (Pub. I, Sec. 2.4). The mean standard deviation across sites $\bar{\sigma}$ was found to be larger than the actual bias, hence the global bias $\bar{b}$ is not considered significant. Moreover, the per station bias turned out to be only significant at some sites, namely Kiruna (NDACC), Bialystok (TCCON), and Bremen (NDACC, TCCON).

(a)

(b)
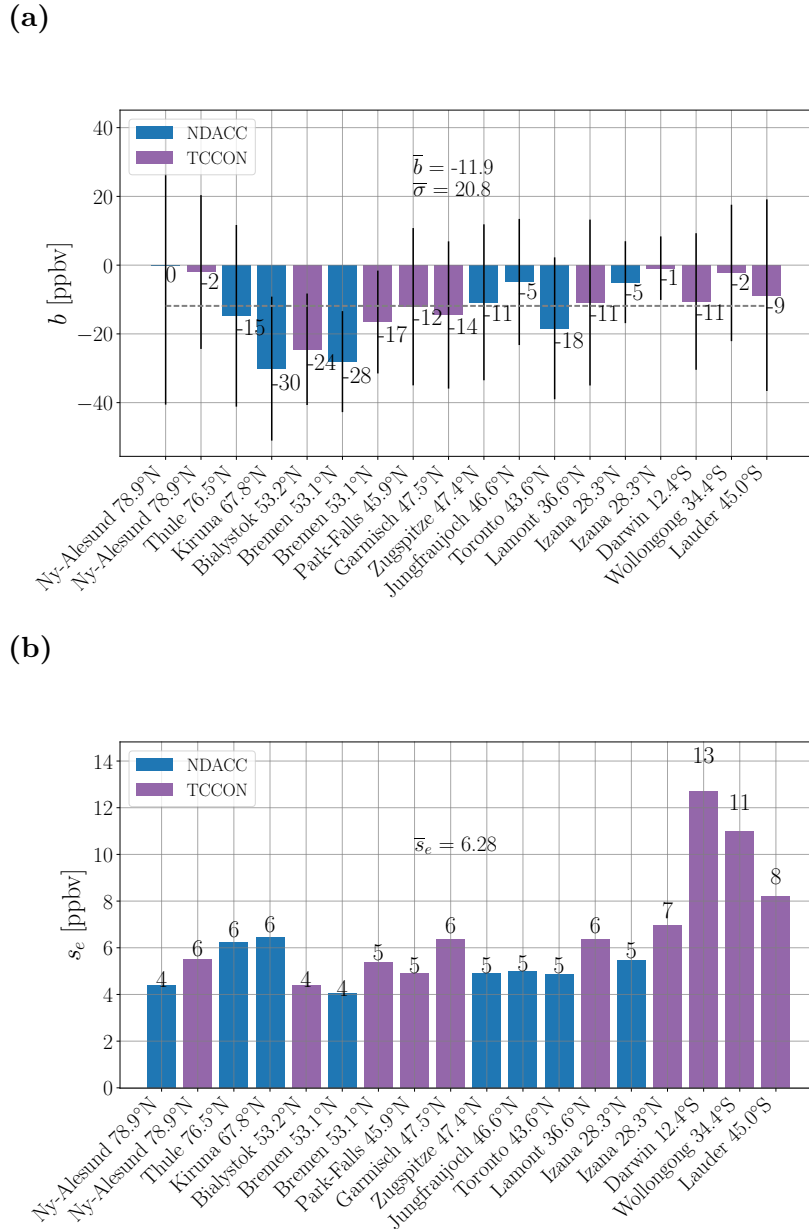
Figure 4.1: Observations within 500 km from the TCCON and NDACC ground stations were used for the calculation of monthly averages upon which the bias was determined. **(a)** Biases b together with their standard deviations on a monthly average bases for SCIAMACHY observations from 2003–2011. The SCIAMACHY CO concentrations were (de-)weighted inverse to the distance squared from the reference site. **b** The standard error of the mean values.

4.2 Averaging kernels

The decision to neglect the column averaging kernel in Pub. I was based on the analysis shown in Fig. 4.2. To quantify the effect of altitude sensitivity on the retrieved CO columns a true ($\boldsymbol{x}_{\text{true}}$) and a prior ($\boldsymbol{x}_{\text{prior}}$) profile were defined. Next, the BIRRA column averaging kernels C (see *Buchwitz et al.*, 2004, Sec. 5) for SCIAMACHY/TROPOMI were calculated for different observer zenith angles (OZA) and solar zenith angles (SZA). Note

that they are almost identical for both instruments given the same retrieval algorithm and spectral interval (see Pub. II, Fig. 8 and Pub. III, Fig. 7). Then a BIRRA retrieval was performed based on a synthetic observations for $\boldsymbol{x}_{\text{true}}$ (US-Standard profile) and a 'wrong' prior profile (initial guess) $\boldsymbol{x}_{\text{prior}}$. The retrieved profiles $\boldsymbol{x}_{\text{rtv}}$, scaled versions of the initial guess profile, were found to be $< 2\,\%$ or $\approx 5 \cdot 10^{16}\,(\text{molec}\,\text{cm}^{-2})$ within the true column value for all examined profiles. After application of the averaging kernels according to Eq. (3.74), the difference in columns was reduced by $\approx 1\,\%$ to that the null space error was quantified to be in the order of $1\,\%$ for CO retrievals from SCIAMACHY clear-sky observations. This is in good agreement with findings by *Borsdorff et al.* (2016, Fig. 2) which represents a minor contribution to the overall SCIAMACHY error budget can hence be neglected in a comparison. A direct comparison was therefore considered adequate.
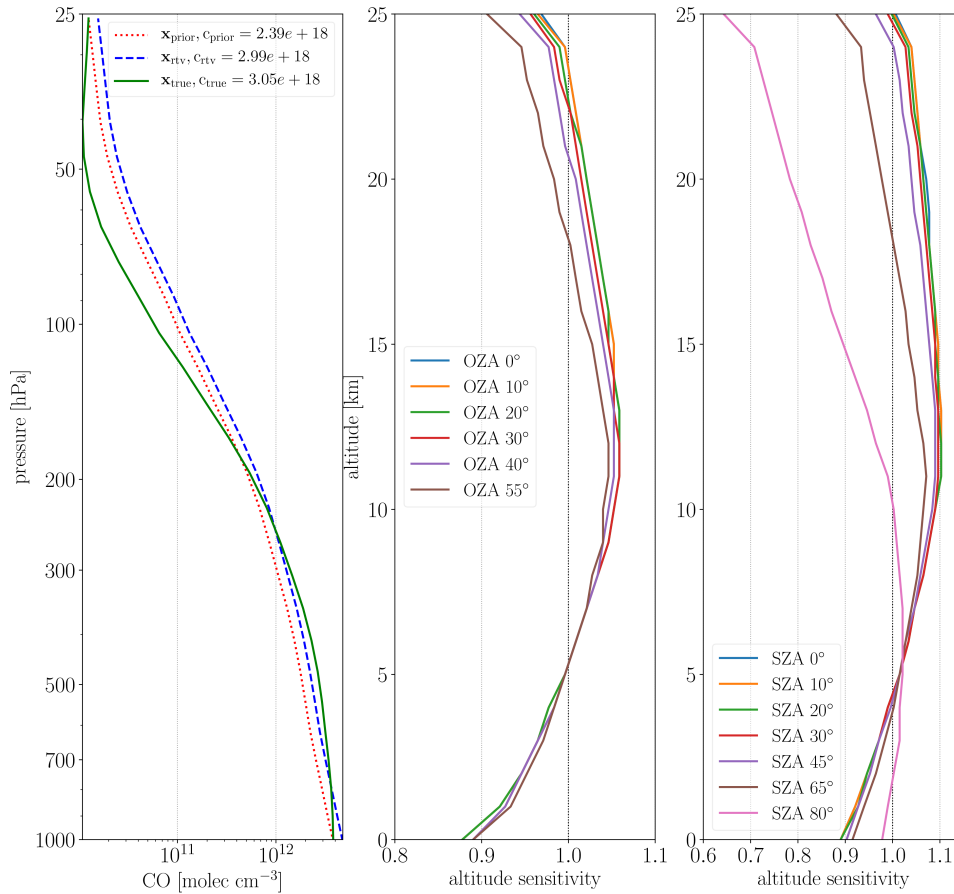
Figure 4.2: (Left) Three different vertical CO profiles. The true profile $\mathbf{x}_{\text{true}}$ was used to simulate the measurement while the retrieved profile $\mathbf{x}_{\text{rtv}}$ is a scaled version of the prior $\mathbf{x}_{\text{rtv}}$. The corresponding $\mathbf{c}_{\text{profile}}$ designates the respective total column. (Center, Right) Column averaging kernels for different SZAs and observer zenith angles (OZAs), respectively.

4.3 Retrieval algorithm upgrade

The development of the scientific BIRRA variant was (more or less) frozen at the status described in *Gimeno García et al.* (2011), however, the development of the GARLIC radiative transfer model was ongoing (by Franz Schreier). In early 2018, version 4.1 was the most current variant of the line-by-line code in the GIT version control system (*Loeliger*, 2010). The forward model that was implemented in BIRRA at that time was based on GARLIC version 2, so it was lagging two version numbers. In order to upgrade BIRRA with GARLIC version 4 major modifications in the BIRRA code were required since GARLIC version 4 incorporates multiple enhancements compared to its predecessor. This includes updated interpolation schemes, optimized convolution routines, the MT-CKD (Mlawer-Tobin-CKD, *Mlawer et al.*, 2012) and CIA (Collision-Induced-Absorption, *Karman et al.*, 2019; *Richard et al.*, 2012; *Borysow*, 2002) continuum absorption models (see Sec. 2.6), as well as multiple additional approximations to compute Rayleigh cross sections (see Sec. 2.5.2). In order to increase its efficiency in the line-by-line calculations, the latest GARLIC variant moreover incorporates changes in the data handling and order of calculations, particularly in the code sections for the computation of molecular absorption. In addition, changes were made with respect to the GARLIC user interface (i.e. FORTRAN (*Adams et al.*, 2008) namelist records).

The BIRRA upgrade was performed in small steps and accompanied by functionality tests of the relevant code segments in order to be able to track issues from the beginning. At this stage it was essentially a software engineering task where in particular subroutines and code blocks needed to be rearranged, new interfaces to be built or existing ones to be redefined, and variable declarations to be adapted. The update was finally completed in the third quarter of 2018 and the upgraded BIRRA variant was designated version 3.0. Although the update was complete from a software engineering point of view, essentially meaning that the program was compiling without errors, the physical results were yet to be examined and verified. In order to do so retrievals on a set of SCIAMACHY observations were performed with the old but validated BIRRA v2.0. The output was then compared to retrievals for the same subset but inferred with BIRRA v3.0. The final result is shown in Fig. 4.3a and 4.3b. It was found that the rather small discrepancies are primarily

(a)   **(b)**
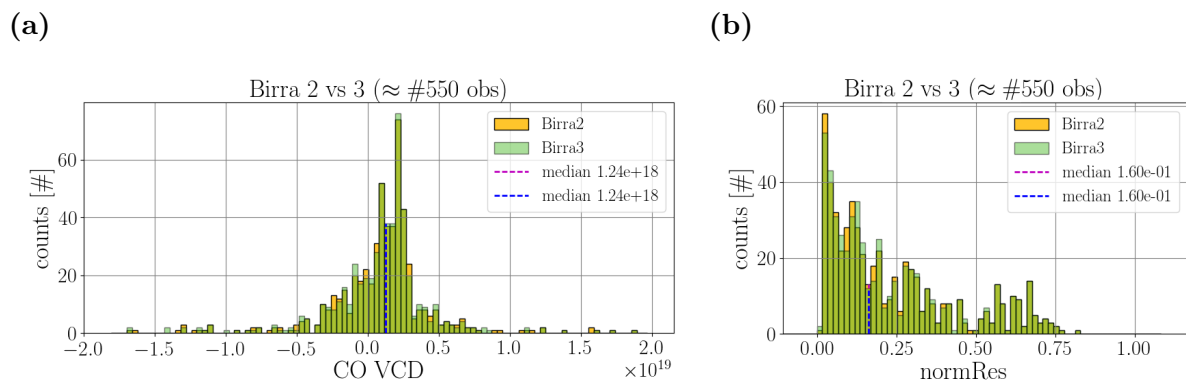
Figure 4.3: Comparison of retrieval results from BIRRA v2.0 and v3.0. **(a)** Carbon monoxide total columns for a subset of SCIAMACHY observations from orbit 13212 on September 9, 2004. **(b)** The corresponding spectral fitting residuals.

caused by different variable declarations in the new GARLIC version. More specifically, some floating point numbers were changed from single to double precision thereby causing numerical effects.

The successful upgrade was a crucial accomplishment for all of the subsequent studies since any deficiency incorporated during the code update could have lead to severely compromised results.

4.4 Publication II: Impact of molecular spectroscopy on CO from SCIAMACHY

In Pub. II the impact of SEOM–IAS spectroscopic information on CO columns from a large subset of SCIAMACHY measurements in 2003, 2004 and 2005 was examined. The well-established HITRAN 2016 and GEISA 2015 line lists were used as a reference upon which the impact was assessed.
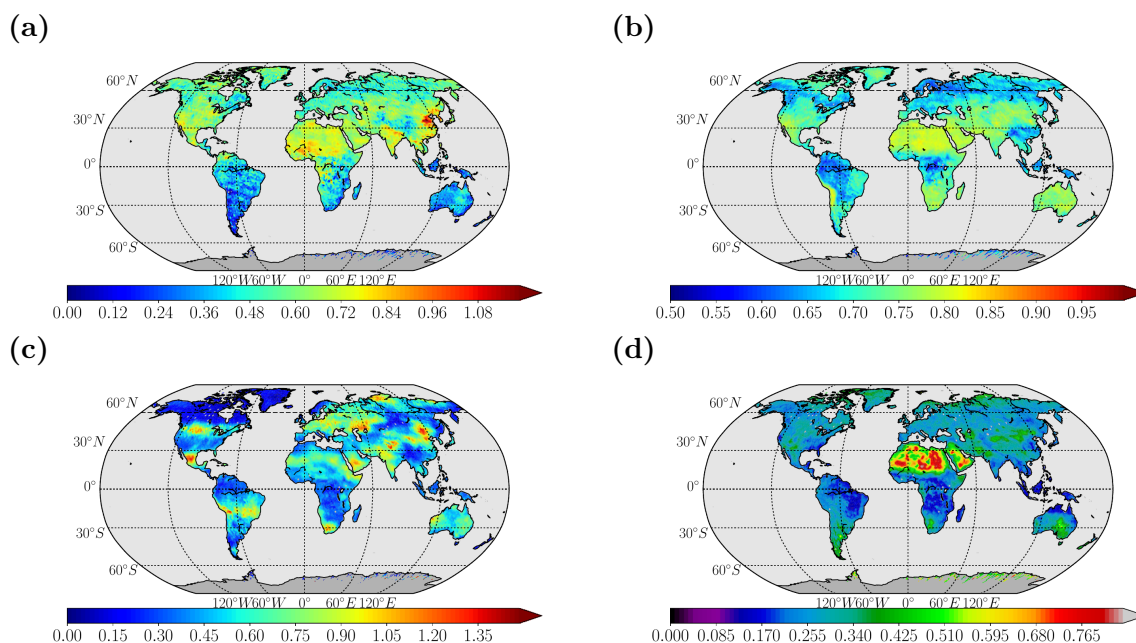
Figure 4.4: Retrieval output of BIRRA v3.0 for SCIAMACHY observations from the second quarter of 2004. The molecular absorption cross sections were calculated with SEOM–IAS line data according to the SDVM line profile. **(a)-(c)** The profile scaling factors of CO, CH_4 and H_2O, respectively. In **(d)** the effective reflectivity r_0 is shown.

It was found that the SEOM–IAS spectroscopy has positive impact on the spectral fitting residuals. This is attributed to both, the updated SEOM–IAS line data as well as more sophisticated line profiles that can be used with the extended set of SEOM–IAS line parameters. The largest effect, however, was found to be attributed to the updated SEOM–IAS line parameters ($\approx 3\%$ on average, up to 15% for individual observations) but not the models. Nonetheless, the best retrieval results for this enhanced spectroscopic dataset is obtained when higher-order effects in molecular absorption described by the so called beyond Voigt profiles are taken into account.

Beside the improvements in the spectral fit quality, the CO mole fractions increased by $4 - 11\%$ for SEOM–IAS spectroscopic information, thereby reducing the bias to both NDACC and TCCON ground-based observations. However, since SCIAMACHY observations show a rather high variability, this difference was found not significant for the majority of observations but only for some spectra over the Sahara with sufficient SNR Fig. 4.4. For those scenes an $\approx 3\%$ increase in the CO column was found.

In conclusion, the outcomes confirm recommendations from earlier investigations, e. g., by *Galli et al.* (2012) or *Checa-García et al.* (2015), i. e., trace gas retrievals in the SWIR will benefit from improved molecular spectroscopy. Although SEOM–IAS has been compiled to meet the accuracy requirements of new operational missions such as TROPOMI the findings suggest that the updated line data and models are beneficial for the retrieval of CO from SCIAMACHY. This is an important aspect, e. g., for the compilation of a multi-mission CO product.

4.5 Publication III: Impact of molecular spectroscopy on CO from TROPOMI

Publication III investigated the impact of the SEOM–IAS spectroscopy on CO mole fractions from TROPOMI SWIR observations.

Similar to Pub. II, it was found that SEOM–IAS line data with the adequate model improves the spectral fit quality by significantly reducing the residuals to TROPOMI measurements compared to both, HITRAN 2016 and GEISA 2015. The magnitude of the improvement varies across climatological regions but range from $\approx 10 - 20\%$ (see Fig. 4.5) and up to 30% for individual observations with respect to GEISA 2015. The improved fit quality was identified to be mainly caused by updates in the H_2O and CH_4 cross sections based on enhanced SEOM–IAS line data and models.

(a) **(b)**
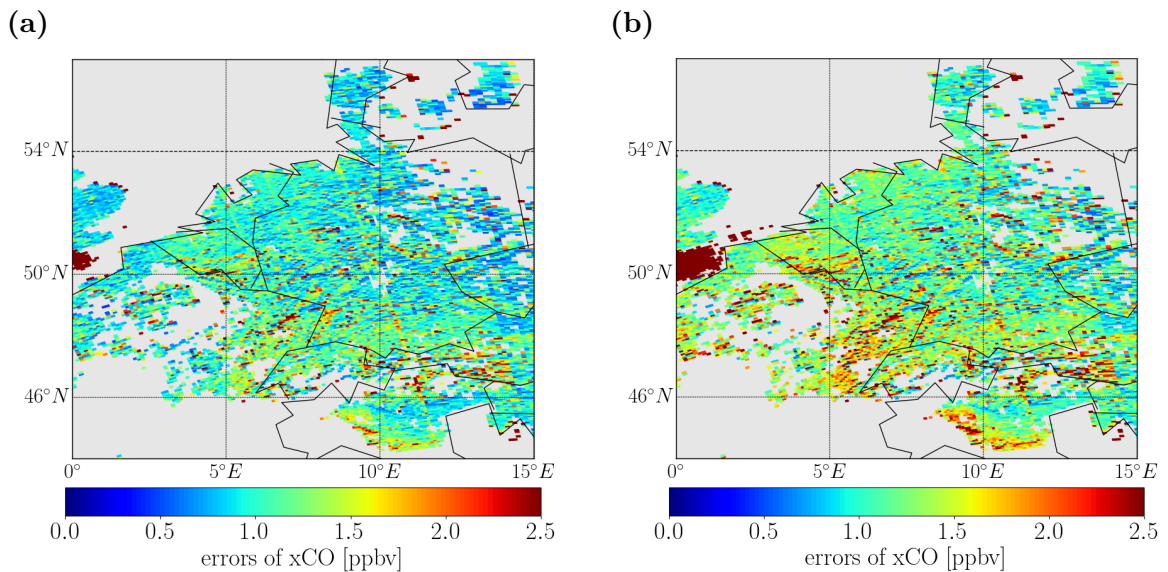
Figure 4.5: Comparison of the CO retrieval error for different molecular spectroscopies. **(a)** Retrieval errors for SEOM–IAS based SDRM cross sections. **(b)** The corresponding errors for HITRAN 2016 with Voigt cross sections.

In contrast to the fitting residuals, the differences in CO columns between SDRM and H16 were found to be rather small across most regions ($\leq 3\,\%$). The comparison to TCCON and NDACC revealed that the smaller retrieval errors in the SEOM–IAS inferred CO concentrations are beneficial when comparing post-processed mole fractions to ground-based references since stricter filter criteria can be applied on the TROPOMI observations within a given distance from the station.

Overall, many aspects of the findings underline recommendations from earlier investigations (*Galli et al.*, 2012; *Checa-García et al.*, 2015) and are in good agreement with findings from Publication II.

4.6   Instrument models and parameter fits

The accurate description of an instrument's spectral response is crucial for spectroscopic measurements (*Beirle et al.*, 2017). In contrast to tabulated response functions that are usually determined during on-ground calibration, parameterized instrument models in general allow for more flexibility in the Level-1b $\rightarrow$ 2 retrieval, e. g., account for small spectral misalignments between the modeled and the recorded spectra by a shift and squeeze (*Gimeno García et al.*, 2011; *Buchwitz et al.*, 2000).

The tabulated TROPOMI response functions as well as some classical instrument line shape parameterizations were presented in Sec. 2.7. In the subsequent paragraphs the tabulated TROPOMI response functions serve as the basis for the assessment of new parameterized instrument models. The latter part of the section examines the fit quality of the TROPOMI CO retrieval when instrument parameters are incorporated into the state vector (i. e., co-retrieved).

A simple method to introduce a flat-topped and skewed distribution is to linearly combine the Gaussian (see Eq. (2.89)) and super-Gaussian (see Eq. (2.93)) models according to

$$\mathcal{S}_{(2,4)}(\nu, \boldsymbol{x}) \;=\; A^* \, \mathcal{S}_2(\nu - \delta_2) \;+\; (1 - A^*) \, \mathcal{S}_4(\nu - \delta_4) \,, \tag{4.1}$$

such that $0 \leq A^* \leq 1$. The state vector for the ISRF fit could then be chosen as $\boldsymbol{x} = (A^*, \delta_2, \gamma_2, \delta_4, \gamma_4)$. The parameters γ and δ represent the respective widths and spectral shifts.

Another approach is to use the generalized normal distribution discussed in *Nadarajah* (2005) and combine it with the error function $\mathrm{erf}(\nu)$ in order to get a generalized skewed normal distribution according to

$$\mathcal{S}_{ks}(\nu, \boldsymbol{x}) \;=\; N^* \, \frac{k \, \sqrt[k]{\ln 2}}{2 \, \gamma \, \Gamma\!\left(\frac{1}{k}\right)} \, \exp\left(-\left|\frac{\nu}{\omega_k}\right|^k\right) \left\{1 \;+\; \mathrm{erf}\left(\xi \left|\frac{\nu}{\omega_k}\right|^{\frac{k}{2}}\right)\right\} . \tag{4.2}$$

The state vector could then be defined as $\boldsymbol{x} = (k, \gamma, \delta, \xi)$. The parameter ξ models the skewness and ω_k is given according to Eq. (2.92). The normalization factor N^* accounts for the finite integral (i. e. $N^* \neq 1$, see Eq. (2.87)).

The accuracy of the parameterized response functions is examined with respect to the tabulated (reference) ISRFs in Figs. 4.6 and 4.7. The figures depict least squares fits of some reference ISRFs via Eqs. (4.1) and (4.2).

While Fig. 4.6 shows the fit for various pixels across the spectral axis at detector pixel zero (left- or rightmost pixel), Fig. 4.7 shows the accuracy for different detector pixels at $4295\,\mathrm{cm}^{-1}$. The results for Eq. (4.1) across the spectral axis of TROPOMI's band

7 in Fig. 4.6a reveal a homogeneously distributed residuum within 3 % of the tabulated ISRF. In Fig. 4.6b the instrument's response is parameterized according to Eq. (4.2) and the variability of the residuum is more pronounced across the three different tabulated responses. Better results are attained for pixels at lower wavenumbers with the residuum increasing up to 6 % for detector pixels corresponding to higher wavenumbers. Largest discrepancies arise in the transition from the center to the wing of the distribution. However, the further away from the peak the disagreements occur the less critical they are for the convolution with the monochromatic spectrum as the spectral response becomes weaker.

The results for different detector pixels at $4295\,\mathrm{cm}^{-1}$ depicted in Figs. 4.7a and 4.7b reveal similar results for both parameterized response functions. The findings suggest that the TROPOMI ISRFs in band 7 could be replaced by the parameterized models without

(a)

(b)
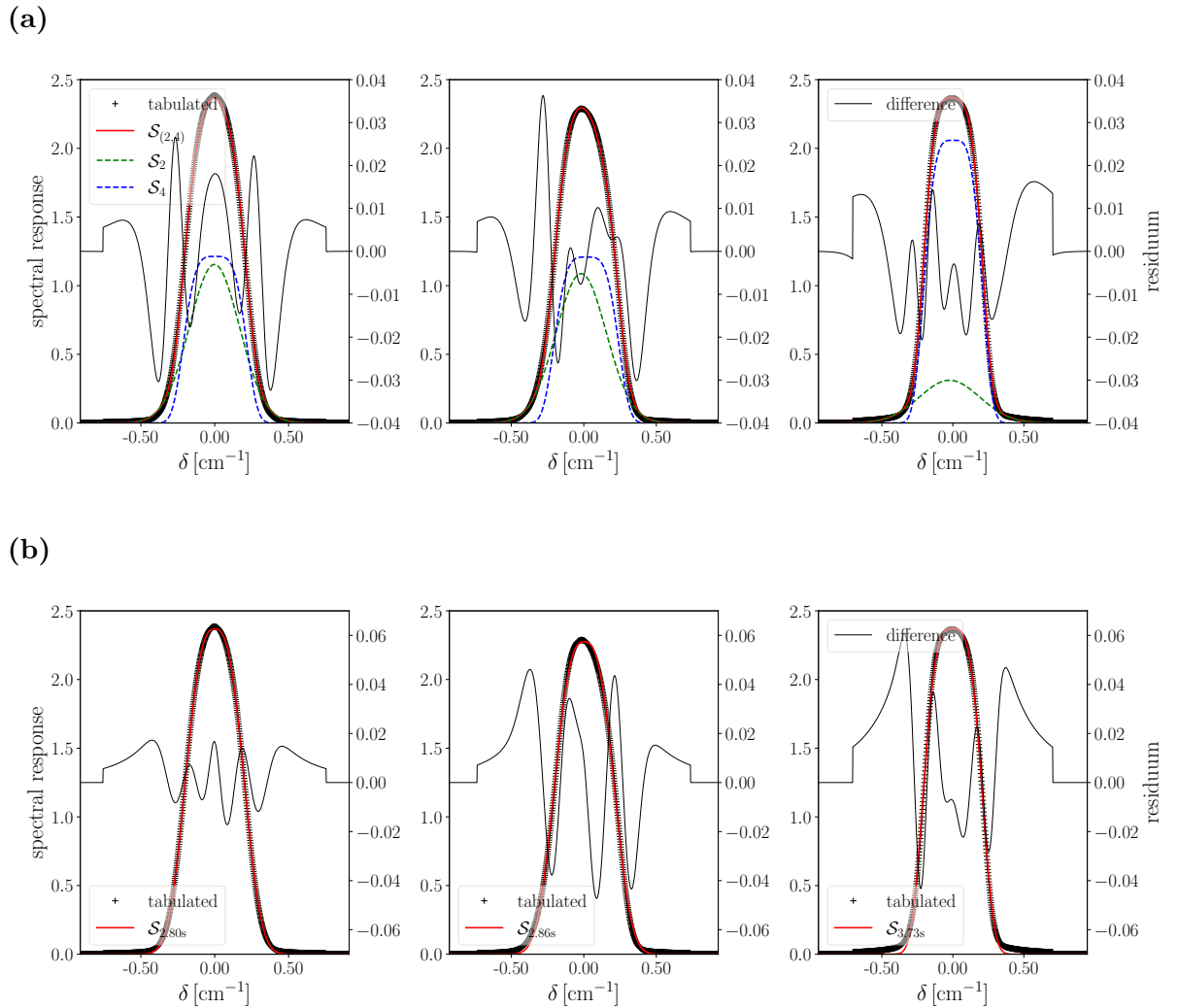
Figure 4.6: Fit of the tabulated TROPOMI spectral response functions for three spectral pixels ($4266.2\,\mathrm{cm}^{-1}$, $4306.6\,\mathrm{cm}^{-1}$, $4351.6\,\mathrm{cm}^{-1}$) of detector pixel 1. Two different parameterization schemes were used. In **(a)** linear combinations of Gauss and super-Gauss models $S_{(2,4)}$ were applied while in **(b)** different realizations of the skewed Gauss/super-Gauss variant $S_{k\mathrm{s}}$ were used.

inducing large spectral errors and hence still give accurate retrieval results.

Another investigation assessed the quality of the fit when the ISRF parameters of e. g. Eq. (4.2) are included in the state vector and co-retrieved with CO columns. As indicated in Fig. 4.8a, it turned out (as expected) that the spectral residuum $\boldsymbol{\rho}$

$$\boldsymbol{\rho} \;=\; \boldsymbol{y} \,-\, \boldsymbol{F}(\boldsymbol{x})\,, \tag{4.3}$$

becomes smaller the more ISRF parameters are included in the state vector. However, the incorporation of multiple additional fit parameters into the state vector makes the retrieval prone to overfitting.

The various state vectors also have major implications on the retrieved CO quantities as shown in Table 4.2. The variations were quantified with respect to the tabulated

(a)

(b)
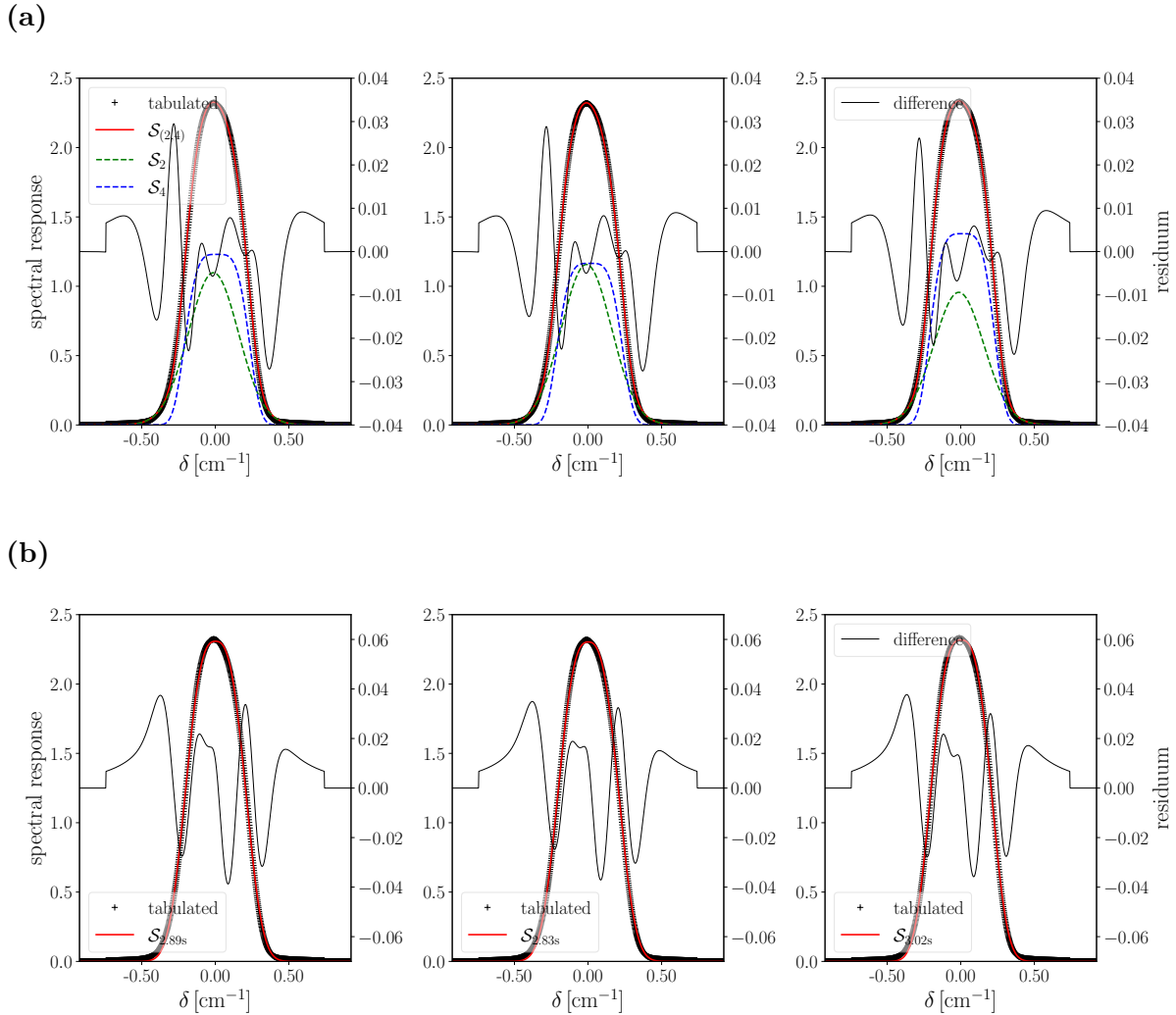
Figure 4.7: Same analysis as in Fig. 4.6 but for three different detector pixels (viewing angles) at $\approx 4295.5\,\mathrm{cm}^{-1}$. The left, center, and right subplots correspond to the detector pixels 1, 128, and 254, respectively. **(a)** Linear combinations of Gaussian response functions according to $S_{(2,4)}$ and **(b)** the response according to $S_{k\mathrm{s}}$.

Table 4.1: Goodness of fit for the TROPOMI observation from Fig. 4.8.

Parameters	Residuum (χ^2)
tabulated	$1.88\,10^{-2}$
$\{\gamma, \delta, \xi\}$	$3.84\,10^{-3}$
$\{\gamma, \xi\}$	$6.21(07)\,10^{-3}$
$\{\gamma, \delta\}$	$6.21(48)\,10^{-3}$
$\{\delta\}$	$2.20\,10^{-2}$

response. The differences in the CO columns were calculated according to

$$\Delta \text{CO} = \frac{\text{CO}_{(S_{2s})} - \text{CO}_{(S_{\text{tbl}})}}{\text{CO}_{(S_{\text{tbl}})}} \, , \tag{4.4}$$

where $\text{CO}_{(S_{2s})}$ and $\text{CO}_{(S_{\text{tbl}})}$ represent the retrieved CO concentrations using either the parameterized or tabulated ISRFs. The corresponding spectral residual with respect to the actual observations is given by

$$\chi^2 \; = \; \sum_{i=1}^{M} \frac{\boldsymbol{\rho}_i^2}{y_i} \, , \tag{4.5}$$

known as the goodness of fit statistic (*Hansen et al.*, 2013). The results in Table 4.2 demonstrate that the inferred columns vary considerably across the various retrieval setups. While most options reveal a positive bias towards to the tabulated ISRF setting, the CO columns with the $\{\gamma, \delta, \xi\}$ option is rather symmetrically distributed. In Table 4.1 the retrievals with the tabulated response functions exhibit a larger χ^2 than the fits with the parameterized ISRF—except for the shifted-only variant ($\{\delta\}$). Although the additional

(a)   **(b)**
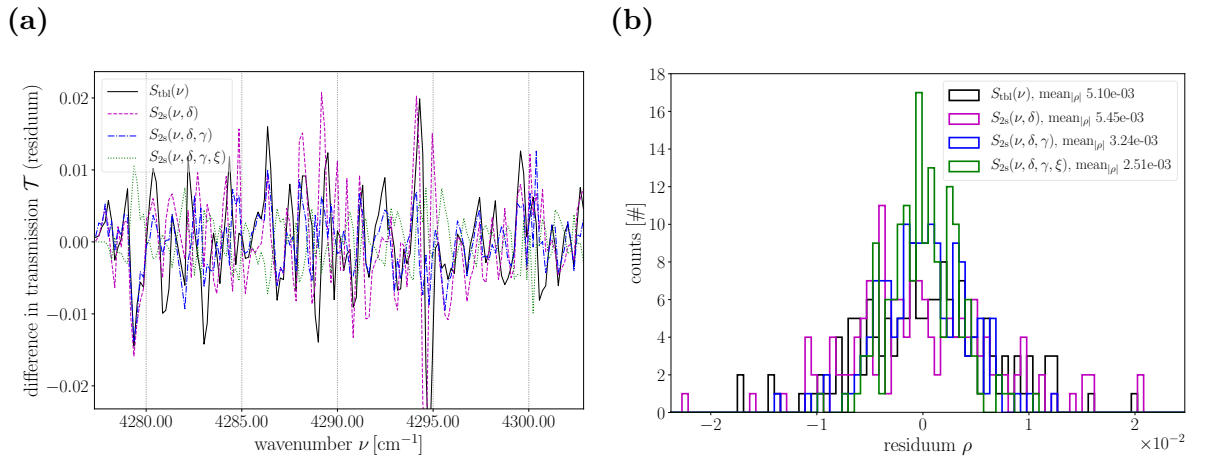
Figure 4.8: Spectral residuum for a randomly chosen TROPOMI observation computed with GEISA 2020 line data over the Sahara on May 07, 2018 (orbit 2923). The spectral shift of the $S_{2s}(\nu, \delta)$ model with respect to the tabulated ISRF can be clearly observed in **(a)** at around $4295\,\text{cm}^{-1}$. **(b)** The corresponding histogram of the spectral residuals. The residuals become smaller the more instrument parameters are included in the state vector.

Table 4.2: Percentage differences of retrieved CO total columns for various ISRF parameter-izations with respect to TROPOMI's tabulated spectral response. The retrievals were performed on a subset of TROPOMI observations over the Sahara on May 07, 2018.

	$\{\gamma, \delta, \xi\}$	$\{\gamma, \xi\}$	$\{\gamma, \delta\}$	$\{\delta\}$
min. (%)	-10.51	1.13	1.28	6.49
max. (%)	8.37	8.59	8.48	18.69
mean (%)	-1.87	3.50	3.68	10.55

state vector elements (fit variables) increase the goodness of the fit by minimizing the residuals, the risk of overfitting increases as well.

The parameterized instrument response function according to Eq. (4.2) (with $k = 4$) was used for the retrieval of CO across Africa in Fig. 4.9. Beside the CO total columns the corresponding errors and cloud fractions are shown. A survey of the spectral residuals over the Sahara revealed smaller discrepancies for the SEOM–IAS spectroscopy compared to HITRAN 2016. This is in good agreement with retrievals using the tabulated ISRFs from Pub. III. Moreover, this indicates that overfitting is not yet that much of an issue. Since the results for the parameterized ISRFs are promising it is proposed that a systematic comparison of TROPOMI retrievals for the tabulated and parameterized versions is conducted in the future.
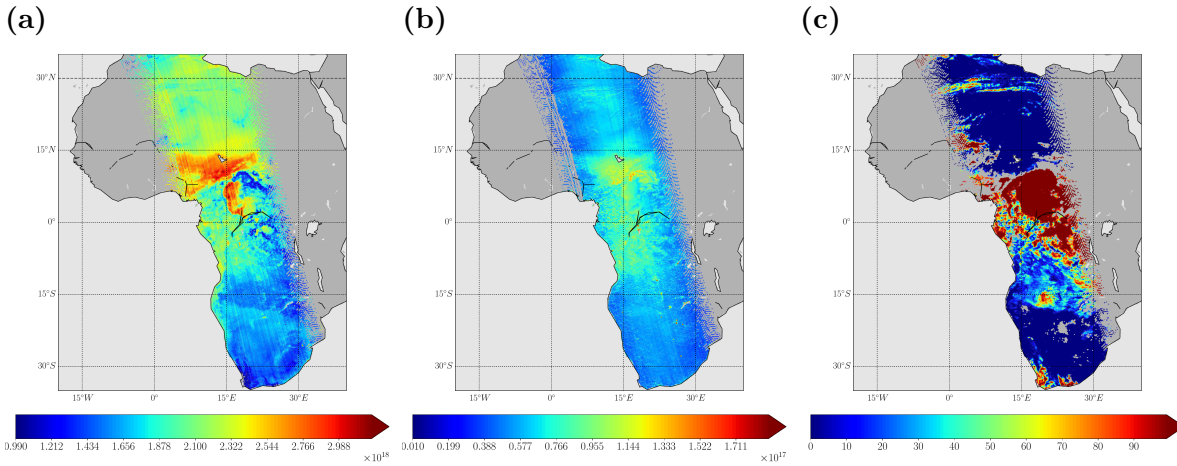
Figure 4.9: CO columns from TROPOMI measurements with the ISRF model $S_{4s}(\nu, \gamma, \delta, \xi)$. **(a)** CO columns in units (molec cm^{-2}) with **(b)** corresponding errors over Africa for a single overpass on May 07, 2018 (orbit 2924). **(c)** Corresponding cloud fractions, given in (%), were taken from the S5P–NPPC product. Cross sections were calculated with SEOM–IAS spectroscopic data and the SDVM line profile and US-Standard *Anderson et al.* (1986) initial guess profiles were used for CO, CH$_4$ and H$_2$O. Vertical temperature profile was taken from CIRA (COSPAR International Reference Atmosphere, *Fleming et al.* (1990)) climatology. Note the hemispheric gradient in CO concentrations.

4.7   Retrieval of aerosol parameters

The co-retrieval of aerosol parameters in the CO fit is based on the extinction of radiation by particles. Therefore, the aerosol optical depth τ_{aer} from Eq. (2.79) is treated as unknown by adding an amplitude scaling factor α_{aer0} and an exponent α_{aer1} to the elements of the state vector

$$\tau_{\mathrm{aer}}(\nu, \alpha_{\mathrm{aer0}}, \alpha_{\mathrm{aer1}}) \ = \ \alpha_{\mathrm{aer0}} \left(N_{\mathrm{air}}^{\downarrow} \ + \ N_{\mathrm{air}}^{\uparrow} \right) 8.833 \, 10^{-33} \, \nu^{\alpha_{\mathrm{aer1}}} \ . \tag{4.6}$$

The feasibility of fitting the extended state vector was assessed by employing the non-linear least squares (NLS) and the separable least squares (SLS) methods from Secs. 3.2.2 and 3.2.3, respectively. The performance of both algorithms was evaluated using TROPOMI observations.

First, the feasibility of the aerosol retrieval was estimated by assessing the Jacobian

(a)

(b)
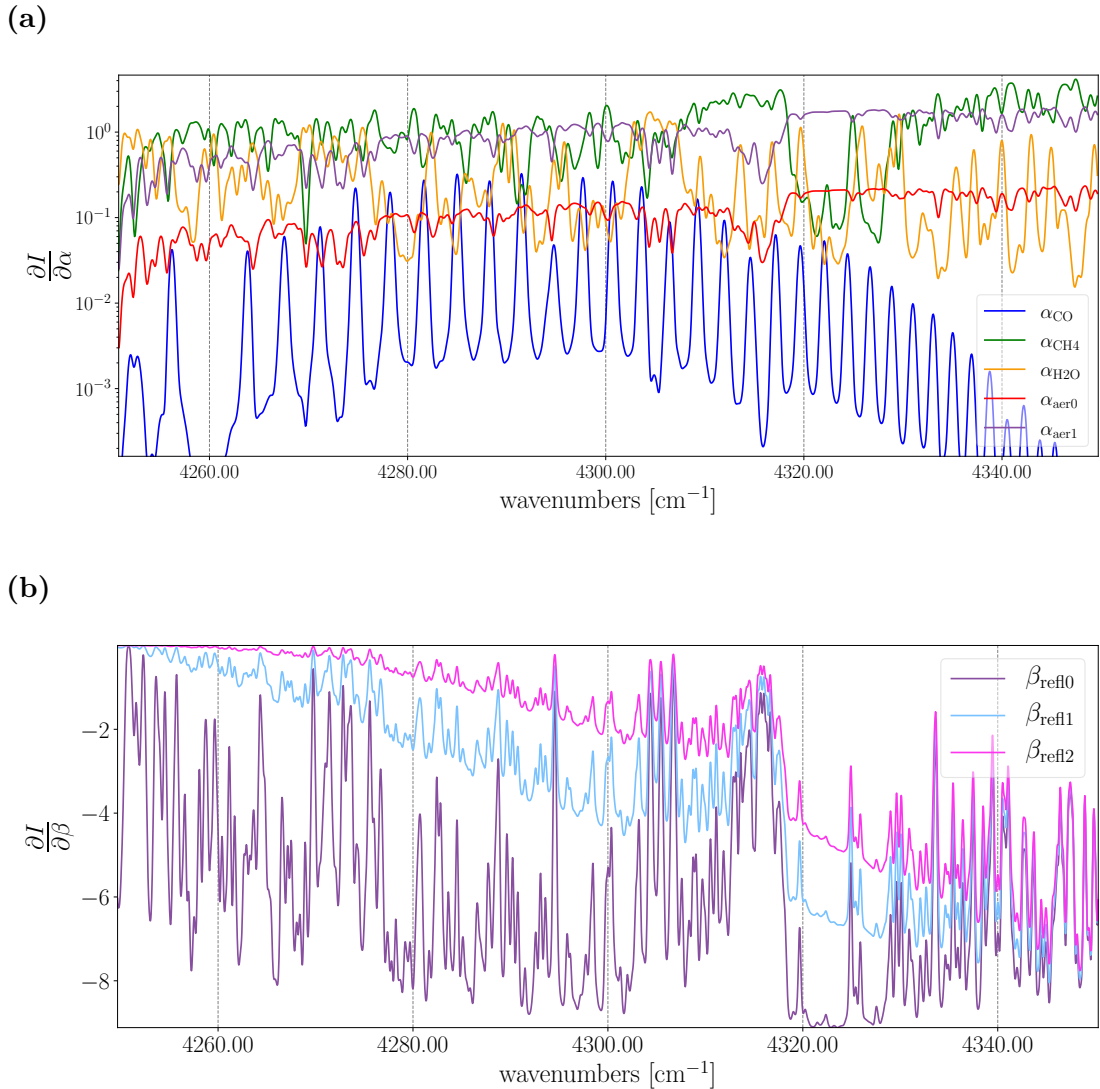
Figure 4.10: (a) Jacobians for the nonlinear molecular and aerosol parameters of the CO retrieval. **(b)** Jacobians for the linear parameters, i. e., the coefficients of the reflectivity polynomial.

matrix (see Eq. (3.55)), i.e. the change of the transmission with respect to retrieval parameters. The outcome is shown in Fig. 4.10, where Fig. 4.10a and Fig. 4.10b include the nonlinear and linear parameters, respectively.

In Fig. 4.10a Jacobians for the molecular scaling factors and aerosol parameters α_{aer0} and α_{aer1} are depicted. The Jacobi matrix was calculated for a double path through the atmosphere in nadir viewing geometry at ToA and for a SZA of 30°. A spectral resolution (FWHM) of 0.25 cm^{-1} was assumed. The derivatives for the aerosol parameters appear to be similar across the spectral range and only differ by a factor $\ln(\nu) \cdot \alpha_{\text{aer0}}$. As a result, the two columns of the Jacobi matrix are close to linear dependence making the inversion of both parameters very sensitive to small perturbations (see Secs. 3.1.1 and 3.1.3). Consequently, as indicated by the very large condition numbers in Fig. 4.11, the matrix represents a very ill-conditioned system when both aerosol parameters are elements in the state vector (see yellow bars).

The top-left of Fig. 4.11 shows various sets for the state vector $\boldsymbol{x}$ of the SLS algorithm. Note that the sets only include nonlinear parameters since the algorithm does not require any linear parameter for the iterative fit (see Sec. 3.2.3). This keeps the condition numbers low, except for the case where both aerosol parameters are included. Interestingly, the inversion with respect to the aerosol exponent α_{aer1} appears to be slightly better conditioned than the inversion with respect to α_{aer0}. This suggests that at least one parameter should be possible to fit along with the molecular scaling factors.
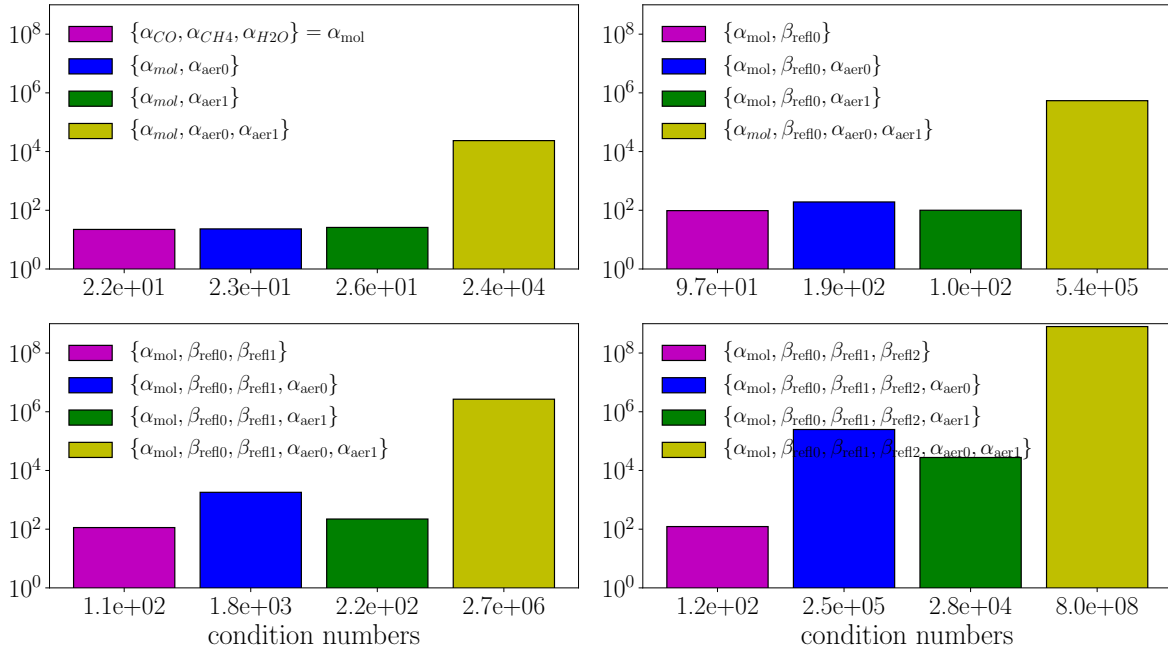
Figure 4.11: Condition numbers for Jacobian matrices that hold different parameter ensembles relevant for the CO retrieval in the SWIR. The group in the top left shows the condition numbers for the nonlinear retrieval parameters α which is hence relevant for the SLS algorithm. The groups in the other figures include the linear parameters β in addition to α which are thus relevant for the NLS solver. As indicated, the condition numbers for the SLS Jacobians in the top left plot are at least (approximately) one order of magnitude better compared to the rest (NLS Jacobians).

Table 4.3: Comparison of 4730 BIRRA CO retrievals over Europe for state vector elements $\{\alpha_{\mathrm{mol}}, \alpha_{\mathrm{aer0}}, \beta_{\mathrm{refl0}}\}$ (center column) and $\{\alpha_{\mathrm{mol}}, \alpha_{\mathrm{aer0}}, \beta_{\mathrm{refl0}}, \beta_{\mathrm{refl1}}, \beta_{\mathrm{refl2}}\}$ (right column) using two different iterative least squares solvers. The SLSB and NLSB fits use the SLS and NLS algorithms with (upper and lower) bounds, respectively.

solver	converged (%)	converged (%)
SLSB	99.0	99.1
NLSB	99.0	73.4

The other three subplots of Fig. 4.11 show different sets for $\boldsymbol{x}$ for the NLS solver. The condition numbers of the respective Jacobians are worse compared to the SLS subplot in the top-left. This is because the algorithm requires derivatives for all elements of $\boldsymbol{x}$ as all fit parameters are estimated in an iterative approach.

The significantly higher condition numbers for the Jacobians in the classical NLS fit asked for a comparison of the SLS and NLS algorithms. A first assessment was conducted in Table 4.3 using the bounded variants of both algorithms, i. e. NLSB and SLSB. A subset with 4730 TROPOMI observations over Europe was selected and the number of converged fits was compared. Strict cloud filtering criteria according to Pub. III (Sec. 2.2.4) were applied. Retrievals were calculated for two different state vectors, i. e. the surface reflectivity was modeled by polynomials of different degrees (a constant and second order reflectivity polynomial). The bounds for α_{aer0} were set to 0 and 10. The GEISA 2020 spectroscopic line list was employed to calculate molecular absorption and positivity constraints (bounds) were enforced on the molecular scaling factors. The result shows that both solvers perform equal for the fits with the constant reflectivity while the SLSB outperforms the NLSB when the reflectivity is modeled by the higher order polynomial.

In Fig. 4.12a the CO mole fractions with the co-retrieved α_{aer1} is depicted. The differ-

(a)   **(b)**
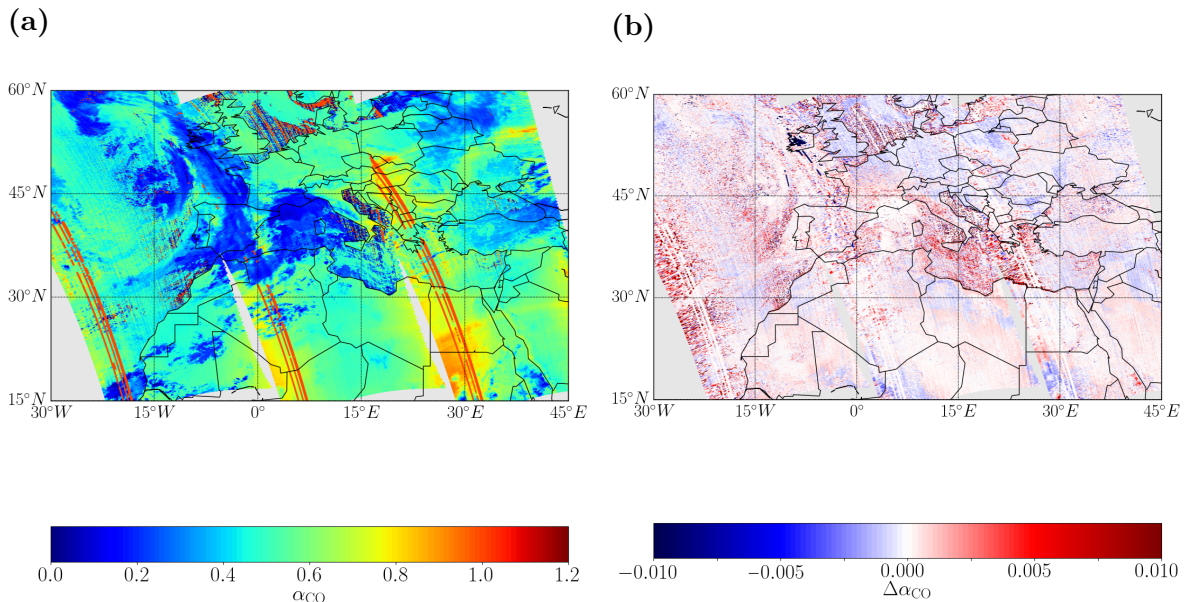
Figure 4.12: **(a)** CO scaling factors over Europe on September 21, 2019 with the α_{aer1} fit enabled in SLS solver. **(b)** The difference in CO when either α_{aer0} or α_{aer1} is co-retrieved is insignificant.

ence of CO for either co-retrieving α_{aer0} or α_{aer1} (respectively) is shown in Fig. 4.12b. The SLS solver was used in both setups and no cloud filter was applied this time. Although the retrieval converges for most observations, the fit of either aerosol parameter has an adverse impact on β_{refl0} ($\gg 1$ for most cases). The results indicate issues due to degeneracies since the distribution of the retrieved aerosol parameter does hardly resemble a physically reasonable distribution and neither least squares solver converges at physically meaningful aerosol values and reflectivity parameters. However, the estimates of either aerosol parameter seems to only have a minor effect on the final CO columns in the SLS fit.

The results presented in this section are not yet considered conclusive, however, they constitute an important step towards an algorithm that considers extinction by particles in addition to molecular absorption. Upcoming studies should also examine other aerosol parameterization schemes for the SWIR (e. g. *Reuter et al.*, 2017).

Chapter 5

Conclusions and Outlook

5.1 Conclusions

The topic of this thesis is the investigation of CO retrievals from SWIR nadir observations by spaceborne observations using the BIRRA code. An assessment and validation of the BIRRA algorithm was performed. This task embarked on the science question (1) in Sec. 1.6 and the outcome along with a detailed description of the methodology is presented in Sec. 4.1 and Pub. I. The inferred CO concentrations from SCIAMACHY SWIR measurements from 2003–2011 revealed to be largely consistent with similar validation studies from other authors using different algorithms. Additional results regarding the averaging kernels are presented in Sec. 4.2 describe the BIRRA column averaging kernels and the decision making process for a direct comparison.

Although the BIRRA software development was not an objective addressed in Sec. 1.6 the algorithm upgrade and its successful verification described in Sec. 4.3 was a crucial aspect for the upcoming studies and prepared the stage for new enhancements in the retrieval algorithm and its forward model.

Next, BIRRA's capabilities were improved in order to accurately retrieve CO from latest missions. A major effort was put in modeling higher order effects 'beyond Voigt' in molecular absorption using latest spectroscopic data. The enhanced physical description of molecular absorption along with updated line data was studied for SCIAMACHY (Sec. 4.4 and Pub. II) and TROPOMI (Sec. 4.5 and Pub. III) measurements and the outcome was compared against TCCON and NDACC ground-based observations. It was found that the new SEOM–IAS line data together with the adequate line profile improves the spectral fit quality by significantly reducing the spectral fitting residuals compared to both, HITRAN 2016 and GEISA 2015. The differences in the retrieved CO concentrations were found to be rather small. However, SEOM-IAS spectroscopy significantly reduces the retrieval error and enhances the precision of the CO product and should therefore be the preferred spectroscopic input for the retrieval of CO in the SWIR.

Further studies that address the objectives (3) and (4) were conducted but (not yet) published in articles of peer-reviewed journals. The studies on instrument line shapes in Sec. 4.6 focus on objective (3). The assessment of the proposed instrument line shapes revealed that both are adequate in fitting the range of tabulated TROPOMI response functions. Moreover, the positive impact of the SEOM–IAS spectroscopy on retrieved CO concentrations was also observed with a proposed parameterized response function.

The retrieval of aerosol parameters presented in Sec. 4.7 deals with objective (4). The

results show that the simple model for aerosol extinction is not sufficient to retrieve a reliable aerosol distribution across scenes. A survey of the Jacobian matrices condition numbers revealed the SLS fit outperforms the NLS fit for the co-retrieval of a single aerosol parameter (i. e. either the aerosol scaling factor or exponent). It was found that the significant higher number of converged fits for the SLS algorithm is attributed to the fact that the derivatives of the nonlinear aerosol parameters are similar to the derivatives of the linear reflectivity coefficients which cause degeneracies and lead to ill-conditioned Jacobians for the NLS fit. However, neither NLS nor SLS converges at physically meaningful values for the reflectivity and aerosol parameters—this is particularly true without setting constraints (bounds). Therefore, more investigations on the co-retrieval of aerosol parameters is proposed before it can be routinely employed in the BIRRA CO retrieval.

5.2   Outlook

It was shown that BIRRA is successful in retrieving CO from spaceborne SWIR measurements and that it is a flexible and versatile tool for a wide range of studies. The enhancements in the BIRRA algorithm yield an improved CO product across instruments. For strictly cloud filtered scenes the BIRRA results show good agreement with both TCCON and NDACC. In order to more accurately account for light path modifications by aerosols or cirrus clouds in the retrieval, investigations on co-retrieved effective aerosol parameters should be intensified.

The developments, in particular the improvements in molecular spectroscopy, are considered a crucial aspect in the algorithm's capabilities to successfully infer concentrations of other atmospheric constituents such as CO_2 and CH_4. Atmospheric concentrations for both are increasing (*Ekwurzel et al.*, 2017; *Yue and Gao*, 2018; *Saunois et al.*, 2020) causing radiative forcing that has the potential to trigger significant changes in climate (*Etminan et al.*, 2016). The majority of anthopogenic CO_2 emissions are concentrated on a small fraction of the globe, primarily on cities and power plants (*Nassar et al.*, 2017; *Kuhlmann et al.*, 2019; *Strandgren et al.*, 2020). The Paris climate agreement sets ambitious goals to reduce CO_2 emissions in order to limit global warming well below 2°C (*United Nations Framework Convention on Climate*, 2015, Article 2). The agreement envisages measures that allow for independent monitoring and verification of nationally reported anthropogenic CO_2 emissions. This includes periodically (5-yearly) global stocktakes for which space-based CO_2 measurements on the local and regional scale should provide input in order to meet the 'monitoring, verification and reporting' requirements.

The NASA mission OCO-2 was the first to demonstrate the potential of detecting and estimating strong CO_2 plumes of megacities and large point sources (*Nassar et al.*, 2017; *Reuter et al.*, 2019). A Copernicus mission that is dedicated to measure the global distribution of CO_2 as well as emissions from large power plants ($> 10\,\mathrm{MtCO2yr}^{-1}$) is CO2M (CO2 Monitoring mission, (*Kuhlmann et al.*, 2019)). There are also proposals for satellite instruments that target localized CO_2 emissions from medium-sized power plants (e. g. $1-10\,\mathrm{MtCO2yr}^{-1}$ *Strandgren et al.*, 2020). Private companies such as GHGSat put their resources on the detection of CH_4 point sources from space (*Cusworth et al.*, 2019; *Jervis et al.*, 2021).

A major challenge for the accurate identification of CO_2 and CH_4 enhancements is scattering by aerosols and cirrus clouds since both modify the light path of the measured back scattered radiation leading to an inaccurate estimation of the true concentration if

not appropriately accounted for (*Aben et al.*, 2007; *Hu et al.*, 2016; *Kuhlmann et al.*, 2019; *Lorente et al.*, 2021). There are basically two ways to account for scattering discussed in the literature. One is to simultaneously retrieve molecular concentrations and particle scattering properties of the atmosphere which is known as a physics-based retrieval (*Butz et al.*, 2009, 2011; *Wu et al.*, 2018, 2019; *O'Dell et al.*, 2018). Beside the physics-based method, the proxy approach proposed by *Frankenberg et al.* (2005) is also an option to account for scattering effects. However, because of different light path sensitivities in different spectral ranges the information of light path modifications from the NIR is often of limited value for the SWIR. A merged fit window approach is often used to enable the transfer of information between separated bands (*Reuter et al.*, 2010; *Gimeno García et al.*, 2011).

Based on those classifications the investigation of aerosol extinction in Sec. 4.7 can be considered a very first step towards a physical-based approach since it retrieves effective aerosol parameters. However, it is important to note that, similar to CH_4 and H_2O in the retrieval of CO (see Pub. III, Sec. 3.2), the co-retrieved effective aerosol parameters are considered a byproducts that is only retrieved to enhance the accuracy of the target species by compensating light path modifications. A study by *Schepers et al.* (2012) compared the performance of the physics- and proxy-based CH_4 retrievals from GOSAT observations and concluded that the two methods are similar in their performance to account for such modifications. It is therefore reasonable to consider both methods in the search for an appropriate description to account for aerosols and clouds in the retrieval of trace gas concentrations.

List of Publications and Contributions

Parts of the developments and results presented in this thesis were published in peer-reviewed articles that are attached in Pub. I, Pub. II, and Pub. III. This section provides an overview of all relevant publications where I am either first author or co-author. Moreover, the scientific context of each paper as well as the contributions from my side are outlined. While an overview of my three first author publications was presented in Secs. 4.1, 4.4, and 4.5, respectively, the co-authored publications are only mentioned here.

Hochstaffl, P., F. Schreier, G. Lichtenberg, and S. Gimeno García, 2018: Validation of carbon monoxide total column retrievals from SCIAMACHY observations with NDACC/TCCON ground-based measurements. *Remote Sens.*, 10 (2), 223, doi: 10.3390/rs10020223.

> **Context:** A dedicated verification and validation with ground-based remote sensing observations was regarded a crucial aspect in assessing the quality of the scientific BIRRA code (*Gimeno García et al.*, 2011). A set of validation tools initially developed by *Hochstaffl* (2015) needed to be extended and refined in order to come up with a dependable and consistent validation across sites and observing systems. Moreover, a thorough analysis of the column averaging kernels for both, SCIAMACHY and ground-based observations was required, in particular since space-based observations differ in viewing geometry as well as the spectral interval (most notably with respect to NDACC). The validation tools for the intercomparison of the three observing systems were written in Python.
>
> **My contributions:** Developed the methods, tools, and strategy for this validation and also performed the retrievals and the intercomparisons. I wrote the manuscript that was then reviewed and commented by the coauthors.

Hochstaffl, P. and F. Schreier, 2020: Impact of molecular spectroscopy on carbon monoxide abundances from SCIAMACHY. *Remote Sens.*, 12 (7), 1084, doi: 10.3390/rs12071084.

> **Context:** A first study regarding the impact of spectroscopic data from HITRAN 2012 and GEISA 2011 on BIRRA retrieved CO from SCIAMACHY has been conducted by *Schmidt* (2014). A new SWIR spectroscopy was compiled at DLR as part of the ESA project SEOM–IAS. The new dataset was of particular interest as it was not yet validated with space-based observations (M. Birk, G. Wagner and F. Schreier, personal communications). Since latest spectroscopic datasets

such as SEOM–IAS use line profiles beyond Voigt new line models needed to be implemented in the radiative transfer code. In order to utilize latest spectroscopic data for line-by-line calculations in BIRRA more sophisticated line profiles (beyond Voigt) had to be implemented its forward model (GARLIC). Technically, this was achieved by using Fortran 2003 type-bound procedures which allowed to bind the corresponding line data and models as attributes and procedures to dedicated derived types. The tools to quantify the differences in the fit quality for various spectroscopies and assess the quality of the retrievals were developed in Python.

My contributions: Developed the methods, tools, and strategy for this study and performed all retrievals. I wrote the manuscript that was then reviewed and commented by the coauthor.

Hochstaffl, P., F. Schreier, M. Birk, G. Wagner, G. D. Feist, J. Notholt, R. Sussmann and Y. Té, 2020: Impact of molecular spectroscopy on carbon monoxide abundances from TROPOMI. *Remote Sens.*, 12 (21), 3486, doi: [10.3390/rs12213486](10.3390/rs12213486).

Context: Beside the updated physical description of molecular absorption, the retrieval algorithm itself as well as the processor framework required to be upgraded and modified in order to facilitate TROPOMI SWIR observations for the retrieval of CO with BIRRA. The capability to use tabulated ISRFs was built and TROPOMI specific data ingestion and management routines written. This included the handling S5P Level-1b Earth radiance and solar irradiance files, application of calibration key data (CKD) such as quality flags and the bad and dead pixel mask. Moreover, a digital elevation model with higher resolution as well as more accurate initial guess data on CO and CH_4 vertical distributions were incorporated and the processor framework upgraded from Python 2 to 3.

My contributions: Developed the methods, tools, framework and strategy for this study and performed all retrievals. I wrote the manuscript that was then reviewed and commented by most of the coauthors.

Schreier, F., S. Gimeno García, **P. Hochstaffl** and S. Städt, 2019: Py4CAtS— PYthon for Computational ATmospheric Spectroscopy. *Atmosphere*, 10 (5), 262, doi: 10.3390/atmos10050262.

Context: Py4CAtS — PYthon scripts for Computational ATmospheric Spectroscopy is a Python re-implementation of the Fortran based infrared radiative transfer code GARLIC. As outlined in the paper, the original intention was to call compute intensive subroutines of GARLIC from Python but thanks to the increased performance of NumPy a Python implementation of lbl cross sections became feasible, i. e., the interface to GARLIC's subroutines became less important. Accordingly, the further development of GARLIC and Py4CAtS became largely independent, and Py4CAtS is now a full line-by-line radiative transfer tool kit delivering absorption cross sections and coefficients, optical depths, transmissions, weighting functions, and radiances.

My contributions: Partly supported the update from Python 2 to 3, assisted in validating the Python scripts and provided comments during draft preparation.

Schreier, F. and <u>P. Hochstaffl</u>, 2021: Computational aspects of Speed-Dependent Voigt and Rautian profiles. _J. Quant. Spectrosc. & Radiat. Transfer_, 137, 29–50, doi: 10.1016/j.jqsrt.2020.107385.

Context: The increasing quality in atmospheric spectroscopy observations has indicated that physical processes beyond pressure and Doppler broadening considered by the Voigt profile should be treated in the computation of molecular absorption cross sections. In order to compute higher-order effects such as the speed-dependence of air broadening or collisional narrowing the imaginary component of the complex error function is employed and differences of two complex error functions are evaluated. Numerical problems due to cancellation errors are discussed and a numerically stable reformulation proposed. Moreover, the impact of various complex error function algorithms such as the Humlíček rational approximation and the Weideman approximation is studied.

My contributions: Verified many computational aspects proposed in the manuscript by own implementations in Pub. II and Pub. III. I suggested the reformulation according to Eq. (12) in the paper (Eq. (2.69) in this thesis) and was involved in the preparation of the manuscript.

Abbreviations and Acronyms

Acronym	Description	Introduced on page
ACE	Atmospheric Chemistry Experiment	1
AIRS	Atmospheric Infrared Sounder	2
BIRRA	Beer InfraRed Retrieval Algorithm	8
CIA	Collision Induced Absorption	9
CKD	Clough-Kneizys-Davies continuum	9
CO2M	CO_2 Monitoring mission	11
CSA	Canadian Space Agency	2
CSC	Copernicus Space Component	8
DAAD	German Academic Exchange Service	99
DLR	German Aerospace Centre	9
ECSS	European Cooperation for Space Standardization	9
ENVISAT	Environmental Satellite	1
ESA	European Space Agency	1
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites	2
FTIR	Fourier Transform InfraRed	7
FTS	Fourier Transform Spectrometer	1
FWHM	Full Width at Half Maximum	40
GARLIC	Generic Atmospheric Radiation Line-by-line Infrared Code	9
GEISA	Gestion et Etude des Informations Spectroscopiques Atmosphériques	9
GOME	Global Ozone Monitoring Experiment	1
GOME-2	Global Ozone Monitoring Experiment-2	1
GOMOS	Global Ozone Monitoring by Occultation of Stars	1
GOSAT	Greenhouse Gases Observing Satellite	2
GOSAT-2	Greenhouse Gases Observing Satellite-2	2
HITRAN	HIgh-resolution TRANsmission molecular absorption database	9
HT	Hartmann-Tran line profile	27
HTM	Hartmann-Tran with line-mixing profile	34
HWHM	Half Width at Half Maximum	22
IASI	Infrared Atmospheric Sounding Interferometer	2
ILS	Instrument Line Shape	10
IMAP-DOAS	Iterativ Maximum A Posteriori	8

Acronym	Description	Introduced on page
IMF	Remote Sensing Technology Institute	9
IMLM	Iterative Maximum Likelihood Method	8
IR	InfraRed	2
ISRF	Instrument Spectral Response Function	11
ISS	International Space Station	2
JAXA	Japan Aerospace Exploration Agency	2
LTE	Local Thermodynamic Equilibrium	13
MAPS	Measurement of Air Pollution from Satellites	6
METOP	METeorological OPerational	2
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding	1
MOPITT	Measurement of Pollution in the Troposphere	2
MT-CKD	Mlawer-Tobin-Clough-Kneizys-Davies continuum	38
NASA	National Aeronautics and Space Administration	1
NIR	Near InfraRed	2
NLS	Nonlinear least squares algorithm/fit	9
NMHC	Non-methane hydrocarbons	4
OCO-2	Orbiting Carbon Observatory-2	2
OCO-3	Orbiting Carbon Observatory-3	2
OMI	Ozone Monitoring Instrument	1
PORT	Portable Optimized . . . mathematical subroutine library	9
Py4CAtS	Python for Computational ATmospheric Spectroscopy	80
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY	1
SDR	Speed-dependent Rautian line profile	33
SDRM	Speed-dependent Rautian with line-mixing profile	29
SDV	Speed-dependent Voigt line profile	32
SDVM	Speed-dependent Voigt with line-mixing profile	34
S5P	Sentinel-5 Precursor	1
S5P–NPPC	Sentinel-5 Precursor National Polar-orbiting Partnership Cloud product	54
SEOM–IAS	Scientific Exploitation of Operational Missions – Improved Atmospheric Spectroscopy	10
SICOR	Shortwave Infrared Carbon Monoxide Retrieval	8
SLS	Separable least squares algorithm/fit	9
SNR	Signal-to-Noise Ratio	3
SSA	Single Scattering Albedo	16
SWIR	Short Wave InfraRed	2
TES	Tropospheric Emissions Spectrometer	6
TIR	Thermal InfraRed	2
TIROS	Television and Infrared Observation Satellite	1
ToA	Top of Atmosphere (at $\approx 120\,$km altitude)	2
TROPOMI	TROPOspheric Monitoring Instrument	1
UV	Ultra Violet	2
VIS	VISible	2

Acronym	Description	Introduced on page
WFM-DOAS	Weighted Function Modified Differential Optical Absorption Spectroscopy	8

Nomenclature

Acronym	Description	Introduced on page
Aqua	A National Aeronautics and Space Administration Earth science satellite mission	2
CH_4	Methane chemical formula	2
CO	Carbon monoxide chemical formula	2
CO_2	Carbon dioxide chemical formula	2
Copernicus	The European Union's Earth observation program	1
GHGSat	A private company in the remote sensing business	2
H_2O	Water chemical formula	10
N_2	Nitrogen chemical formula	22
NO_x	Nitrogen oxides, in particular nitric oxide and nitrogen dioxide	4
NO_2	Nitrogen dioxide chemical formula	4
NO	Nitric oxide chemical formula	4
O_2	Oxygen chemical formula	22
Sentinel	The European Space Agency's Earth observation missions for the Copernicus Space Component	1

Symbols

Sign	Meaning	Units	Introduced on page
A_{21}	Einstein A_{21} coefficient	s^{-1}	19
$\boldsymbol{\alpha}$	Nonlinear state vector elements of molecular and aerosol scaling factors	1	53
$A_{\perp}$	Projected area	cm^2	15
B_{if}	Einstein B coefficients for absorption B_{12} and stimulated emission B_{21}	$\mathrm{erg}^{-1}\,\mathrm{cm}^3\,\mathrm{s}^{-2}$	18
$\boldsymbol{\beta}$	Linear state vector elements	-	52
ω_{10}	Bohr angular frequency	s^{-1}	21
$\gamma_{\mathrm{L}}^{(\mathrm{air})}$	Air-pressure broadening coefficient	$\mathrm{cm}^{-1}\,\mathrm{atm}^{-1}$	22
$\mathbf{C}$	Column averaging kernel matrix	1	56
C	Noise (co-)variance matrix	1	49
Γ	Line broadening coefficient	cm^{-1}	18
c	Speed of light in vacuum	$\mathrm{cm}\,\mathrm{s}^{-1}$	13
Γ	Line half-width	cm^{-1}	18
$\delta_p^{(\mathrm{air})}$	Pressure dependence of air-pressure induced line shift	$\mathrm{cm}^{-1}\,\mathrm{atm}^{-1}$	22
$\delta_T^{(\mathrm{air})}$	Temperature dependence of air-pressure induced line shift	$\mathrm{cm}^{-1}\,\mathrm{atm}^{-1}\,\mathrm{K}^{-1}$	22
d_j	Degeneracy of the vibrational state j	1	19
$\boldsymbol{M}$	Oscillating dipole moment	$\mathrm{dyn}^{\frac{1}{2}}\,\mathrm{cm}^2$	21
Δ_{L}	Pressure-induced line shift	cm^{-1}	22
η	Partial correlation of velocity changes due to collisions and the speed-dependence of relaxation rates	1	27
E	Spectral irradiance	$\mathrm{erg}\,\mathrm{s}^{-1}\,\mathrm{cm}^{-3}$	15
E_n	Energy of the harmonic oscillator mode n	erg	14
$\boldsymbol{\varepsilon}$	Errors of the measurement vector	$\mathrm{erg}\,\mathrm{s}^{-1}\,\mathrm{cm}^{-3}\,\mathrm{sr}^{-1}$	45
$\boldsymbol{\eta}$	Nonlinear state vector elements	-	52
$|f\rangle$	Final state	1	18
$\boldsymbol{F}$	Forward model	$\mathrm{erg}\,\mathrm{s}^{-1}\,\mathrm{cm}^{-3}\,\mathrm{sr}^{-1}$	3
$f_{n>1}$	Overtone frequencies	s^{-1}	14
f	Frequency	s^{-1}	13
f_1	Fundamental (first harmonic) frequency	s^{-1}	14
g_f	Degeneracy of the final energy state	1	19
g_i	Degeneracy of the initial energy state	1	19

Sign	Meaning	Units	Introduced on page
g	Line profile function (normalized to unity)	1	18
h	Planck constant	$\mathrm{erg\,s}$	13
$\lvert i \rangle$	Initial state	1	18
I	Spectral radiance	$\mathrm{erg\,s^{-1}\,cm^{-3}\,sr^{-1}}$	15
J'	Rotational transition number	1	5
J	Jacobi matrix	1	51
J	Emission coefficient	$\mathrm{erg\,s^{-1}\,g^{-1}}$	16
J_S	Scattering source function	$\mathrm{erg\,s^{-1}\,cm^{-3}\,sr^{-1}}$	16
k	Absorption cross section; also designates the order the ISRF	$\mathrm{cm^2}$	17
k_B	Boltzmann constant	$\mathrm{erg\,K^{-1}}$	14
L	Likelihood for the observation (of M measurements)	1	49
μ_a	Absorption coefficient	$\mathrm{cm^{-1}}$	13
μ_s	Scattering coefficient	$\mathrm{cm^{-1}}$	16
$\hat{\nu}$	Transition wavenumber (line position)	$\mathrm{cm^{-1}}$	18
$\mathcal{N}$	Null space (the kernel of a linear map)	1	45
n_m	Number density of molecule m	$\mathrm{molec\,cm^{-3}}$	17
Ω	Solid angle	1	15
P_B	Boltzmann probability distribution	1	14
P_M	Joint probability function for measurements M	1	49
P_S	Normalized scattering phase function	1	16
Q	Total internal partition sum	1	19
B	Spectral radiance	$\mathrm{erg\,s^{-1}\,cm^{-3}\,sr^{-1}}$	15
$\mathcal{R}$	Column space of a matrix i.e. its range	1	45
R_{if}	Dipole moment operator (weighted transition moment squared)	$\mathrm{erg\,cm^3}$	19
r	Reflectivity	1	18
$\boldsymbol{\rho}$	Spectral residuum vector	$\mathrm{erg\,s^{-1}\,cm^{-3}\,sr^{-1}}$	67
ρ	Mass density	$\mathrm{g\,cm^{-3}}$	16
s	Slant distance	cm	15
σ	Standard deviation	$\mathrm{cm^{-1}}$	40
σ_l	Singular values	1	47
$\mathcal{S}$	Instrument spectral response function	1	38
S	Spectral line intensity (line strength)	$\mathrm{cm^{-1}\,(molec\,cm^{-2})^{-1}}$	18
τ	Molecular optical depth	1	17
T	Thermodynamic temperature	K	14
θ	Solar zenith angle	$^\circ$	18
$\mathcal{T}$	Monochromatic transmission	1	17
v'	Vibrational transition number	1	5
ν	Wavenumber	$\mathrm{cm^{-1}}$	13
$\boldsymbol{x}$	State vector containing the fit parameters	-	3
X	A unknown function of interest	1	43
Y	Rosenkranz line-mixing parameter	1	27

Sign	Meaning	Units	Introduced on page
$\boldsymbol{y}$	Measurement vector (radiance spectrum)	$\mathrm{erg\,s^{-1}\,cm^{-3}\,sr^{-1}}$	3
z	Altitude (vertical distance)	cm	18

Appendix A

Hartmann-Tran for TROPOMI

The HTM (Sec. 2.4.2.1) line profile was examined in the retrieval of CO from TROPOMI measurements and compared to results inferred with the SDRM (Sec. 2.4.2.4) line shape model. Both models consider line-mixing according to the Rosenkranz approximation from Sec. 2.4.2.5. The SEOM–IAS line data was used to calculate the absorption cross sections in either case. For the HTM profile a fixed partial correlation $\eta = 0.5$ was assumed since the parameter is not provided in the SEOM–IAS database.

Table A.1: Retrieved scaling factors for the HTM and SDRM line profiles based on a subset of TROPOMI measurements around $25\,\text{N}^\circ$ over Egypt on May 07, 2018 (orbit 2923).

	median		mean		variance	
	HTM	**SDRM**	**HTM**	**SDRM**	**HTM**	**SDRM**
α_{CO}	0.9566	0.9572	0.9338	0.9344	0.0463	0.0461
α_{CH_4}	1.0475	1.0485	1.0467	1.0478	0.0157	0.0158
$\alpha_{\text{H}_2\text{O}}$	1.3450	1.3555	1.3929	1.3945	0.4495	0.4487

Table A.1 shows the retrieval results for the HTM (see Sec. 2.4.2.1) and SDRM (see Sec. 2.4.2.4) line shape models. The outcome reveals that the retrieved scaling factors α_m are very similar for both line profiles ($< 1\,\%$). Nonetheless, the effect of partial correlation can become crucial in cases where the target needs to be determined with very high accuracy, such as in current or upcoming CO_2 missions (*Oyafuso et al.*, 2017; *Kuhlmann et al.*, 2019).

Appendix B

BIRRA CO maps from TROPOMI

Subsequent maps depict CO mole fractions and corresponding errors over various continents across the globe. Figure B.1 shows the CO concentrations over America for three consecutive orbits (9538–9540). In Fig. B.3 CO concentrations over Australia for orbits 11516–11518 are shown. Finally, Fig. B.2 shows the distribution of CO over Africa and Europe for orbits 10045–10047.

(a)　　　　　　　　　　　　　　　　　　　　**(b)**
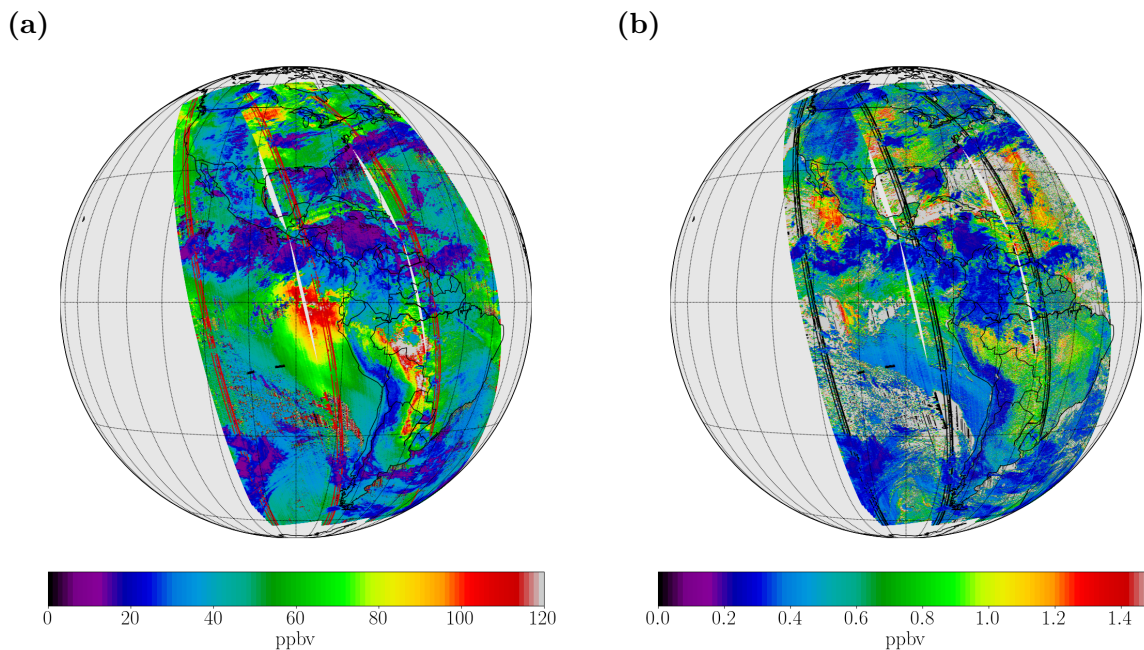
Figure B.1: **(a)** BIRRA retrieved CO from TROPOMI observations over North and South America on August 16, 2019. **(b)** Corresponding CO errors.
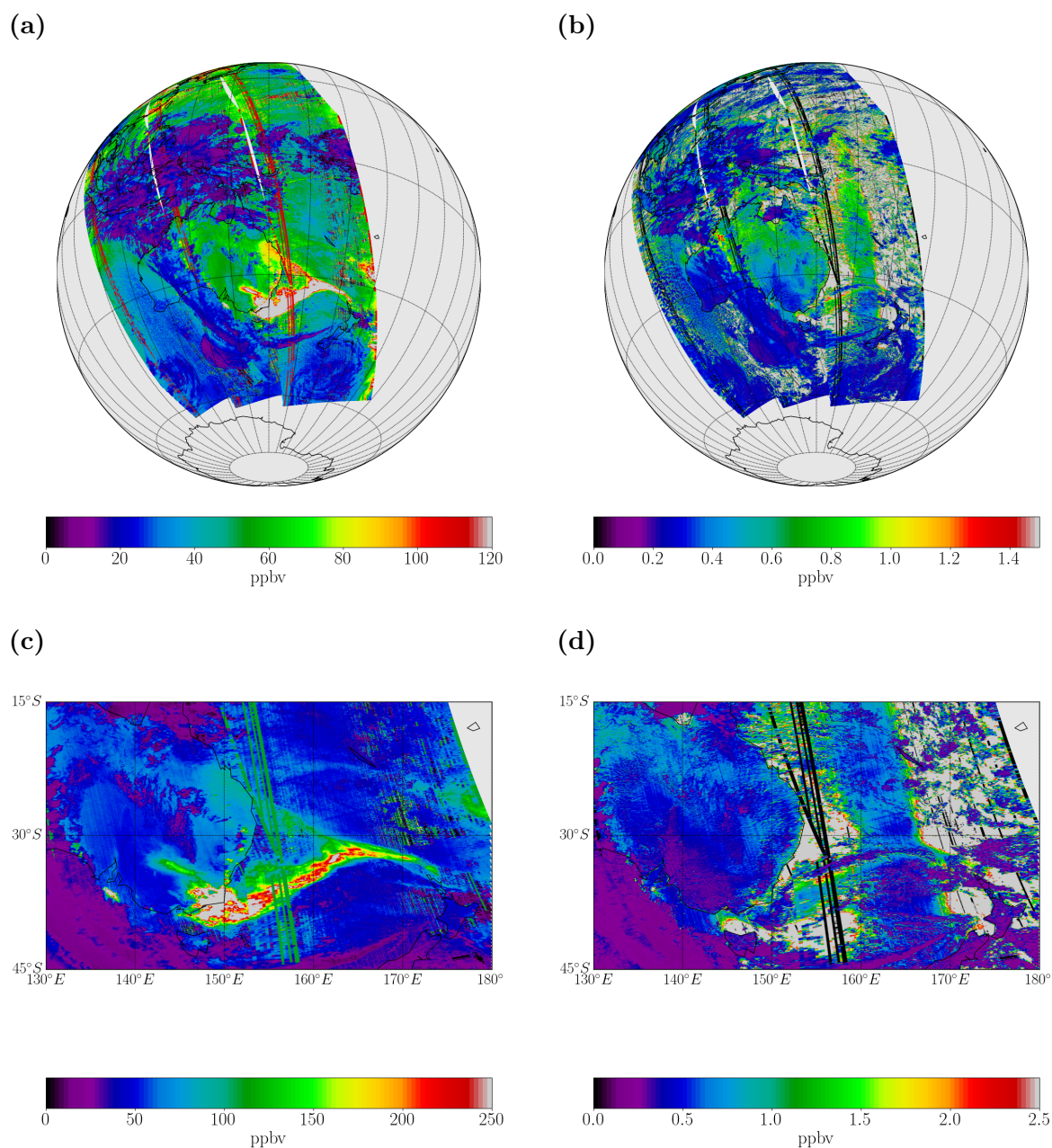
Figure B.2: **(a)** BIRRA inferred CO mole fractions from TROPOMI over Australia during the wildfires on January 3, 2020. **(c)** Zoom into the eastern parts of Australia, the Tasman Sea and parts of New Zealand. **(b, d)** Corresponding mole fraction errors. Small errors over the ocean across the Tasman Sea (from north to south) correspond to sun glint observations. The striping pattern is an artefact.

(a)
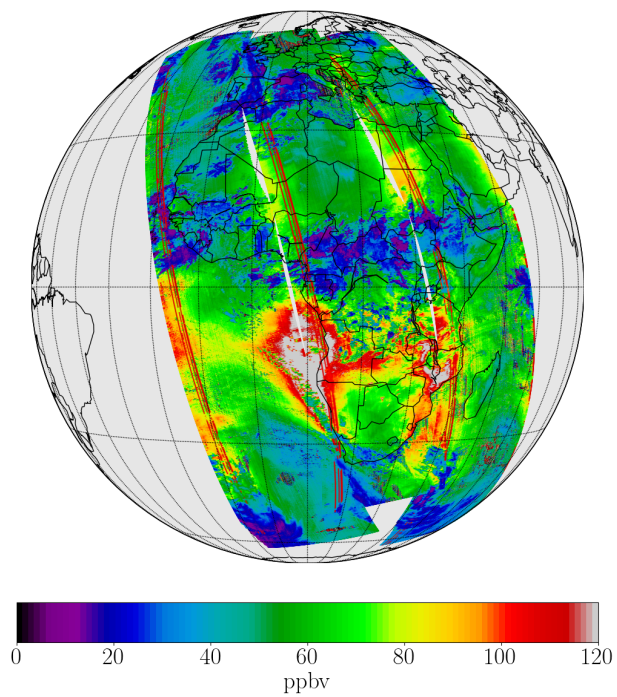
(b)
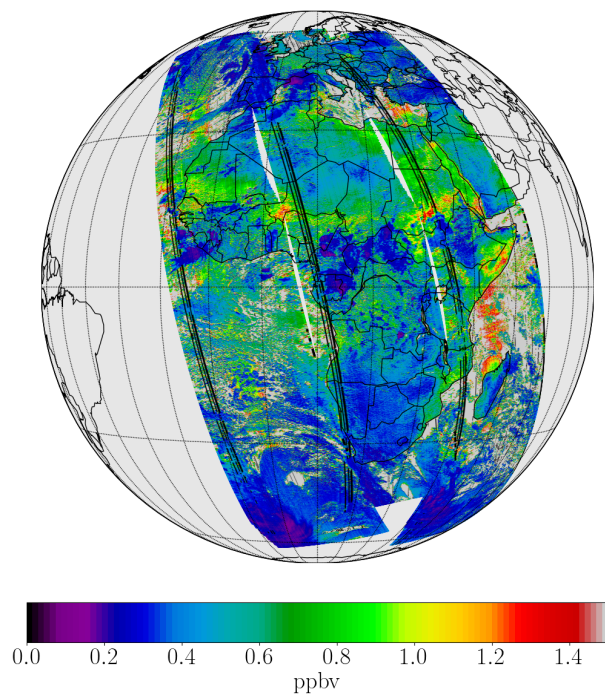
Figure B.3: **(a)** BIRRA retrieved CO mole fractions over Africa and Europe on September 21, 2019. **(b)** Corresponding errors.

Acknowledgments

I would like to express my gratitude to *Prof. Dr. Mark Wenig* and *Prof. Dr. Thomas Trautmann* for their interest in and support of this work, for their patience and flexibility and, last but not least, for reviewing this thesis. A special thanks to the latter for giving me the opportunity to work at his department within the DLR-DAAD (German Academic Exchange Service) research fellowship program. Doing research in this framework was an excellent experience.

Furthermore, I sincerely thank *Dr. Franz Schreier* for supervising the thesis and sharing his outstanding competency and scientific knowledge. Without his guidance and advice this research would not have been possible. I also want to thank Sebastian Gimeno García for his advice and support in the initial phase of the program.

I greatly appreciate the fruitful discussions with *Dr. Günter Lichtenberg, Dr. Manfred Birk, Dr. Georg Wagner* and *Dr. Dietrich G. Feist.* I thank *Prof. Dr. André Butz* for the personal communications and I want to thank *Bernd Aberle* and *Walter Zimmer* for the technical support that was provided.

Finally, I want to thank my defense committee, *Prof. Dr. Hartmut Zohm, Prof. Dr. Barbara Ercolano, Prof. Dr. Otmar Biebel* and *Prof. Dr. Lode Pollet.*

It was a pleasure meeting people across the field and having many constructive discussions with them at conferences and workshops.

Bibliography

Aben, I., O. Hasekamp, and W. Hartmann, 2007: Uncertainties in the space-based measurements of CO_2 columns due to scattering in the Earth's atmosphere. *J. Quant. Spectrosc. & Radiat. Transfer*, **104 (3)**, 450–459, doi: 10.1016/j.jqsrt.2006.09.013.

Abramowitz, M., and I. Stegun, 1964: *Handbook of Mathematical Functions*. National Bureau of Standards, AMS55, New York.

Adams, J. C., W. S. Brainerd, R. A. Hendrickson, R. E. Maine, J. T. Martin, and B. T. Smith, 2008: *The Fortran 2003 Handbook: The Complete Syntax, Features and Procedures*. Springer London.

Anderson, G., S. Clough, F. Kneizys, J. Chetwynd, and E. Shettle, 1986: AFGL atmospheric constituent profiles (0 – 120 km). Tech. Rep. TR-86-0110, AFGL.

Armstrong, B., 1967: Spectrum line profiles: The Voigt function. *J. Quant. Spectrosc. & Radiat. Transfer*, **7**, 61–88, doi: 10.1016/0022-4073(67)90057-X.

Aster, R., B. Borchers, and C. Thurber, 2018: *Parameter Estimation and Inverse Problems*. Elsevier Science.

Beirle, S., J. Lampel, C. Lerot, H. Sihler, and T. Wagner, 2017: Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives. *Atmos. Meas. Tech.*, **10 (2)**, 581–598, doi: 10.5194/amt-10-581-2017.

Berman, P., 1972: Speed-dependent collisional width and shift parameters in spectral profiles. *J. Quant. Spectrosc. & Radiat. Transfer*, **12**, 1321–1342, doi: 10.1016/0022-4073(72)90189-6.

Bernath, P., 2017: The atmospheric chemistry experiment (ACE). *J. Quant. Spectrosc. & Radiat. Transfer*, **186**, 3 – 16, doi: 10.1016/j.jqsrt.2016.04.006.

Bernath, P., and Coauthors, 2005: Atmospheric Chemistry Experiment (ACE): Mission overview. *Geophys. Res. Letters*, **32**, L15S01, doi: 10.1029/2005GL022386.

Bernath, P. F., 2016: *Spectra of Atoms and Molecules*. Oxford University Press.

Birk, M., and G. Wagner, 2016: Voigt profile introduces optical depth dependent systematic errors — detected in high resolution laboratory spectra of water. *J. Quant. Spectrosc. & Radiat. Transfer*, **170**, 159–168, doi: 10.1016/j.jqsrt.2015.11.008.

Birk, M., G. Wagner, J. Loos, D. Mondelain, and A. Campargue, 2017a: ESA SEOM–IAS - Measurement database — 2.3 μm region [Data set]. Zenodo. Zenodo, doi: 10.5281/zenodo.1009122.

Birk, M., G. Wagner, J. Loos, D. Mondelain, and A. Campargue, 2017b: ESA SEOM–IAS
- Spectroscopic parameters database — 2.3 μm region [Data set]. Zenodo. Zenodo, doi:
10.5281/zenodo.1009126.

Bodhaine, B., N. Wood, E. Dutton, and J. Slusser, 1999: On Rayleigh Optical Depth
Calculations. *J. Atmos. Oceanic Techn.*, **16 (11)**, 1854–1861, doi: 10.1175/1520-
0426(1999)016¡1854:ORODC¿2.0.CO;2.

Boone, C., K. Walker, and P. Bernath, 2007: Speed–dependent Voigt profile for water
vapor in infrared remote sensing applications. *J. Quant. Spectrosc. & Radiat. Transfer*,
105, 525–532, doi: 10.1016/j.jqsrt.2006.11.015.

Boone, C., K. Walker, and P. Bernath, 2011: An efficient analytical approach for calcu-
lating line mixing in atmospheric remote sensing applications. *J. Quant. Spectrosc. &
Radiat. Transfer*, **112 (6)**, 980 – 989, doi: 10.1016/j.jqsrt.2010.11.013.

Borsdorff, T., J. Aan de Brugh, H. Hu, I. Aben, O. Hasekamp, and J. Landgraf,
2018: Measuring Carbon Monoxide With TROPOMI: First Results and a Compari-
son With ECMWF-IFS Analysis Data. *Geophys. Res. Letters*, **45 (6)**, 2826–2832, doi:
10.1002/2018GL077045.

Borsdorff, T., J. aan de Brugh, H. Hu, P. Nédélec, I. Aben, and J. Landgraf, 2017: Carbon
monoxide column retrieval for clear-sky and cloudy atmospheres: a full-mission data
set from SCIAMACHY 2.3 μm reflectance measurements. *Atmos. Meas. Tech.*, **10 (5)**,
1769–1782, doi: 10.5194/amt-10-1769-2017.

Borsdorff, T., P. Tol, J. E. Williams, J. de Laat, J. aan de Brugh, P. Nédélec, I. Aben,
and J. Landgraf, 2016: Carbon monoxide total columns from SCIAMACHY 2.3 μm at-
mospheric reflectance measurements: towards a full-mission data product (2003–2012).
Atmos. Meas. Tech., **9 (1)**, 227–248, doi: 10.5194/amt-9-227-2016.

Borysow, A., 2002: Collision-induced absorption in the infrared: A data base for modelling
planetry and stellar atmospheres. URL http://www.astro.ku.dk/~aborysow/programs/
index.html.

Bovensmann, H., J. Burrows, M. Buchwitz, J. Frerick, S. Noël, V. Rozanov, K. Chance,
and A. Goede, 1999: SCIAMACHY: Mission objectives and measurement mode. *J.
Atmos. Sci.*, **56**, 127–150.

Bransden, B., C. Joachain, and T. Plivier, 2003: *Physics of Atoms and Molecules.* Pearson
Education, Prentice Hall.

Brooks, R. L., 2014: *The Fundamentals of Atomic and Molecular Physics.* Undergraduate
Lecture Notes in Physics, Springer New York.

Bucholtz, A., 1995: Rayleigh-scattering calculations for the terrestrial atmosphere. *Appl.
Opt.*, **34 (15)**, 2765–2773.

Buchwitz, M., and J. Burrows, 2003: Retrieval of CH_4, CO, and CO_2 total column
amounts from SCIAMACHY near-infrared nadir spectra: Retrieval algorithm and first
results. *10. International Symposium Remote Sensing — Remote Sensing of Clouds and
the Atmosphere VIII*, Proceedings of SPIE, Vol. 5235.

Buchwitz, M., R. de Beek, K. Bramstedt, S. Noël, H. Bovensmann, and J. P. Burrows,

2004: Global carbon monoxide as retrieved from SCIAMACHY by WFM-DOAS. *Atm. Chem. Phys.*, **4**, 1945–1960, doi: 10.5194/acp-4-1945-2004.

Buchwitz, M., I. Khlystova, H. Bovensmann, and J. P. Burrows, 2007a: Three years of global carbon monoxide from SCIAMACHY: comparison with MOPITT and first results related to the detection of enhanced CO over cities. *Atm. Chem. Phys.*, **7**, 2399–2411, doi: 10.5194/acp-7-2399-2007.

Buchwitz, M., V. Rozanov, and J. Burrows, 2000: A near-infrared optimized DOAS method for the fast global retrieval of atmospheric CH_4, CO, CO_2, H_2O, and N_2O total column amounts from SCIAMACHY Envisat-1 nadir radiances. *J. Geophys. Res.*, **105 (D12)**, 15 231–15 245, doi: 10.1029/2000JD900191.

Buchwitz, M., O. Schneising, J. P. Burrows, H. Bovensmann, M. Reuter, and J. Notholt, 2007b: First direct observation of the atmospheric CO_2 year-to-year increase from space. *Atm. Chem. Phys.*, **7**, 4249–4256.

Buchwitz, M., and Coauthors, 2005: Carbon monoxide, methane and carbon dioxide columns retrieved from SCIAMACHY by WFM-DOAS: year 2003 initial data set. *Atm. Chem. Phys.*, **5**, 3313–3329, doi: 10.5194/acp-5-3313-2005.

Burrows, J. P., U. Platt, and P. Borrell, 2011: *The Remote Sensing of Tropospheric Composition from Space.* Physics of Earth and Space Environment, Springer-Verlag, Berlin Heidelberg, doi: 10.1007/978-3-642-14791-3.

Burrows, J. P., and Coauthors, 1999: The Global Ozone Monitoring Experiment (GOME): Mission Concept and First Scientific Results. *J. of the Atmos. Sci.*, **56 (2)**, 151–175, doi: 10.1175/1520-0469(1999)056¡0151:TGOMEG¿2.0.CO;2.

Butz, A., O. P. Hasekamp, C. Frankenberg, and I. Aben, 2009: Retrievals of atmospheric CO_2 from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects. *Appl. Opt.*, **48 (18)**, 3322–3336, doi: 10.1364/AO.48.003322.

Butz, A., and Coauthors, 2011: Toward accurate CO_2 and CH_4 observations from GOSAT. *Geophys. Res. Letters*, **38 (14)**, L14 812, doi: 10.1029/2011GL047888.

Chahine, M., 1991: AIRS: The atmospheric infrared sounder. *Optics & Photonics News*, 25–27.

Chahine, M., and Coauthors, 2006: AIRS: Improving weather forecasting and providing new data on greenhouse gases. *Bull. Am. Met. Soc.*, **87**, 911–926, doi: 10.1175/BAMS-87-7-911.

Checa-García, R., and Coauthors, 2015: Mapping spectroscopic uncertainties into prospective methane retrieval errors from Sentinel-5 and its precursor. *Atmos. Meas. Tech.*, **8 (9)**, 3617–3629, doi: 10.5194/amt-8-3617-2015.

Chen, Y., and J. Surmont, 1976: Iterative algorithms for nonlinear problems in remote sensing. *Appl. Math. Comp.*, **2**, 197–228.

Clerbaux, C., D. P. Edwards, M. Deeter, L. Emmons, J.-F. Lamarque, X. Tie, S. T. Massie, and J. Gille, 2008: Carbon monoxide pollution from cities and urban areas observed by the Terra/MOPITT mission. *Geophys. Res. Letters*, **35**, L03 817, doi:

10.1029/2007GL032300.

Clerbaux, C., and Coauthors, 2005: Carbon monoxide distribution from the ACE-FTS solar occultation measurements. *Geophys. Res. Letters*, **32**, L16S01, doi: 10.1029/2005GL022394.

Clerbaux, C., and Coauthors, 2009: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atm. Chem. Phys.*, **9 (16)**, 6041–6054, doi: 10.5194/acp-9-6041-2009.

Clough, S., F. Kneizys, and R. Davies, 1989: Line shape and the water vapor continuum. *Atm. Res.*, **23**, 229–241, doi: 10.1016/0169-8095(89)90020-3.

Crisp, D., and Coauthors, 2004: The Orbiting Carbon Observatory (OCO) mission. *Adv. Space Res.*, **34 (4)**, 700 – 709, doi: 10.1016/j.asr.2003.08.062.

Cusworth, D. H., and Coauthors, 2019: Potential of next-generation imaging spectrometers to detect and quantify methane point sources from space. *Atmos. Meas. Tech.*, **12 (10)**, 5655–5668, doi: 10.5194/amt-12-5655-2019.

Dakin, J., M. Gunning, P. Chambers, and Z. Xin, 2003: Detection of gases by correlation spectroscopy. *Sensors and Actuators B: Chemical*, **90 (1)**, 124 – 131, doi: 10.1016/S0925-4005(03)00043-1, proceedings of the 6th European Conference on Optical Chemical Sensors and Biosensors EUROPT(R)ODE VI.

de Laat, A., A. Gloudemans, I. Aben, M. Krol, J. Meirink, G. van der Werf, and H. Schrijver, 2007: Scanning Imaging Absorption Spectrometer for Atmospheric Chartography carbon monoxide total columns: Statistical evaluation and comparison with chemistry transport model results. *J. Geophys. Res.*, **112**, D12 310, doi: 10.1029/2006JD008256.

de Laat, A., A. Gloudemans, H. Schrijver, M. van den Broek, J. Meirink, I. Aben, and M. Krol, 2006: Quantitative analysis of SCIAMACHY carbon monoxide total column measurements. *Geophys. Res. Letters*, **33**, L07 807, doi: 10.1029/2005GL025530.

de Laat, A. T. J., and Coauthors, 2010: Validation of five years (2003 – 2007) of SCIAMACHY CO total column measurements using ground–based spectrometer observations. *Atmos. Meas. Tech.*, **3 (5)**, 1457–1471, doi: 10.5194/amt-3-1457-2010.

Deeter, M., D. Edwards, J. Gille, and J. Drummond, 2009: CO retrievals based on MOPITT near-infrared observations. *J. Geophys. Res.*, **114**, D04 303, doi: 10.1029/2008JD010872.

Deeter, M. N., and Coauthors, 2003: Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument. *J. Geophys. Res.*, **108 (D14)**, 4399, doi: 10.1029/2002JD003186.

Dicke, R. H., 1953: The effect of collisions upon the Doppler width of spectral lines. *Phys. Rev.*, **89**, 472–473, doi: 10.1103/PhysRev.89.472.

Dils, B., and Coauthors, 2006: Comparisons between SCIAMACHY and ground–based FTIR data for total columns of CO, CH_4, CO_2 and N_2O. *Atm. Chem. Phys.*, **6**, 1953–1976, doi: 10.5194/acp-6-1953-2006.

DLMF, 2021: *NIST Digital Library of Mathematical Functions.* National Institute of

Standards and Technology, URL http://dlmf.nist.gov/, online companion to *Olver et al.* (2010).

Edwards, D., 1988: Atmospheric transmittance and radiance calculations using line–by–line computer models. *Modelling of the Atmosphere*, Proc. SPIE, Vol. 928, 94–116, doi: 10.1117/12.975622.

Edwards, D., and Coauthors, 2004: Observations of carbon monoxide and aerosols from the Terra satellite: Northern hemisphere variability. *J. Geophys. Res.*, **109**, D24 202, doi: 10.1029/2004JD004727.

Ekwurzel, B., J. Boneham, M. W. Dalton, R. Heede, R. J. Mera, M. R. Allen, and P. C. Frumhoff, 2017: The rise in global atmospheric CO_2, surface temperature, and sea level from emissions traced to major carbon producers. *Climatic Change*, **144 (4)**, 579–590, doi: 10.1007/s10584-017-1978-0.

Eldering, A., T. E. Taylor, C. W. O'Dell, and R. Pavlick, 2019: The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data. *Atmos. Meas. Tech.*, **12 (4)**, 2341–2370, doi: 10.5194/amt-12-2341-2019.

Elsey, J., M. D. Coleman, T. D. Gardiner, K. P. Menang, and K. P. Shine, 2020: Atmospheric observations of the water vapour continuum in the near-infrared windows between 2500 and 6600 cm^{-1}. *Atmos. Meas. Tech.*, **13 (5)**, 2335–2361, doi: 10.5194/amt-13-2335-2020.

Eriksson, P., 1999: Microwave radiometric observations of the middle atmosphere: Simulations and inversions. Ph.D. thesis, Department of Radio and Space Science — Chalmers University of Technology, Göteborg, Sweden.

Etminan, M., G. Myhre, E. J. Highwood, and K. P. Shine, 2016: Radiative forcing of carbon dioxide, methane, and nitrous oxide: A significant revision of the methane radiative forcing. *Geophys. Res. Letters*, **43 (24)**, 12,614–12,623, doi: 10.1002/2016GL071930.

Fischer, H., and F. Hase, 2015: CHEMISTRY OF THE ATMOSPHERE — Observations for Chemistry (Remote Sensing): IR/FIR (Satellite, Balloon and Ground). *Encyclopedia of Atmospheric Sciences (Second Edition)*, G. R. North, J. Pyle, and F. Zhang, Eds., second edition ed., Academic Press, Oxford, 401–410, doi: 10.1016/B978-0-12-382225-3.00270-X.

Fischer, H., and Coauthors, 2008: MIPAS: an instrument for atmospheric and climate research. *Atm. Chem. Phys.*, **8 (8)**, 2151–2188, doi: 10.5194/acp-8-2151-2008.

Fleming, E., S. Chandra, J. Barnett, and M. Corney, 1990: Zonal mean temperature, pressure, zonal wind and geopotential height as functions of latitude. *Adv. Space Res.*, **10 (12)**, 11–59, doi: 10.1016/0273-1177(90)90386-E.

Fox, P., A. Hall, and N. Schryer, 1978: The PORT mathematical subroutine library. *ACM Trans. Math. Soft.*, **4 (2)**, 104–126, doi: 10.1145/355780.355789.

Frankenberg, C., U. Platt, and T. Wagner, 2005: Retrieval of CO from SCIAMACHY onboard ENVISAT: detection of strongly polluted areas and seasonal patterns in global CO abundances. *Atm. Chem. Phys.*, **5**, 1639–1644, doi: 10.5194/acp-5-1639-2005.

Galli, A., and Coauthors, 2012: CH_4, CO, and H_2O spectroscopy for the Sentinel-5

Precursor mission: an assessment with the Total Carbon Column Observing Network measurements. *Atmos. Meas. Tech.*, **5 (2)**, 1387–1398, doi: 10.5194/amt-5-1387-2012.

Gay, D., 1990: Usage summary for selected optimization routines (PORT mathematical subroutine library, optimization chapter). Computing Science Technical Report 153, AT&T Bell Laboratories, Murray Hill, NJ 07974. Available at http://netlib.bell-labs.com/cm/cs/cstr/153.pdf.

Gentle, J., 2017: *Matrix Algebra: Theory, Computations and Applications in Statistics.* Springer Texts in Statistics, Springer International Publishing.

George, M., and Coauthors, 2009: Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors. *Atm. Chem. Phys.*, **9 (21)**, 8317–8330, doi: 10.5194/acp-9-8317-2009.

George, M., and Coauthors, 2015: An examination of the long-term CO records from MOPITT and IASI: comparison of retrieval methodology. *Atmos. Meas. Tech.*, **8 (10)**, 4313–4328, doi: 10.5194/amt-8-4313-2015.

Gill, P. E., W. Murray, and M. H. Wright, 2019: *Practical Optimization.* Classics in Applied Mathematics, SIAM, Society for Industrial and Applied Mathematics.

Gimeno García, S., F. Schreier, G. Lichtenberg, and S. Slijkhuis, 2011: Near infrared nadir retrieval of vertical column densities: methodology and application to SCIAMACHY. *Atmos. Meas. Tech.*, **4 (12)**, 2633–2657, doi: 10.5194/amt-4-2633-2011.

Gloudemans, A., A. de Laat, H. Schrijver, I. Aben, J. Meirink, and G. van der Werf, 2009: SCIAMACHY CO over land and oceans: 2003–2007 interannual variability. *Atm. Chem. Phys.*, **9 (2)**, 3799–3813, doi: 10.5194/acp-9-3799-2009.

Gloudemans, A., H. Schrijver, O. Hasekamp, and I. Aben, 2008: Error analysis for CO and CH_4 total column retrievals from SCIAMACHY $2.3\,\mu m$ spectra. *Atm. Chem. Phys.*, **8**, 3999–4017, doi: 10.5194/acp-8-3999-2008.

Gloudemans, A., and Coauthors, 2005: The impact of SCIAMACHY near-infrared instrument calibration on CH_4 and CO total columns. *Atm. Chem. Phys.*, **5**, 2369–2383, doi: 10.5194/acp-5-2369-2005.

Golub, G., and V. Pereyra, 2003: Separable nonlinear least squares: the variable projection method and its applications. *Inverse Problems*, **19**, R1–R26, doi: 10.1088/0266-5611/19/2/201.

Golub, G., and C. Van Loan, 2013: *Matrix Computations.* Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press.

Goody, R., and Y. Yung, 1995: *Atmospheric Radiation: Theoretical Basis.* Oxford University Press.

Gordon, I., and Coauthors, 2017: The HITRAN2016 molecular spectroscopic database. *J. Quant. Spectrosc. & Radiat. Transfer*, **203**, 3 – 69, doi: 10.1016/j.jqsrt.2017.06.038.

Gottwald, M., and H. Bovensmann, Eds., 2011: *SCIAMACHY — Exploring the Changing Earth's Atmosphere.* Springer, Dordrecht, NL, doi: 10.1007/978-90-481-9896-2.

Hamidouche, M., S. Gimeno García, F. Schreier, M. Meringer, G. Lichtenberg,

P. Hochstaffl, and T. Trautmann, 2016: Atmospheric methane with SCIAMACHY: operational level 2 data analysis and verification. *Proceedings Living Planet Symposium 2016*, ESA, Vol. SP-740.

Hanel, R. A., B. J. Conrath, D. E. Jennings, and R. E. Samuelson, 2003: *Exploration of the Solar System by Infrared Remote Sensing.* Cambridge University Press.

Hansen, P., V. Pereyra, and G. Scherer, 2013: *Least Squares Data Fitting with Applications.* Johns Hopkins University Press.

Hanson, R., and J. Phillips, 1975: An adaptive numerical method for solving linear Fredholm integral equations of the first kind. *Num. Math.*, **24**, 291–307.

Hartmann, J., C. Boulet, and D. Robert, 2008: *Collisional Effects on Molecular Spectra: Laboratory Experiments and Models, Consequences for Applications.* Elsevier Science.

Hartmann, J.-M., H. Tran, and G. Toon, 2009: Influence of line mixing on the retrievals of atmospheric CO_2 from spectra in the 1.6 and 2.1μm regions. *Atm. Chem. Phys.*, **9**, 7303–7312, doi: 10.5194/acp-9-7303-2009.

Hochstaffl, P., 2015: Validation of carbon monoxide total columns from SCIAMACHY near infrared nadir spectra with NDACC/TCCON ground-based measurements. Master thesis, University of Innsbruck, Institute of Atmospheric and Cryospheric Sciences.

Holloway, T., H. Levy II, and P. Kasibhatla, 2000: Global distribution of carbon monoxide. *J. Geophys. Res.*, **105 (D10)**, 12 123–12 147, doi: 10.1029/1999JD901173.

Hu, H., and Coauthors, 2016: The operational methane retrieval algorithm for TROPOMI. *Atmos. Meas. Tech.*, **9 (11)**, 5423–5440, doi: 10.5194/amt-9-5423-2016.

Hui, A., B. Armstrong, and A. Wray, 1978: Rapid computation of the Voigt and complex error functions. *J. Quant. Spectrosc. & Radiat. Transfer*, **19**, 509–516, doi: 10.1016/0022-4073(78)90019-5.

Humlíček, J., 1979: An efficient method for evaluation of the complex probability function: the Voigt function and its derivatives. *J. Quant. Spectrosc. & Radiat. Transfer*, **21**, 309–313, doi: 10.1016/0022-4073(79)90062-1.

Humlíček, J., 1982: Optimized computation of the Voigt and complex probability function. *J. Quant. Spectrosc. & Radiat. Transfer*, **27**, 437–444, doi: 10.1016/0022-4073(82)90078-4.

Hummel, T., C. Meister, J. Krauser, and M. Wenig, 2021: Slit homogenizer introduced performance gain analysis based on Sentinel-5/UVNS spectrometer. *Atmospheric Measurement Techniques Discussions*, **2021**, 1–18, doi: 10.5194/amt-2021-15.

Jacob, D. J., 1999: *Introduction to Atmospheric Chemistry.* Princeton University Press.

Jacquinet-Husson, N., and Coauthors, 2008: The GEISA spectroscopic database: Current and future archive for Earth and planetary atmosphere studies. *J. Quant. Spectrosc. & Radiat. Transfer*, **109**, 1043–1059, doi: 10.1016/j.jqsrt.2007.12.015.

Jacquinet-Husson, N., and Coauthors, 2016: The 2015 edition of the GEISA spectroscopic database. *J. Mol. Spectrosc.*, **327**, 31 – 72, doi: 10.1016/j.jms.2016.06.007, new Visions of Spectroscopic Databases, Volume II.

Jervis, D., and Coauthors, 2021: The GHGSat-D imaging spectrometer. *Atmos. Meas. Tech.*, **14 (3)**, 2127–2140, doi: 10.5194/amt-14-2127-2021.

Kaltenegger, L., and W. Traub, 2009: Transits of Earth-like planets. *Astrophys. J.*, **698 (1)**, 519, doi: 10.1088/0004-637X/698/1/519.

Karman, T., and Coauthors, 2019: Update of the HITRAN collision-induced absorption section. *Icarus*, **328**, 160 – 175, doi: 10.1016/j.icarus.2019.02.034.

Kerzenmacher, T., and Coauthors, 2012: Validation of IASI FORLI carbon monoxide retrievals using FTIR data from NDACC. *Atmos. Meas. Tech.*, **5 (11)**, 2751–2761, doi: 10.5194/amt-5-2751-2012.

Kidder, S. Q., and T. H. Vonder Haar, 2010: Observing Weather from Space. *Science*, **327 (5969)**, 1085–1086, doi: 10.1126/science.1185867.

Kleipool, Q., and Coauthors, 2018: Pre-launch calibration results of the TROPOMI payload on-board the Sentinel-5 Precursor satellite. *Atmos. Meas. Tech.*, **11 (12)**, 6439–6479, doi: 10.5194/amt-11-6439-2018.

Kochanov, V., 2012: On systematic errors in spectral line parameters retrieved with the Voigt line profile. *J. Quant. Spectrosc. & Radiat. Transfer*, **113 (12)**, 1635 – 1641, doi: 10.1016/j.jqsrt.2012.03.024.

Kopacz, M., and Coauthors, 2010: Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES). *Atm. Chem. Phys.*, **10 (3)**, 855–876, doi: 10.5194/acp-10-855-2010.

Kramer, H. J., 2021: Copernicus: Sentinel-5P (Precursor - Atmospheric Monitoring Mission). European Space Agency (ESA), accessed: 2021-03-26, https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/copernicus-sentinel-5p.

Kuhlmann, G., G. Broquet, J. Marshall, V. Clément, A. Löscher, Y. Meijer, and D. Brunner, 2019: Detectability of CO_2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO_2 Monitoring (CO2M) mission. *Atmos. Meas. Tech.*, **12 (12)**, 6695–6719, doi: 10.5194/amt-12-6695-2019.

Kurucz, R., 2005: New atlases for solar flux, irradiance, central intensity, and limb intensity. *ATLAS12 and related codes*, Memorie della Socienta Astronomica Italiana Supplementi, Vol. 8, 69-71.

Kuze, A., H. Suto, M. Nakajima, and T. Hamazaki, 2009: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.*, **48 (35)**, 6716–6733, doi: 10.1364/AO.48.006716.

Kuze, A., and Coauthors, 2016: Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space. *Atmos. Meas. Tech.*, **9 (6)**, 2445–2461, doi: 10.5194/amt-9-2445-2016.

Kyrölä, E., and Coauthors, 2004: GOMOS on Envisat: an overview. *Adv. Space Res.*, **33 (7)**, 1020–1028, doi: 10.1016/S0273-1177(03)00590-8.

Landgraf, J., and Coauthors, 2016: Carbon monoxide total column retrievals from

TROPOMI shortwave infrared measurements. *Atmos. Meas. Tech.*, **9 (10)**, 4955–4975, doi: 10.5194/amt-9-4955-2016.

Lawson, C., and R. Hanson, 1995: *Solving Least Squares Problems.* 2nd ed., SIAM, Philadelphia, PA.

Letchworth, K., and D. Benner, 2007: Rapid and accurate calculation of the Voigt function. *J. Quant. Spectrosc. & Radiat. Transfer*, **107**, 173–192, doi: 10.1016/j.jqsrt.2007.01.052.

Levelt, P., and Coauthors, 2006: The Ozone Monitoring Instrument. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1093–1101, doi: 10.1109/TGRS.2006.872333.

Lévy, A., N. Lacome, and C. Chackerian, Jr., 1992: Collisional line mixing. *Spectroscopy of the Earth's atmosphere and interstellar medium*, K. Rao, and A. Weber, Eds., Academic Press, 261 – 338.

Lichtenberg, G., S. Gimeno García, F. Schreier, S. Slijkhuis, R. Snel, and H. Bovensmann, 2010: Impact of level 1 quality on SCIAMACHY level 2 retrieval. *38. COSPAR Scientific Assembly.*

Liou, K.-N., 2002: *An Introduction to Atmospheric Radiation.* 2nd ed., Academic Press.

Lisak, D., G. Rusciano, and A. Sasso, 2004: An accurate comparison of lineshape models on H_2O lines in the spectral region around $3 \, \mu$m. *J. Mol. Spectrosc.*, **227**, 162–171, doi: 10.1016/j.jms.2004.06.006.

Loeliger, J., 2010: *Versionskontrolle mit Git.* O'Reilly.

Loos, J., M. Birk, and G. Wagner, 2017: Measurement of positions, intensities and self-broadening line shape parameters of H_2O lines in the spectral ranges $1850 - 2280$ cm^{-1} and $2390 - 4000$ cm^{-1}. *J. Quant. Spectrosc. & Radiat. Transfer*, **203**, 119 – 132, doi: 10.1016/j.jqsrt.2017.02.013, HITRAN2016 Special Issue.

Loos, J., M. B. G. Wagner, and Coauthors, 2015: Spectroscopic database for TROPOMI/Sentinel-5 Precursor. *ATMOS 2015 — Advances in Atmospheric Science and Applications*, ESA.

López-Puertas, M., and F. Taylor, 2001: *Non-LTE Radiative Transfer in the Atmosphere.* Series on atmospheric, oceanic and planetary physics, World Scientific.

Lorente, A., and Coauthors, 2021: Methane retrieved from TROPOMI: improvement of the data product and validation of the first 2 years of measurements. *Atmos. Meas. Tech.*, **14 (1)**, 665–684, doi: 10.5194/amt-14-665-2021.

Luo, M., and Coauthors, 2007: Comparison of carbon monoxide measurements by TES and MOPITT: Influence of a priori data and instrument characteristics on nadir atmospheric species retrievals. *J. Geophys. Res.*, **112**, D09 303, doi: 10.1029/2006JD007663.

Margerum, E. A., 1983: A mathematical method for inversion in atmospheric remote sensing (ETL - 0346). Tech. Rep. 2, US Army Corps of Engineers Engineer Topographic Laboratories, Ft. Belvoir, Virginia, USA, 10 pp.

Martinie, R. J., J. J. Bultema, M. N. Vander Wal, B. J. Burkhart, D. A. Vander Griend, and R. L. DeKock, 2011: Bond Order and Chemical Properties of BF, CO, and N_2. *J.*

of Chem. Edu., **88 (8)**, 1094–1097, doi: 10.1021/ed100758t.

McMillan, W., and Coauthors, 2005: Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. *Geophys. Res. Letters*, **32**, L11 801, doi: 10.1029/2004GL021821.

Mlawer, E., V. Payne, J.-L. Moncet, J. Delamere, M. Alvarado, and D. Tobin, 2012: Development and recent evaluation of the MT-CKD model of continuum absorption. *Philos. Trans. Roy. Soc. London, Ser. A*, **370 (1968)**, 2520–2556, doi: 10.1098/rsta.2011.0295.

Müller-Kirsten, H., 2013: *Basics Of Statistical Physics (Second Edition)*. World Scientific Publishing Company.

Munro, R., and Coauthors, 2016: The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing – an overview. *Atmos. Meas. Tech.*, **9 (3)**, 1279–1301, doi: 10.5194/amt-9-1279-2016.

Nadarajah, S., 2005: A generalized normal distribution. *J. Appl. Stat.*, **32 (7)**, 685–694, doi: 10.1080/02664760500079464.

Nassar, R., T. G. Hill, C. A. McLinden, D. Wunch, D. B. A. Jones, and D. Crisp, 2017: Quantifying CO_2 Emissions From Individual Power Plants From Space. *Geophys. Res. Letters*, **44 (19)**, 10,045–10,053, doi: 10.1002/2017GL074702.

Neto, F., and A. da Silva Neto, 2012: *An Introduction to Inverse Problems with Applications*. Springer Berlin Heidelberg.

Newell, R. E., S. T. Shipley, V. S. Connors, and J. Reichle, H. G., 1988: Regional studies of potential carbon monoxide sources based on Space Shuttle and aircraft measurements. *J. of Atm. Chem.*, **6**, 61–81, doi: 10.1007/BF00048332.

Ngo, N., D. Lisak, H. Tran, and J.-M. Hartmann, 2013: An isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes. *J. Quant. Spectrosc. & Radiat. Transfer*, **129**, 89 – 100, doi: 10.1016/j.jqsrt.2013.05.034, erratum: JQSRT 134, 105 (2014).

Ngo, N., D. Lisak, H. Tran, and J.-M. Hartmann, 2014: Erratum to "an isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes" [j. quant. spectrosc. radiat. transf. 129 (2013) 89-100]". *J. Quant. Spectrosc. & Radiat. Transfer*, **134**, 105, doi: 10.1016/j.jqsrt.2013.10.016.

Ngo, N. H., H. Tran, R. R. Gamache, and J. M. Hartmann, 2012: Pressure effects on water vapour lines: beyond the Voigt profile. *Philos. Trans. Roy. Soc. London, Ser. A*, **370 (1968)**, 2495–2508, doi: 10.1098/rsta.2011.0272.

Nicolet, M., 1984: On the molecular scattering in the terrestrial atmosphere: An empirical formula for its calculation in the homosphere. *Planet. Space Sci.*, **32 (11)**, 1467–1468, doi: 10.1016/0032-0633(84)90089-8.

Nikitin, A., and Coauthors, 2010: GOSAT-2009 methane spectral line list in the 5550 – 6236 cm^{-1} range. *J. Quant. Spectrosc. & Radiat. Transfer*, **111 (12-13)**, 2211–2224, doi: 10.1016/j.jqsrt.2010.05.010.

Nikitin, A., and Coauthors, 2015: GOSAT-2014 methane spectral line list. *J. Quant.*

Spectrosc. & Radiat. Transfer, **154**, 63 – 71, doi: 10.1016/j.jqsrt.2014.12.003.

Noël, S., K. Bramstedt, H. Bovensmann, J. P. Burrows, C. Standfuss, E. Dufour, and B. Veihelmann, 2012: Quantification and mitigation of the impact of scene inhomogeneity on Sentinel-4 UVN UV-VIS retrievals. *Atmos. Meas. Tech. Disc.*, **5 (2)**, 2043–2075, doi: 10.5194/amtd-5-2043-2012.

Norton, R., and C. Rinsland, 1991: ATMOS data processing and science analysis methods. *Appl. Opt.*, **30**, 389–400, doi: 10.1364/AO.30.000389.

O'Dell, C. W., and Coauthors, 2018: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm. *Atmos. Meas. Tech.*, **11 (12)**, 6539–6576, doi: 10.5194/amt-11-6539-2018.

O'Leary, D. P., and B. W. Rust, 2013: Variable projection for nonlinear least squares problems. *Computational Optimization and Applications*, **54 (3)**, 579–593, doi: 10.1007/s10589-012-9492-9.

Olver, F., D. Lozier, R. Boisvert, and C. Clark, Eds., 2010: *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, print companion to DLMF.

Oyafuso, F., and Coauthors, 2017: High accuracy absorption coefficients for the Orbiting Carbon Observatory-2 (OCO-2) mission: Validation of updated carbon dioxide cross-sections using atmospheric spectra. *J. Quant. Spectrosc. & Radiat. Transfer*, **203**, 213 – 223, doi: 10.1016/j.jqsrt.2017.06.012, HITRAN2016 Special Issue.

Penndorf, R., 1957: Tables of the Refractive Index for Standard Air and the Rayleigh Scattering Coefficient for the Spectral Region between 0.2 and $20.0\,\mu$m and Their Application to Atmospheric Optics. *J. Opt. Soc. Am.*, **47 (2)**, 176–182, doi: 10.1364/JOSA.47.000176.

Pierrehumbert, R. T., 2010: *Principles of Planetary Climate*. Cambridge University Press.

Pine, A., 1994: Line shape asymmetries in ar-broadened HF (v=1-0) in the Dicke-narrowed regime. *J. Chem. Phys.*, **101**, 3444.

Pine, A. S., and T. Gabard, 2000: Speed-dependent broadening and line mixing in CH_4 perturbed by Ar and N_2 from multispectrum fits. *J. Quant. Spectrosc. & Radiat. Transfer*, **66 (1)**, 69–92, doi: 10.1016/S0022-4073(99)00222-8.

Poisson, N., M. Kanakidou, and P. Crutzen, 2000: Impact of Non-Methane Hydrocarbons on Tropospheric Chemistry and the Oxidizing Power of the Global Troposphere: 3-Dimensional Modelling Results. *J. of Atm. Chem.*, **36 (2)**, 157–230, doi: 10.1023/A:1006300616544.

Ralston, A., and P. Rabinowitz, 1978: *A First Course in Numerical Analysis*. 2nd ed., McGraw–Hill Book Company.

Rautian, S. G., 1999: On the absorption line profile of a molecular gas. *Optics and Spectroscopy*, **86**, 334–336.

Reichle Jr., H. G., and Coauthors, 1999: Space shuttle based global CO measurements during April and October 1994, MAPS instrument, data reduction, and data validation.

J. of Geophys. Res.: Atm., **104 (D17)**, 21 443–21 454, doi: 10.1029/97JD03299.

Reuter, M., M. Buchwitz, O. Schneising, J. Heymann, H. Bovensmann, and J. P. Burrows, 2010: A method for improved SCIAMACHY CO_2 retrieval in the presence of optically thin clouds. *Atmos. Meas. Tech.*, **3 (1)**, 209–232, doi: 10.5194/amt-3-209-2010.

Reuter, M., M. Buchwitz, O. Schneising, S. Krautwurst, C. W. O'Dell, A. Richter, H. Bovensmann, and J. P. Burrows, 2019: Towards monitoring localized CO_2 emissions from space: co-located regional CO_2 and NO_2 enhancements observed by the OCO-2 and S5P satellites. *Atm. Chem. Phys.*, **19 (14)**, 9371–9383, doi: 10.5194/acp-19-9371-2019.

Reuter, M., M. Buchwitz, O. Schneising, S. Noël, V. Rozanov, H. Bovensmann, and J. P. Burrows, 2017: A fast atmospheric trace gas retrieval for hyperspectral instruments approximating multiple scattering — part 1: Radiative transfer and a potential OCO-2 XCO2 retrieval setup. *Remote Sens.*, **9 (11)**, 1159, doi: 10.3390/rs9111159.

Richard, C., and Coauthors, 2012: New section of the HITRAN database: Collision-induced absorption (CIA). *J. Quant. Spectrosc. & Radiat. Transfer*, **113 (11)**, 1276–1285, doi: 10.1016/j.jqsrt.2011.11.004.

Richter, A., 2010: Satellite remote sensing of tropospheric composition - principles, results, and challenges. *EPJ Web of Conferences*, **9**, 181–189, doi: 10.1051/epj-conf/201009014.

Rinsland, C., and Coauthors, 2006: Nadir measurements of carbon monoxide distributions by the Tropospheric Emission Spectrometer onboard the Aura spacecraft: Overview of analysis approach and examples of initial results. *Geophys. Res. Letters*, **33 (22)**, L22 806, doi: 10.1029/2006GL027000.

Rodgers, C., 2000: *Inverse Methods for Atmospheric Sounding: Theory and Practise.* World Scientific, Singapore.

Rodgers, C., and B. Connor, 2003: Intercomparison of remote sounding instruments. *J. Geophys. Res.*, **108 (D3)**, 4116, doi: 10.1029/2002JD002299.

Rohart, F., G. Wlodarczak, J.-M. Colmont, G. Cazzoli, L. Dore, and C. Puzzarini, 2008: Galatry versus speed-dependent Voigt profiles for millimeter lines of O_3 in collision with N_2 and O_2. *J. Mol. Spectrosc.*, **251 (1-2)**, 282 – 292, doi: 10.1016/j.jms.2008.03.005, special issue dedicated to the pioneering work of Drs. Edward A. Cohen and Herbert M. Pickett on spectroscopy relevant to the Earth's atmosphere and astrophysics.

Rosenkranz, P., 1975: Shape of the 5 mm oxygen band in the atmosphere. *IEEE Transactions on Antennas and Propagation*, **23 (4)**, 498–506, doi: 10.1109/TAP.1975.1141119.

Rothman, L., and Coauthors, 1987: The HITRAN database: 1986 edition. *Appl. Opt.*, **26**, 4058, doi: 10.1364/AO.26.004058.

Rothman, L., and Coauthors, 1998: The HITRAN molecular spectroscopic database and HAWKS (Hitran Atmospheric Workstation): 1996 edition. *J. Quant. Spectrosc. & Radiat. Transfer*, **60**, 665–710.

Rothman, L., and Coauthors, 2009: The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. & Radiat. Transfer*, **110 (9-10)**, 533 – 572, doi:

10.1016/j.jqsrt.2009.02.013.

Rust, B. W., 2001a: Fitting nature's basic functions Part I: Polynomials and linear least squares. *Computing in Science & Eng.*, **3 (5)**, 84–89, doi: 10.1109/MCISE.2001.947111.

Rust, B. W., 2001b: Fitting nature's basic functions Part II: Estimating uncertainties and testing hypotheses. *Computing in Science & Eng.*, **3 (6)**, 60–64, doi: 10.1109/5992.963429.

Rust, B. W., 2002: Fitting nature's basic functions Part III: Exponentials, sinusoids, and nonlinear least squares. *Computing in Science & Eng.*, **4 (4)**, 72–77, doi: 10.1109/MCISE.2002.1014982.

Rust, B. W., 2003: Fitting nature's basic functions Part IV: The variable projection algorithm. *Computing in Science & Eng.*, **5 (2)**, 74–79, doi: 10.1109/MCISE.2003.1182965.

Rybicki, G., and A. Lightman, 2008: *Radiative Processes in Astrophysics.* Physics textbook, Wiley.

Saunois, M., and Coauthors, 2020: The Global Methane Budget $2000 - 2017$. *Earth Syst. Sci. Data*, **12 (3)**, 1561–1623, doi: 10.5194/essd-12-1561-2020.

Schepers, D., and Coauthors, 2012: Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: Performance comparison of proxy and physics retrieval algorithms. *J. Geophys. Res.*, **117 (D10)**, D10 307, doi: 10.1029/2012JD017549.

Schmidt, D., 2014: Carbon Monoxide from SCIAMACHY Near Infrared Nadir Spectra: Impact of Retrieval Settings. Master thesis, TU München.

Schneider, M., F. Hase, J.-F. Blavier, G. Toon, and T. Leblanc, 2011: An empirical study on the importance of a speed-dependent Voigt line shape model for tropospheric water vapor profile remote sensing. *J. Quant. Spectrosc. & Radiat. Transfer*, **112 (3)**, 465–474, doi: 10.1016/j.jqsrt.2010.09.008.

Schreier, F., 2011: Optimized implementations of rational approximations for the Voigt and complex error function. *J. Quant. Spectrosc. & Radiat. Transfer*, **112 (6)**, 1010–1025, doi: 10.1016/j.jqsrt.2010.12.010.

Schreier, F., 2016: An assessment of some closed-form expressions for the Voigt function. *J. Quant. Spectrosc. & Radiat. Transfer*, **176**, 1–5, doi: 10.1016/j.jqsrt.2016.02.016.

Schreier, F., 2017: Computational aspects of speed-dependent Voigt profiles. *J. Quant. Spectrosc. & Radiat. Transfer*, **187**, 44–53, doi: 10.1016/j.jqsrt.2016.08.009.

Schreier, F., 2018: The Voigt and complex error function: Humlíček's rational approximation generalized. *Mon. Not. Roy. Astron. Soc.*, **479 (3)**, 3068–3075, doi: 10.1093/mnras/sty1680.

Schreier, F., S. Gimeno García, P. Hedelt, M. Hess, J. Mendrok, M. Vasquez, and J. Xu, 2014: GARLIC – a general purpose atmospheric radiative transfer line-by-line infrared-microwave code: Implementation and evaluation. *J. Quant. Spectrosc. & Radiat. Transfer*, **137**, 29–50, doi: 10.1016/j.jqsrt.2013.11.018.

Schreier, F., and P. Hochstaffl, 2021: Computational aspects of speed-dependent Voigt

and Rautian profiles. *J. Quant. Spectrosc. & Radiat. Transfer*, **258**, 107 385, doi: https://doi.org/10.1016/j.jqsrt.2020.107385.

Seager, S., 2010: *Exoplanet Atmospheres: Physical Processes.* Princeton Series in Astrophysics, Princeton University Press.

Serov, E., T. Odintsova, M. Tretyakov, and V. Semenov, 2017: On the origin of the water vapor continuum absorption within rotational and fundamental vibrational bands. *J. Quant. Spectrosc. & Radiat. Transfer*, **193**, 1 – 12, doi: 10.1016/j.jqsrt.2017.02.011.

Shine, K., I. Ptashnik, and G. Rädel, 2012: The water vapour continuum: Brief history and recent developments. *Surv. Geophys.*, **33 (3-4)**, 535–555, doi: 10.1007/s10712-011-9170-y.

Smith, E. W., 1981: Absorption and dispersion in the O_2 microwave spectrum at atmospheric pressures. *J. Chem. Phys.*, **74**, 6658–6673.

Sneep, M., and W. Ubachs, 2005: Direct measurement of the Rayleigh scattering cross section in various gases. *J. Quant. Spectrosc. & Radiat. Transfer*, **92 (3)**, 293 – 310, doi: 10.1016/j.jqsrt.2004.07.025.

Solimini, D., 2016: *Understanding Earth Observation: The Electromagnetic Foundation of Remote Sensing.* Remote Sensing and Digital Image Processing, Springer International Publishing.

Stamnes, K., G. Thomas, and J. Stamnes, 2017: *Radiative Transfer in the Atmosphere and Ocean.* Cambridge University Press.

Strandgren, J., and Coauthors, 2020: Towards spaceborne monitoring of localized CO_2 emissions: an instrument concept and first performance assessment. *Atmos. Meas. Tech.*, **13 (6)**, 2887–2904, doi: 10.5194/amt-13-2887-2020.

Strow, L. L., D. C. Tobin, and S. E. Hannon, 1994: A compilation of first-order line-mixing coefficients for CO_2 Q-branches. *J. Quant. Spectrosc. & Radiat. Transfer*, **52 (3)**, 281–294, doi: https://doi.org/10.1016/0022-4073(94)90158-9.

Struve, W., 1989: *Fundamentals of Molecular Spectroscopy.* Wiley.

Sussmann, R., and M. Buchwitz, 2005: Initial validation of ENVISAT/SCIAMACHY columnar CO by FTIR profile retrievals at the Ground-Truthing Station Zugspitze. *Atm. Chem. Phys.*, **5**, 1497–1503, doi: 10.5194/acp-5-1497-2005.

Suto, H., and Coauthors, 2021: Thermal and near-infrared sensor for carbon observation Fourier transform spectrometer-2 (TANSO-FTS-2) on the Greenhouse gases Observing SATellite-2 (GOSAT-2) during its first year in orbit. *Atmos. Meas. Tech.*, **14 (3)**, 2013–2039, doi: 10.5194/amt-14-2013-2021.

Taylor, T. E., and Coauthors, 2020: OCO-3 early mission operations and initial (vEarly) XCO_2 and SIF retrievals. *Remote Sensing of Environment*, **251**, 112 032, doi: 10.1016/j.rse.2020.112032.

Tennyson, J., 2005: *An Introduction to the Atomic and Molecular Physics of Astronomical Spectra.* Astronomical Spectroscopy, Imperial College Press.

Tennyson, J., and Coauthors, 2014: Recommended isolated-line profile for representing

high-resolution spectroscopic transitions (IUPAC technical report). *Pure Appl. Chem.*, **86 (12)**, 1931–1943, doi: 10.1515/pac-2014-0208.

Thalman, R., K. J. Zarzana, M. A. Tolbert, and R. Volkamer, 2014: Rayleigh scattering cross-section measurements of nitrogen, argon, oxygen and air. *J. Quant. Spectrosc. & Radiat. Transfer*, **147**, 171–177, doi: 10.1016/j.jqsrt.2014.05.030.

Tran, H., J.-M. Hartmann, G. Toon, L. Brown, C. Frankenberg, T. Warneke, P. Spietz, and F. Hase, 2010: The $2\nu_3$ band of CH_4 revisited with line mixing: Consequences for spectroscopy and atmospheric retrievals at $1.67\,\mu$m. *J. Quant. Spectrosc. & Radiat. Transfer*, **111**, 1344–1356, doi: 10.1016/j.jqsrt.2010.02.015.

Tran, H., N. Ngo, and J.-M. Hartmann, 2013: Efficient computation of some speed-dependent isolated line profiles. *J. Quant. Spectrosc. & Radiat. Transfer*, **129**, 199 – 203, doi: 10.1016/j.jqsrt.2013.06.015, erratum: JQSRT 134, 104 (2014).

Tran, H., N. Ngo, and J.-M. Hartmann, 2014: Erratum to "efficient computation of some speed-dependent isolated line profiles" [j. quant. spectrosc. radiat. transfer 129 (2013) 199-203]". *J. Quant. Spectrosc. & Radiat. Transfer*, **134**, 104, doi: 10.1016/j.jqsrt.2013.10.015.

Tran, H., B. Picquet-Varrault, C. Boursier, C. Viatte, M. Eremenko, F. Hase, and J.-M. Hartmann, 2011: Non-Voigt line-shape effects on retrievals of atmospheric ozone: Line-mixing effects. *J. Quant. Spectrosc. & Radiat. Transfer*, **112 (14)**, 2287–2295, doi: 10.1016/j.jqsrt.2011.06.001.

United Nations Framework Convention on Climate, 2015: Paris Agreement to the United Nations Framework Convention on Climate Change. Accessed: 2021-04-01, https://unfccc.int/files/meetings/paris_nov_2015/application/pdf/paris_agreement_english_.pdf.

van Hees, R. M., and Coauthors, 2018: Determination of the TROPOMI-SWIR instrument spectral response function. *Atmos. Meas. Tech.*, **11 (7)**, 3917–3933, doi: 10.5194/amt-11-3917-2018.

van Kempen, T. A., R. M. van Hees, P. J. J. Tol, I. Aben, and R. W. M. Hoogeveen, 2019: In-flight calibration and monitoring of the Tropospheric Monitoring Instrument (TROPOMI) short-wave infrared (SWIR) module. *Atmos. Meas. Tech.*, **12 (12)**, 6827–6844, doi: 10.5194/amt-12-6827-2019.

Vargas-Rodríguez, E., and H. Rutt, 2009: Design of CO, CO_2 and CH_4 gas sensors based on correlation spectroscopy using a Fabry-Perot interferometer. *Sensors and Actuators B: Chemical*, **137 (2)**, 410 – 419, doi: 10.1016/j.snb.2009.01.013.

Varghese, P., and R. Hanson, 1984: Collisional narrowing effects on spectral line shapes measured at high resolution. *Appl. Opt.*, **23 (14)**, 2376–2385, doi: 10.1364/AO.23.002376.

Varon, D. J., D. J. Jacob, J. McKeever, D. Jervis, B. O. A. Durak, Y. Xia, and Y. Huang, 2018: Quantifying methane point sources from fine-scale satellite observations of atmospheric methane plumes. *Atmos. Meas. Tech.*, **11 (10)**, 5673–5686, doi: 10.5194/amt-11-5673-2018.

Veefkind, J., and Coauthors, 2012: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, **120**, 70 – 83, doi: 10.1016/j.rse.2011.09.027, the Sentinel Missions - New Opportunities for Science.

Vonk, F., 2017: Input/output data specification for the TROPOMI L01b data processor. Technical report, https://sentinel.esa.int/documents/247904/3119978/Sentinel-5P-Level-01B-input-output-data-specification, The Royal Netherlands Meteorological Institute KNMI.

Warner, J., M. Comer, C. Barnet, W. McMillan, W. Wolf, E. Maddy, and G. Sachse, 2007: A comparison of satellite tropospheric carbon monoxide measurements from AIRS and MOPITT during INTEX-A. *J. Geophys. Res.*, **112**, D12S17, doi: 10.1029/2006JD007925.

Weideman, J., 1994: Computation of the complex error function. *SIAM J. Num. Anal.*, **31**, 1497–1518, doi: 10.1137/0731077.

Weinberg, S., 2015: *Lectures on Quantum Mechanics*. Cambridge University Press.

Wu, L., O. Hasekamp, H. Hu, J. aan de Brugh, J. Landgraf, A. Butz, and I. Aben, 2019: Full-physics carbon dioxide retrievals from the OCO-2 satellite by only using the 2.06 μm band. *Atmos. Meas. Tech. Disc.*, **2019**, 1–15, doi: 10.5194/amt-2019-218.

Wu, L., and Coauthors, 2018: Carbon dioxide retrieval from OCO-2 satellite observations using the RemoTeC algorithm and validation with TCCON measurements. *Atmos. Meas. Tech.*, **11 (5)**, 3111–3130, doi: 10.5194/amt-11-3111-2018.

Wunch, D., and Coauthors, 2010: Calibration of the Total Carbon Column Observing Network using aircraft profile data. *Atmos. Meas. Tech.*, **3 (5)**, 1351–1362, doi: 10.5194/amt-3-1351-2010.

Wunderlich, F., M. Scheucher, M. Godolt, J. Grenfell, F. Schreier, and H. Rauer, 2021: Detectability of biosignatures on LHS1140b. *Astron. & Astrophys.*, **647**, A48, doi: 10.1051/0004-6361/202039663.

Yan, F., R. Fosbury, M. Petr-Gotzens, G. Zhao, W. Wang, L. Wang, Y. Liu, and E. Pallé, 2015: High-resolution transmission spectrum of the Earth's atmosphere — seeing Earth as an exoplanet using a lunar eclipse. *Int. J. Astrobiology*, **14**, 255–266, doi: 10.1017/S1473550414000172.

Yue, X.-L., and Q.-X. Gao, 2018: Contributions of natural systems and human activity to greenhouse gas emissions. *Adv. in Clim. Change Res.*, **9 (4)**, 243–252, doi: 10.1016/j.accre.2018.12.003.

Zdunkowski, W., T. Trautmann, and A. Bott, 2007: *Radiation in the Atmosphere: A Course in Theoretical Meteorology*. Cambridge University Press.

Zhang, F., 2011: *Matrix Theory: Basic Results and Techniques*. Universitext, Springer New York.

Publication I

remote sensing

Article

Validation of Carbon Monoxide Total Column Retrievals from SCIAMACHY Observations with NDACC/TCCON Ground-Based Measurements

Philipp Hochstaffl [1,*] , **Franz Schreier** [1] , **Günter Lichtenberg** [1]
and **Sebastian Gimeno García** [1,2]

[1] DLR—German Aerospace Center, Remote Sensing Technology Institute, 82234 Oberpfaffenhofen, Germany;
    franz.schreier@dlr.de (F.S.); guenter.lichtenberg@dlr.de (G.L.); sebastian.gimenogarcia@eumetsat.int (S.G.G.)
[2] EUMETSAT—European Organisation for the Exploitation of Meteorological Satellites,
    64283 Darmstadt, Germany
*   Correspondence: philipp.hochstaffl@dlr.de; Tel.: +49-8153-28-3056

Abstract: The objective was to validate the carbon monoxide (CO) total column product inferred from Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) full-mission (2003–2011) short-wave infrared (SWIR) nadir observations using the Beer InfraRed Retrieval Algorithm (BIRRA). Globally distributed Network for the Detection of Atmospheric Composition Change (NDACC) and Total Carbon Column Observing Network (TCCON) ground-based (g-b) measurements were used as a true reference. Weighted averages of SCIAMACHY CO observations within a circle around the g-b observing system were utilized to minimize effects due to spatial mismatch of space-based (s-b) and g-b observations, i.e., disagreements due to representation errors rather than instrument and/or algorithm deficiencies. In addition, temporal weighted averages were examined and then the unweighted (classical) approach was compared to the weighted (non-classical) method. The delivered distance-based filtered SCIAMACHY data were in better agreement with respect to CO averages as compared to square-shaped sampling areas throughout the year. Errors in individual SCIAMACHY retrievals have increased substantially since 2005. The global bias was determined to be in the order of -10 parts per billion in volume (ppbv) depending on the reference network and validation strategy used. The largest negative bias was found to occur in the northern mid-latitudes in Europe and North America, and was partly caused by insufficient a priori estimates of CO and cloud shielding. Furthermore, no significant trend was identified in the global bias throughout the mission. The global analysis of the CO columns retrieved by the BIRRA shows results that are largely consistent with similar investigations in previous works.

Keywords: SCIAMACHY; BIRRA; NDACC; TCCON; validation; retrieval; carbon monoxide; mixing ratios; weighted averages

1. Introduction

Obtaining space-based (s-b) measurements of the state of Earth's atmosphere is costly compared to many other atmospheric research activities. Once the instruments needed to take these measurements are in orbit, thorough verification and validation of the delivered data is required.

Therefore, considerable efforts have been invested in establishing a ground-based (g-b) validation infrastructure for atmospheric composition, temperature, cloud, and aerosol data acquired by satellite-based remote sensing instruments. Additionally, the stringent requirements of upcoming missions such as Sentinel-5P/4/5 require comprehensive validation campaigns and solid strategies.

1.1. The Environmental Satellite (ENVISAT) and Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) Instrument

In March 2002 the ENVISAT was launched into a low-Earth orbit. The routine operation comprising the nominal measurement program began in August 2002 and lasted until April 2012 when contact was lost [1]. The satellite comprised 10 instruments with the SCIAMACHY, the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), and Global Ozone Monitoring by Occultation of Stars (GOMOS) dedicated to studying the Earth's atmosphere. The SCIAMACHY was an ultra-violet (UV) to short-wave infrared (SWIR, 1.4–3 μm) absorption spectrometer that observed the scattered and reflected solar spectral radiance transmitted through the atmosphere [2].

The SCIAMACHY's major research objective was the acquisition of information from Earth radiance spectra of various trace gases in the troposphere and the stratosphere. Among these gases were O_3, CO, CH_4, H_2O, SO_2, CO_2, and NO_2. Thus, the SCIAMACHY comprised eight detectors in the wavelength range between 214 and 2386 nm. Channel 8 observed radiance in the SWIR spectral range from 2259.38 nm to 2386.07 nm (4426–4191 cm^{-1}) at a resolution of 0.26 nm, which is equivalent to a resolving power of 8689 to 9177.

1.2. Carbon Monoxide (CO) from Channel 8 of SCIAMACHY

In atmospheric remote sensing, trace gas concentrations are retrieved from measured radiance or transmission spectra. The accurate retrieval of CO from Channel 8 observations of the SCIAMACHY is demanding due to the low optical depth of CO compared to the total depth in this spectral region. To be more specific, for a vertical path through the atmosphere, only about one percent of absorption is due to CO molecules. In addition, the precision of the cloud-free SCIAMACHY measurements is strongly influenced by the reflectance (albedo) of the observed ground-pixel, because it determines to a large extent the signal-to-noise ratio (SNR) of the corresponding observed spectra. The retrieval of CO can therefore only deliver total column amounts without any information on the vertical distribution. In addition, the Channel 8 detector shows temporal degradation due to an accumulation of ice, which significantly affects the throughput. This makes gas retrieval an even more challenging task [3]. Moreover, pixel degradation due to, e.g., solar radiation, is an ongoing process. According to Lichtenberg et al. [4] Channel 8 contains a substantial number of dead and bad pixels, with 40% deemed unusable by June 2009. For the CO fitting window from 4280 to 4305 cm^{-1} the number of usable pixels was reduced to only about 50 in the period between 2003 and 2005, and to even fewer afterwards.

1.3. Retrieval Codes

The retrieval of trace gas concentrations from radiance or transmission spectra poses an inverse problem that is typically solved by least-squares algorithms. Several codes have been developed for SCIAMACHY nadir SWIR spectra at different European institutes, e.g., the Weighted Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS) algorithm [5,6], the Iterative Maximum A Posteriori (IMAP)-DOAS [7] method, the Iterative Maximum Likelihood Method (IMLM) [3], and the Beer InfraRed Retrieval Algorithm (BIRRA) [8]. Recently, SCIAMACHY spectra have also been processed with the Shortwave Infrared CO Retrieval (SICOR) algorithm developed for the operational data processing of the Tropospheric Monitoring Instrument (TROPOMI) that was launched on the European Space Agency's (ESA) Sentinel-5 Precursor (S5P) mission [9,10]. Concurrent validation is a crucial task in processor/algorithm development in order to confirm applicability for scientific tasks and ultimately qualify operational algorithms or scientific tools. Based on such studies, processors can be applied on new data sets from other platforms with reasonable confidence. However, new fields of application require further validation or verification efforts.

1.4. The Network for the Detection of Atmospheric Composition Change (NDACC) and Total Carbon Column Observing Network (TCCON) Ground Truthing Networks

Fourier transform spectroscopy (FTS) is a well-established technique in passive remote sensing and observes thermal emissions (e.g., using the Infrared Atmospheric Sounding Interferometer (IASI) [11] or MIPAS [12]) or absorption (e.g., using the Atmospheric Chemistry Experiment—Fourier Transform Spectrometer (ACE-FTS) [13] or the Greenhouse Gases Observing Satellite (GOSAT) [14], which also observes in the thermal infrared range (TIR range, 8–15 µm)). High-resolution FTS is also used by numerous g-b observatories of the NDACC and TCCON [15]. The instruments associated with these networks routinely record thermal emission spectra in the mid-infrared range (MIR range, 3–8 µm) (NDACC) and solar absorption spectra in the SWIR range (TCCON) at a number of stations worldwide (see Figure 1, Table 1) and utilize these to infer information about atmospheric constituent columns and concentration profiles (NDACC only), including CO.
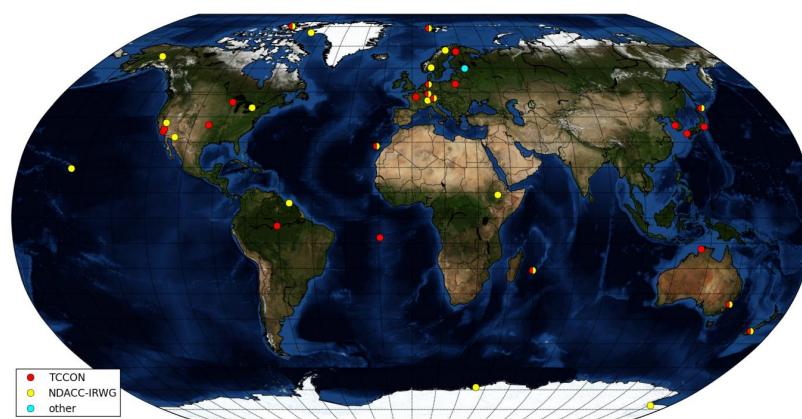

Figure 1. World map created by Feist [16] showing stations affiliated to ground-based (g-b) observing networks routinely measuring trace gases such as CO in the mid-infrared (MIR) and short-wave infrared (SWIR) ranges. The background color scheme also provides some information on the ground reflectance in the SWIR range (e.g., the variations of the reflectance on the continents and also the difference between land and the oceans).

Table 1. The Network for the Detection of Atmospheric Composition Change (NDACC) and Total Carbon Column Observing Network (TCCON) g-b Fourier transform infrared (FTIR) stations used in this validation. The last two columns indicate the time span of the g-b data used for the comparison to the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY). Further details see Appendix A and B.

Stations	Lat N [°]	Lon E [°]	Altitude [m]	NDACC (years)	TCCON (years)
Bialystok	53.23	23.03	160	-	2009–2011
Bremen	53.10	8.85	30	2003–2011	2005–2011
Darwin	−12.42	130.89	30	-	2005–2011
Garmisch	47.48	11.06	743	-	2007–2011
Izana	28.30	−16.48	2367	2003–2011	2007–2011
Jungfraujoch	46.55	7.98	3580	2003–2011	-
Kiruna	67.84	20.40	420	2003–2011	-
Lamont	36.60	−97.49	320	-	2008–2011
Lauder	−45.04	169.68	370	-	2004–2010
Ny Alesund	78.92	11.93	15	2003–2011	2005–2011
Parkfalls	45.95	−90.27	442	-	2004–2011
Thule	76.52	−68.77	220	2003–2011	-
Toronto	43.66	−79.40	174	2003–2011	-
Wollongong	−34.41	150.88	30	-	2008–2011
Zugspitze	47.42	10.98	2964	2003–2011	-

The TCCON is the reference network for the validation of greenhouse gas satellite retrievals and enables the linking of satellite retrievals to the World Meteorological Organization (WMO) reference scale [17]. In this study, the GGG2014 release of TCCON data is used. GGG stands for the whole software package. According to the TCCON website the error budget for CO is around 8.7% in total (see TCCON [18]). NDACC Fourier transform infrared (FTIR) data were obtained via the publicly available website (see http://www.ndacc.org).

1.5. Validation

In general terms, validation is defined as the process of evaluating the performance of a system against some equivalent information that is regarded as a 'true' reference [19]. In the field of remote sensing from space, compliance of observations from satellite platforms with the actual state of the atmosphere is usually assessed through validation studies. In that sense, validation means comparing measurements acquired by s-b instruments to other measurements utilizing different measuring methods, e.g., in-situ measurements or g-b soundings.

1.5.1. General Aspects

With respect to a direct comparison it stands to reason that the closer in time and space the reference measurements are acquired, the better they quantify differences due to errors in the acquisition or retrieval process, i.e., the instrument or algorithm performance. However, space-borne and reference measurements do not exactly match in time and space, nor do they address the same volumes of air. Hence, direct comparison of observations from different observing systems is affected by representation errors.

In fact, the column measured by the SCIAMACHY is an average column above the area covered by a SCIAMACHY pixel (32 $\times$120 km^2 for observations in Channel 8) which greatly extends beyond the location of the point-like g-b station. According to Verhoelst et al. [20], non-perfect co-location in space is therefore a consequence of both a difference in sampling (i.e., a satellite pixel center generally does not coincide exactly with a ground station), and a difference in the way each instrument has a smoothed perception of the real, non-homogeneous atmosphere. The air mass of sensitivity for the s-b nadir measurements contains the ground pixel footprint, an extension towards the satellite, and, especially in the SWIR range, an extension in the direction of the Sun. In the SWIR range, these extensions correspond to a good approximation to the optical light path between the Sun, the surface reflection point, and the sensor.

Moreover, with respect to the SCIAMACHY, Borsdorff et al. [9] mention another issue, whereby for low radiance scenes, the retrieval noise error for SCIAMACHY measurements can exceed 100% of the retrieved column. That is, the comparison of individual CO columns using the SCIAMACHY to a reference is not sufficient at all. According to Gimeno García et al. [8], De Laat et al. [21], and de Laat et al. [22] averaging over an ensemble of pixels is necessary to reduce the instrument-noise error.

Finally, all of these issues need to be considered in an appropriate validation strategy of the CO product inferred from a retrieval algorithm. For mutually exclusive aspects a compromise must be made between, on the one hand, the abundance of comparison pairs, and on the other hand, non-instrumental comparison errors due to non-perfect co-location in space and time of satellite and ground-based measurements.

1.5.2. The SCIAMACHY CO Product

Early validation studies of the CO product retrieved from SCIAMACHY SWIR measurements by g-b data have been conducted by Sussmann and Buchwitz [23], Sussmann et al. [24], and Dils et al. [25]. In recent years, additional validation efforts have been carried out by Borsdorff et al. [9,10], De Laat et al. [21], and Schneising et al. [26]. Borsdorff et al. [9] additionally included Tracer Model version 5 (TM5) data for large-scale (northern and southern hemispheric Africa) intercomparison and Measurement

of Ozone on Airbus In-service Aircraft (MOZAIC)/In-service Aircraft for Global Observing System (IAGOS) measurements.

In all of these studies, important aspects of the comparison of g-b and s-b data were presented. With respect to methodology, many considerations described in those articles are not only true for SCIAMACHY but are valid for all s-b to g-b comparisons. The former studies particularly include validation of CO data against measurements from g-b FTIR spectrometers from either the NDACC, the TCCON, or both. As SCIAMACHY measurements do not provide any information about the vertical distribution of CO, it is reasonable to use TCCON columns for SCIAMACHY validation. In addition, integrals of vertical CO profiles provided by the NDACC are utilized for validation of the s-b total column retrievals. Usually, if total columns calculated from vertical profiles of different remote sounding instruments are to be compared, averaging kernels have to be taken into account. Nonetheless, due to the small differences in vertical sensitivity between g-b FTIR and SCIAMACHY, and because the representation error dominates, a direct comparison (i.e., a comparison without applying total column averaging kernels) of the columns was considered possible without introducing significant intercomparison errors [23].

So far, only Borsdorff et al. [9,10] have presented full-mission data set validation of SCIAMACHY CO vertical column densities employing the SICOR algorithm. Dils et al. [25] considered SCIAMACHY data that covered the January to December 2003 time period that were produced using three different retrieval algorithms, namely the WFM-DOAS algorithm (version 0.5 for CO), the IMAP-DOAS method (version 0.9 for CO), and the IMLM (version 6.3). Sussmann and Buchwitz [23] used CO vertical profile retrievals from g-b solar FTIR measurements at Zugspitze, Germany (January–October 2003) to validate columnar CO measurements retrieved from SCIAMACHY spectra with the WFM-DOAS algorithm version 0.4.

1.6. Goals and Structure of this Study

The objective of this study is to validate the BIRRA level 2 prototype processor using the full-mission level 1b SCIAMACHY data set. Both land and ocean pixels are used in this intercomparison. In addition, the study places emphasisis on mitigating errors induced by representation deficiencies due to averaging and analyses the effect of different comparison methodologies—a crucial step in thoroughly quantifying the performance (accuracy and precision) of a retrieval algorithm or instrument.

The paper is organized in five sections. Our methodology is described in Section 2 and the results in Section 3. Finally, the discussion and conclusion are provided in Sections 4 and 5.

2. Methodology

The analysis of a solar absorption spectrum essentially reveals the slant columns of the various absorbing gases. However, the dry air mole fractions are actually advantageous for comparison. These quantities are independent of surface pressure and humidity and are therefore much more useful for satellite validation. Therefore, before comparing the observations for validation, CO needs to be harmonized across the observing systems.

2.1. Ground-Based Product Definition

The vertical column density (VCD) in [mol cm^{-2}] is the number of molecules above the surface defined as

$$N_\mu = \int_{z_{srf}}^{z_{\text{TOA}}} n_\mu(z)\, dz\,. \tag{1}$$

with n_μ the number density [mol cm^{-3}] of molecule μ. The surface pressure p is given by

$$p = \int_{z_{srf}}^{z_{TOA}} m\, n(z)\, g(z)\, \mathrm{d}z$$

$$\approx m\, g \int_{z_{srf}}^{z_{TOA}} n(z)\, \mathrm{d}z \,, \tag{2}$$

where m is the molar mass and g the gravitational acceleration of the Earth. In Equation (2), $m\, g$ is the column-averaged value. Within the homosphere, roughly including the troposphere, stratosphere, and mesosphere, the molar mass of air and gravity is a good approximation constant.

The column mixing ratio (CMR) f_μ of molecule μ is related to the total column according to

$$N_\mu = f_\mu\, N_{air}^{wet}\,, \tag{3}$$

with $N_{air}^{wet} = \sum N_\mu + N_{H_2O} = N_{air} + N_{H_2O}$. As indicated above, it is useful to discuss dry CMR d_μ, i.e., $\mu \neq H_2O$ rather than the true (wet) mixing ratios because dry CMR is a better tracer, not being subject to strongly varying H_2O variations. As

$$\sum_\mu f_\mu + f_{H_2O} = 1 \tag{4}$$

$$\sum_\mu d_\mu = 1 \tag{5}$$

the mole fraction and dry air mole fraction are connected according to

$$f_\mu = d_\mu\, (1 - f_{H_2O}). \tag{6}$$

Note that the molar masses of $m_{H_2O} = 18.02$ g mol^{-1} and $m_{air} = 28.96$ g mol^{-1} are different. $\overline{m} = m_{air}\, (1 - f_{H_2O}) + f_{H_2O}\, m_{H_2O}$ accounts for this by correcting the mass for the dry air mixing ratio with the fraction of the H_2O mixing ratio, as H_2O contributes less to the surface pressure p on a per molecule basis than air. All this together leads to Equation (3), reading as

$$N_\mu = d_\mu\, (1 - f_{H_2O})\, \frac{p}{\overline{m}\, g}\,, \tag{7}$$

and therefore

$$d_\mu = \frac{N_\mu}{N_{air}^{wet} - N_{H_2O}} = \frac{N_\mu}{N_{air}} \quad . \tag{8}$$

Given that $f_{H_2O} \ll 1$, Equation (7) can be approximated according to

$$d_\mu = N_\mu \left(\frac{p}{m_{air}\, g} - N_{H_2O}\, \frac{m_{H_2O}}{m_{air}} \right)^{-1}\,, \tag{9}$$

which constitutes the dry air volume mixing ratio for gas μ. The dry air CMR d_μ is frequently designated as xCO and given in parts per billion in volume (ppbv).

2.1.1. NDACC

NDACC sites retrieve vertical profiles for a variety of molecules (see Table 2). Since the actual measurement obtains the slant column and not the vertical column, $\zeta = \cos(\theta)$ accounts for the local zenith angle $0 \leq \theta < \frac{\pi}{2}$. $\frac{1}{\zeta}$ is commonly known as the geometrical airmass factor. Therefore, when

calculating CO dry air column mixing ratio (xCO) for the respective site, Equation (9) has to be taken into account according to

$$\mathrm{xCO} = N_{\mathrm{CO}} \left(\frac{1}{\zeta} \frac{p}{m_{\mathrm{air}} \, g} - N_{\mathrm{H_2O}} \frac{m_{\mathrm{H_2O}}}{m_{\mathrm{air}}} \right)^{-1}. \tag{10}$$

Table 2. Center wavelengths (for the NDACC) and spectral windows (for the TCCON and Beer InfraRed Retrieval Algorithm, BIRRA) used for the retrieval of CO. The NDACC relies on narrow spectral fitting regions [27], while the retrieval strategy of the TCCON prefers much wider spectral regions [28]. The BIRRA uses one spectral fitting window in Channel 8. The NDACC uses the Scale Invariant Feature Transform (SFIT) algorithm, and the TCCON applies GFIT to acquire the GGG2014 data set.

Algorithm	Wavenumber [cm^{-1}]	Interfering Species
SFIT	2057.8575	NO, COF$_2$ ...
(NDACC)	2069.6559	NO, COF$_2$ ...
	2111.5430	OCS, N$_2$ ...
	2158.2997	OCS, N$_2$ ...
GFIT	4208.70–4257.30	H$_2$O, CH$_4$
(TCCON)	4262.00–4318.80	H$_2$O, CH$_4$
BIRRA	4280.00–4305.00	H$_2$O, CH$_4$
(SCIAMACHY)		

2.1.2. The TCCON

According to Wunch et al. [15], the TCCON scales a priori profiles of retrieved molecules

$$n_\mu(z) = \alpha_\mu \, n_\mu^{\mathrm{ref}}(z) \tag{11}$$

similarly to the BIRRA. Hence, the shape of the vertical profile is invariant with respect to variations in molecular atmospheric densities. For TCCON sites, Equation (9) is used in addition to CO with respect to the co-retrieved oxygen O_2. Assuming a dry air mole fraction of $d_{O_2} = 0.2095$, this is equivalent to the dry air column mixing ratio (details see Wunch et al. [29]). Rationing the CO and O_2 equations eliminates the denominator in Equation (9) to make d_{CO} independent of surface pressure and gravity, leading to the equation

$$\mathrm{xCO} = 0.2095 \, \frac{N_{\mathrm{CO}}}{N_{O_2}}. \tag{12}$$

2.2. The BIRRA

The BIRRA was developed at the Deutsches Zentrum für Luft- und Raumfahrt (DLR) and serves the operational SCIAMACHY processor for CO and CH$_4$ retrievals using Channel 8 and Channel 6 spectra, respectively. The CO total column amounts are inferred simultaneously with methane (CH$_4$) and water vapour columns and a Lambertian surface albedo from individual SCIAMACHY measurements assuming a non-scattering atmosphere. The validation study reported here is based on the BIRRA prototype version.

2.2.1. Algorithm

The BIRRA comprises the line-by-line forward model Generic Atmospheric Radiation Line-by-line InfraRed Code (GARLIC) [30,31] coupled to a least-squares [32] inversion algorithm for trace gas retrieval in the SWIR spectral region(see Gimeno García et al. [8]). In case of SWIR nadir observations,

the radiative transfer equation through Earth's atmosphere with molecules μ, including H_2O, reduces to Beer's law, describing the radiance

$$
\begin{aligned}
I(\mathbf{x}, \nu) \quad &\equiv \quad \frac{r(\nu)}{\pi} \mu_\odot I_{\text{sun}}(\nu) \times \exp\left(-\sum_\mu \alpha_\mu \, \tau_\mu^{\text{ref}}(\nu)\right) \\
&\otimes S(\nu, \gamma, \delta) + b(\nu).
\end{aligned}
\tag{13}
$$

$\mu_\odot = \cos\theta_\odot$ and I_{sun} describe the solar zenith angle and incoming radiation from the sun, respectively. The 'true' (= to be estimated by the fit) optical depth $\tau_\mu(\nu)$ is hence given by

$$
\tau_\mu(\nu) \ = \ \alpha_\mu \, \tau_\mu^{\text{ref}}.
\tag{14}
$$

For the BIRRA, the state vector $\mathbf{x} = (\boldsymbol{\eta}, \boldsymbol{\beta})$ includes nonlinear ($\boldsymbol{\eta}$) and linear ($\boldsymbol{\beta}$) parameters to be estimated. This separation of parameters enables the algorithm to utilize a separable least-squares fit—also known as variable projection [33]—in order to estimate the unknown quantities. The molecular scaling factor(s) α_μ, the half width γ and the wavelength shift δ of the instrumental slit function S are estimated using a nonlinear least-squares fit; furthermore, the coefficients of the surface albedo r, modeled by a second-order polynomial depending on wavenumber, and the optional baseline correction b (again a polynomial but not used in this study) show up linearly within the model and are estimated using a linear least-squares fit. The model for the least-squares problem is therefore

$$
y_i \ = \ \sum_{j=1}^{n} \phi_{i,j}(\boldsymbol{\eta}) \, \beta_j, \qquad i = 1, 2, \ldots, m
\tag{15}
$$

with n representing the numbers of linear parameters. m is the number of observations (spectral pixels) in the microwindow chosen for the retrieval. Notice that in the separable least-squares approach it is assumed that the model functions $\phi_{i,j}(\boldsymbol{\eta})$ depend on the nonlinear parameters $\boldsymbol{\eta}$, but not on the linear parameters $\boldsymbol{\beta}$. Thus, for any given $\boldsymbol{\eta}$ a matrix $\boldsymbol{\Phi}(\boldsymbol{\eta})$ comprising the model functions, $\phi_{i,j}(\boldsymbol{\eta})$ is defined according to

$$
\boldsymbol{\Phi}(\boldsymbol{\eta}) \ = \
\begin{pmatrix}
\phi_{1,1}(\boldsymbol{\eta}) & \phi_{1,2}(\boldsymbol{\eta}) & \cdots & \phi_{1,n}(\boldsymbol{\eta}) \\
\phi_{2,1}(\boldsymbol{\eta}) & \phi_{2,1}(\boldsymbol{\eta}) & \cdots & \phi_{2,n}(\boldsymbol{\eta}) \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{m,1}(\boldsymbol{\eta}) & \phi_{m,2}(\boldsymbol{\eta}) & \cdots & \phi_{m,n}(\boldsymbol{\eta})
\end{pmatrix}.
\tag{16}
$$

Now the objective function to be minimized with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$ is given by

$$
\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\eta}}) \ = \ \min_{\boldsymbol{\eta}} \left\{ \min_{\boldsymbol{\beta}} \left\{ \left\| \mathbf{y} - \boldsymbol{\Phi}(\boldsymbol{\eta})\boldsymbol{\beta} \right\|^2 \right\} \right\},
\tag{17}
$$

where the inner minimization problem forms a linear least-squares problem. The overall minimization problem can therefore be represented only in terms of the nonlinear parameters $\boldsymbol{\eta}$ according to

$$
\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\eta}}) \ = \ \min_{\boldsymbol{\eta}} \left\{ \left\| \mathbf{y} - \boldsymbol{\Phi}(\boldsymbol{\eta})\boldsymbol{\Phi}^\dagger(\boldsymbol{\eta})\mathbf{y} \right\|^2 \right\},
\tag{18}
$$

with the generalized inverse $\boldsymbol{\Phi}^\dagger(\boldsymbol{\eta}) = \left[\boldsymbol{\Phi}^T(\boldsymbol{\eta})\boldsymbol{\Phi}(\boldsymbol{\eta})\right]^{-1}\boldsymbol{\Phi}^T(\boldsymbol{\eta})$. The advantages of the separable least-squares approach are that for the linear parameters $\boldsymbol{\beta}$ no initial estimate is necessary and the size of the Jacobian matrix is reduced which improves the condition and minimizes computing times. Only for nonlinear parameters such as the scale factors of the optical depth of molecule μ is a priori information on the number density (see Equation (1)) required. For CO and CH_4 this information is taken from the Air Force Geophysical Laboratory (AFGL) atmospheric constituent profiles [34],

while H_2O and the auxiliary parameter temperature and pressure are taken from reanalysis data provided by the National Center for Environmental Prediction (NCEP).

2.2.2. Product Definition

Recall that in order to obtain atmospheric CO abundances, the BIRRA scales a representative CO reference profile. In general, to calculate the actual wet CMR of CO the inferred scaling factors are used according to

$$f_{CO} = \frac{\alpha_{CO} \, N_{CO}^{ref}}{\alpha_{air} \, N_{air}^{ref} + \alpha_{H_2O} \, N_{H_2O}^{ref}} \quad . \tag{19}$$

However, there is actually no parameter α_{air} showing up in the retrieval to account for the dry air CMR. Instead, differences between the retrieved and a priori CH_4 columns are used as a proxy to account for variations in the dry air column density according to

$$N_{CO} = \frac{\alpha_{CO}}{\alpha_{CH_4}} \, N_{CO}^{ref}. \tag{20}$$

This approach assumes that CH_4 is a well-mixed gas and that natural CH_4 variations in the atmosphere are small compared to variations of CO, with CH_4 changes produced solely by light path modifications (mainly by cloud shielding, for details see Gloudemans et al. [35]). Since CH_4 has strong absorption lines across the CO spectral fitting window in Channel 8, which allow is to determine the amount of CH_4 with good accuracy, this method is also suitable to detect optically thick clouds in the SCIAMACHY observations. In general, the error of the CH_4 scaling factors is 1–2 orders of magnitude smaller than the error of the CO scaling factors, depending on the scene. Intending to eliminate most of the uncertainties arising in the level-2 processing, postprocessing included the composition of quality criteria based on multiple parameters from the BIRRA output. For example, non-converging fits were filtered out, and errors of CO, H_2O and CH_4 were used to eliminate data with extremely low signal-to-noise ratios. This was done particularly for measurements on the ocean benefit from the presence of clouds due to the low reflectance of water in the SWIR range (see Figure 1, Gloudemans et al. [35]). This combination of filter criteria allows for the selection of optically thick clouds over the oceans and both cloudless scenes with acceptable small errors and cloudy scenes over land. If the light path was enhanced by,for example, aerosols, and exceeded a certain threshold (a 10% the enhancement of the CH_4 scaling factor), observations were rejected. Further details on the quality criteria chosen for the CO retrievals in this study are described in Gimeno García et al. [8] Section 2.3. The CO dry air CMR defined is therefore actually given by

$$xCO = d_{CO} = \frac{\alpha_{CO} \, N_{CO}^{ref}}{\alpha_{CH_4} \, N_{air}^{ref}} = \frac{\alpha_{CO}}{\alpha_{CH_4}} \, d_{CO}^{ref}. \tag{21}$$

2.3. Weighted Averages

It stands to reason that averaging over lots of measurements requires accounting for incorporating representation deficiencies. For example, if an average of 10 measurements is already deemed trustworthy, representation errors might not be an issue. However, in case of the SCIAMACHY's CO product, the number of measurements to be averaged for a representative mixing ratio (with respect to error of the mean) is at least in the order of 10^2.

A representative average value $\langle xCO \rangle$ for both s-b and g-b data with respect to space ρ and time τ is the weighted mean according to

$$\langle xCO \rangle = \frac{\sum\limits^{M} \omega \, xCO}{\sum\limits^{M} \omega}, \tag{22}$$

with M representing the number of observations, and $\omega = \rho^\kappa \tau$ the respective weight. The weights include a spatial $\rho = r_{gb} - r_{sb}$ and temporal $\tau = t_{gb} - t_{sb}$ component, with r and t designating location and time of the observation, respectively. The CO spatial variation of column density is much larger than that of CO_2 and CH_4 [36,37]. Many NDACC and TCCON sites are located in remote regions while some sites are located in polluted areas (Table 1). Since the validation aims to compare s-b observations for both background and enhanced CO levels, the exponent κ is introduced to account for the representation deficiencies introduced by the great variability of CO in space. Typically $\langle xCO \rangle$ is the monthly, seasonal, or annual mean. The corresponding standard deviation $\langle xCO \rangle_e$ is defined according to

$$\langle xCO \rangle_e = \sqrt{\frac{\sum\limits^{M} \omega \, (xCO - \langle xCO \rangle)^2}{\sum\limits^{M} \omega}} \, . \tag{23}$$

In general, if weighting is applied it needs to be applied on both datasets. Note that Equation (22) is valid with respect to both the spatial and the temporal domain.

2.3.1. Space

In the spatial domain the calculation of the averages for the SCIAMACHY data accounts for the position of acquisition relative to the location of the g-b data set. Therefore, $\tau = 1$ and $\omega = \rho^\kappa$. Note that since the location of the g-b reference site r_{gb} is fixed in time, $r_{gb} = const$.

2.3.2. Time

In order to calculate a representative average with respect to the temporal domain, the context was changed with $\omega = \tau$. In addition, a running average of t_{gb} needs to be introduced since t_{gb} is not constant. The interval selected for the running average is centered at the respective t_{sb} and chosen to extend the same time span as the corresponding weighted average. For example, a 30-day weighted average of the s-b CMR corresponds a running average of t_{gb} that incorporates g-b observations within $t_{sb} \pm 15$ days.

2.4. Bias

In order to quantify the retrieval accuracy per reference site, the bias was calculated as an error weighted offset according to

$$b = \sum\limits^{M} \langle xCO_{sb} \rangle_e^{-1} \left(\langle xCO_{sb} \rangle - \langle xCO_{gb} \rangle \right) / \sum\limits^{M} \langle xCO_{sb} \rangle_e \, , \tag{24}$$

with $\langle xCO_{sb} \rangle_e$ representing the standard deviation of the SCIAMACHY xCO in a certain time interval and M being the number of s-b and g-b averages (see Section 2.3), respectively.

In addition, the average of the s-b standard deviations at a specific site was used to characterize the accuracy of the bias b according to the standard error of the mean $s_e = \langle xCO_{sb} \rangle_e / \sqrt{M}$, with M the number of measurements incorporated in $\langle xCO_{sb} \rangle_e$. To estimate the overall performance, the global mean bias $\bar{b}$ was determined as the average of all station biases weighted by the standard deviation σ_b of the respective biases. $\bar{s}_e$ is the global (for all sites) standard error of the means. A similar approach was applied by Borsdorff et al. [9].

2.5. Averaging Multiple Years of CO

In order to exploit more observations within a given sampling area, multiannual averages for CO mixing ratios were calculated and refered to a common reference. Within the time interval 2003–2011, parameters for intercept a_n and slope b_n of xCO were estimated for each year n for both measurement systems and referred to a common reference. For this purpose, linear least-squares was applied.

Subsequently, the annual estimates for intercept and slope were averaged yielding the parameters $\bar{a}$ and $\bar{b}$. The detrended data points $\hat{x}CO$ were then calculated according to

$$\hat{x}CO \rightarrow xCO_n - (a_n - \bar{a}) - (b_n - \bar{b}) t_n. \tag{25}$$

The detrended observations $\hat{x}CO$ are largely independent with respect to the year of acquisition. Therefore more measurements belong to the common interval (e.g., a month, a year) compared to the single year analysis, and thus stricter thresholds can be imposed on filter criteria during post processing.

3. Results

The reference sites were selected with respect to the temporal data coverage and continuity of measurements as well as the reliability and completeness of auxiliary parameters such as surface pressure. The former two criteria were met if there were at least three years of observations covering the time period 2003–2011 with at least 10 months of observations. In case of NDACC stations all sites fulfilled this criteria, however, some TCCON sites such as Jet Propulsion Laboratory (JPL), Reunion Island, etc. had to be omitted. The latter criteria regarding auxiliary parameters posed an issue for some NDACC sites such as Arrival Heights, Lauder, or Wollongong. However, no TCCON sites were skipped due to this filter criterion.

3.1. Averaging of Measurements

The single most limiting factor in the global SCIAMACHY CO column product is its significant variability due to noise in the recorded spectrum. This clearly shows up in the CO values of individual observations in Figure 2. Furthermore, because Channel 8's pixel has a footprint on the Earth's surface of around 32×120 km^2 for the $180°$ nadir, a direct comparison of BIRRA-retrieved CO columns was not reasonable. It was found that even comprehensive filtering of the CO product based on the norm of the residuum ([8] Section 4) was not able to deliver the sufficient quality required for the comparison of a single or few (i.e., tens) of CO observations.

Figure 3 reveals that errors of individual SCIAMACHY retrievals have increased substantially since 2006. This seems to be primarily caused by omitting regular decontamination procedures from 2006 onward to dispose of the ice accumulated on the detector in Channel 8 (see [9] Table 1 and [2]). Furthermore, measurements conducted during this decontamination phase were commonly not suited for the adequate retrieval of parameters and most of them were filtered out during postprocessing. This effect is clearly visible in Figure 2 displaying the lack of proper CO retrievals at the turn of the years from 2003 through 2005. Also, the decreasing number of good spectral points (pixels) on the sensor plays a role (see [8] Figure 3).

Monthly-mean averages of CO mixing ratios constitute an adequate trade-off since the basic features of the CO annual cycle are preserved and at the same time they include enough measurements to get reasonable statistics for both SCIAMACHY and NDACC/TCCON observations (see Figure 4). Note that the validation studies cited in Section 1.5.2 also use a temporal interval of one month. Figure 4 also demonstrates that the adequate size of the temporal interval varies with the size of the sampling area. Moreover, the results indicate that the adequate selection depends on the quality of the s-b data and the problem investigated, hence also seasonal and annual averages were examined accordingly.
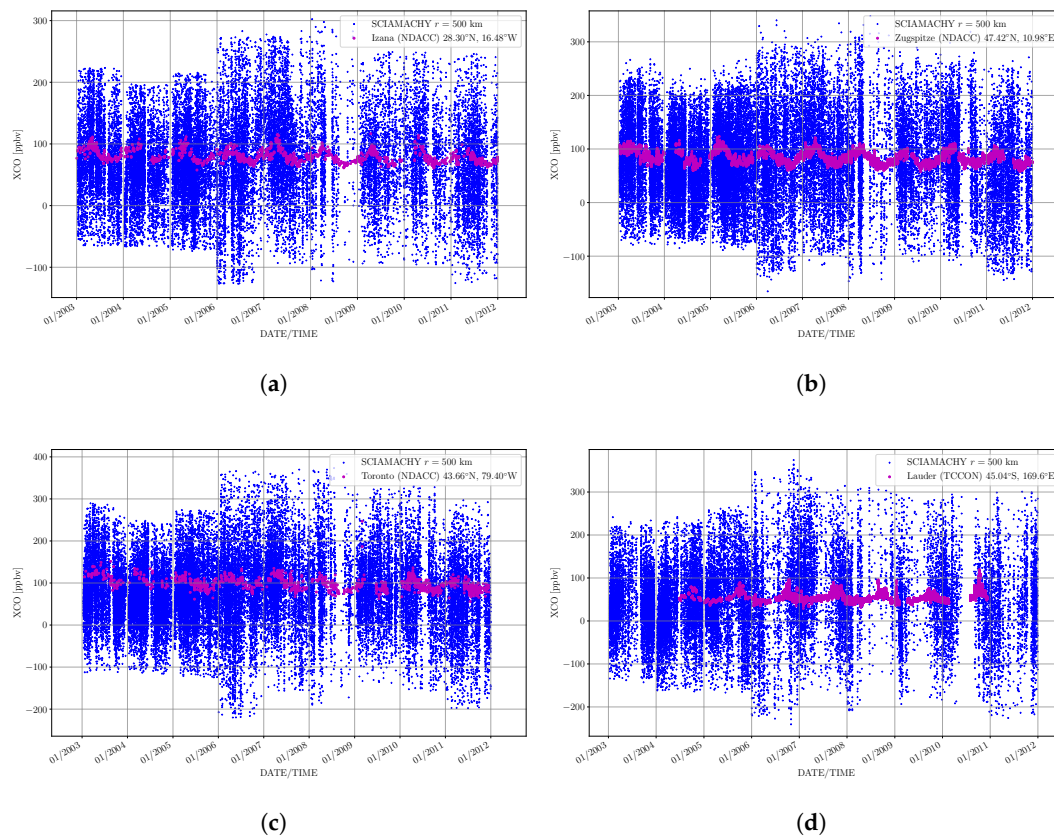
(a)

(b)

(c)

(d)

Figure 2. The SCIAMACHY and g-b reference site dry air column mixing ratios of carbon monoxide (xCO). The plot includes the column mixing ratios (CMRs) after postprocessing within 500 km of the reference sites (**a**) Izana, (**b**) Zugspitze, (**c**) Toronto and (**d**) Lauder from 2003 to 2011. The large scatter of the individual SCIAMACHY CO columns is mainly caused by measurement noise, while averaging of observations within 500 km of the reference site has only a minor contribution.

Since local enhancements shall be preserved, a related issue that also demands an adequate trade-off is spatial averaging of the s-b observations within a given radius from the g-b reference site. In order to include equidistant s-b measurements with respect to a reference site, latitude and longitude coordinates of the measurements were converted to the great circle distance on the Earth's surface. In a first step, the chosen distances for spatial averaging were 500 km, 1000 km, and 2000 km, since previous validation studies of SCIAMACHY CO used collocation criteria within this range (e.g., Borsdorff et al. [9], De Laat et al. [21], and Dils et al. [25]). The results in Figure 4 suggest that the smaller the sampling area, the larger the interval for temporal averaging that should be chosen. So far, only Sussmann et al. [24] sampled SCIAMACHY observations within 2000 km of Zugspitze. Other preceding validation studies validated s-b data within a square- or rectangle-shaped area around the g-b site. The effect of the different methodologies are displayed in Figure 5. Spatial filtering with respect to distance delivers a smoother course of s-b CO averages throughout the year.

(**a**)   (**b**)
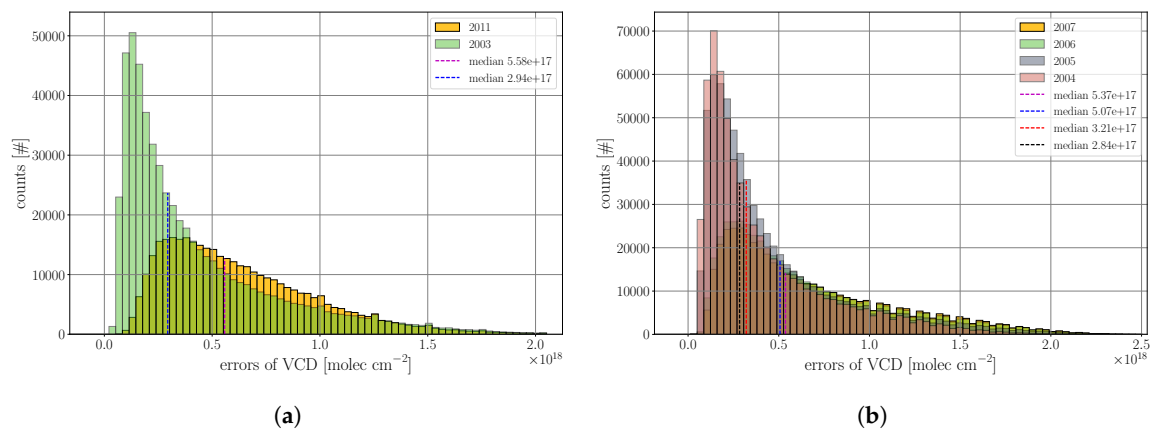
Figure 3. Histogram of the errors of the CO columns (defined in Equation (20)), represented as the unnormalized error probability function (also see Section 4.2.1 in Gimeno García et al. [8]). The data content is the global fraction of CO columns that survived after postprocessing of the BIRRA retrieval output. The dashed lines indicate the respective median. Note the different range in the x-axis. (**a**) The first and last year of the nominal SCIAMACHY operation covering the complete time span; (**b**) The occurrence frequency of total column errors for the years 2004 through 2007 indicate a substantial degradation of the quality of the CO product from 2006 onward (blue and magenta dashed lines). The years 2003 through 2005, however, show similar and significantly smaller errors.

(**a**)   (**b**)
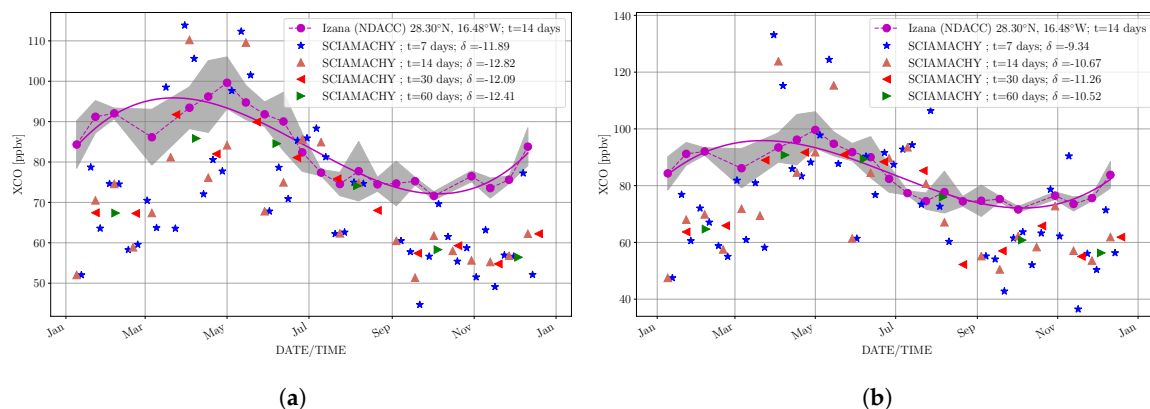
Figure 4. xCO averages for various temporal intervals within (**a**) 1000 km and (**b**) 500 km of Izana in 2003. δ designates the offset of the s-b averages to the 14-days mean of Izana. The gray shaded area and magenta solid line represent the standard deviation of the g-b means and the third-order polynomial fit on the individual observations, respectively. The SCIAMACHY xCO averages become smoother in larger temporal intervals and in the share of averages. This is in agreement within the standard deviation of the g-b site increases. Note the different range of the y-axis.

The SCIAMACHY observations within various-sized areas around two the NDACC and TCCON g-b sites, namely Bremen and Izana (see Figure 1), are compared in Figure 6 and the results demonstrate that the adequate spatial and temporal range of averaging depends on the precision of the s-b observations at a specific site. The former station is located in the plains of northern Germany, while the latter is situated on an island in the sub-tropical region in the Atlantic ocean at about 2370 m above mean sea level. With respect to Izana, a station largely surrounded by water, observations of optically thick clouds were preferred in order to get an acceptable SNR for the retrieval. Note that the bias *b* designated in the legend is the deviation of the s-b to the g-b CO averages weighted by the respective s-b standard deviation (details on the bias analysis in the subsequent section).
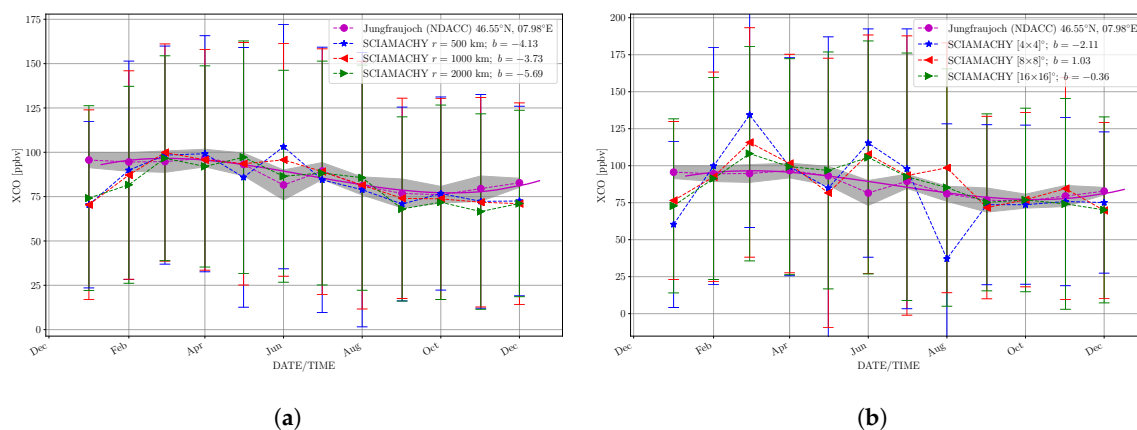
(a)　　　　　　　　　　　　　　　　　　(b)

Figure 5. xCO averages for circle- and square-shaped sampling areas. (**a**) Averages of SCIAMACHY xCO values calculated for various distances from the reference site Jungfraujoch in 2003; (**b**) xCO averages within given latitudes and longitudes of the g-b site. Here, $[4 \times 4]°$ designates a square-shaped object with an extent of $\pm 2°$ in latitude and longitude from the reference site. SCIAMACHY observations within the circle-shaped sampling areas show better agreement for most cases in 2003. Note the different range of the y-axis.

Another important aspect that needs to be considered is that for spatially weighted averages (used throughout the subsequent sections), the effective radius becomes seemingly smaller since close measurements account most while observations far away from the g-b measurement site have much weaker impact on the representative average. This basically means that in the case of weighted averages, the area significantly contributing to the average of s-b measurements is reduced. Therefore, we consider averaged CO values of 1 month and within 500 km of the g-b site as an appropriate trade-off for SCIAMACHY CO validation.

(a)　　　　　　　　　　　　　　　　　　(b)
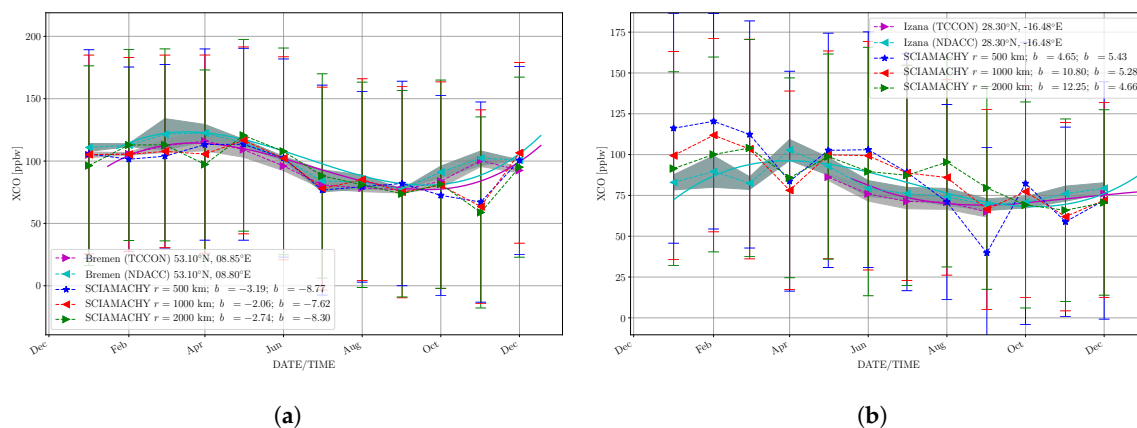
Figure 6. xCO monthly averages around the collocated NDACC and TCCON sites in 2007. (**a**) SCIAMACHY xCO values calculated for various distances from the reference site in Bremen. Measurements of the SCIAMACHY within 1000 km show the smallest bias *b*; (**b**) xCO averages with respect to Izana in 2003. On this site the SCIAMACHY observations within 500 km of the reference site show the smallest bias in total for the TCCON.

3.2. Bias and Weighting

The bias of SCIAMCHY observations to the g-b reference was analyzed using two different approaches. Initially, in Section 3.2.1, the classical unweighted averages for SCIAMACHY and g-b observations are used, while distance weighted averaging is treated in Section 3.2.2.

3.2.1. Unweighted Bias

Overall, Figure 7 shows that the bias of CO mixing ratios is ranging from almost 0 ppbv to -27 ppbv. In over 80% the biases range from around -5 to -15%. The global bias $\bar{b} = -12.1$ ppbv represents the mean bias of all stations weighted by the respective standard deviation. The global standard error of the mean $\bar{s}_e = 6.77$ ppbv is the average of all the stations standard errors of the mean.

In general, the results in Figure 7 agree with the findings by Borsdorff et al. [9,10] within around -5 to -10% for most sites. However, the CO mixing ratios calculated from the BIRRA retrieval have consistently low bias with respect to all sites in both networks, while Borsdorff et al. [9,10] found slightly positive biases at some sites. The data show different magnitudes of biases for different sites. Notice that some negative bias in the northern mid-latitudes in Europe and North America are significant with respect to standard deviation. These findings suggest that the comparison crucially depends on the CO a priori information which is not well characterized in polluted areas. Sites located in remote areas such as Ny Alesund, Izana, Wollongong, and Lauder do therefore show much better agreement (Figure 7).

(**a**)   (**b**)
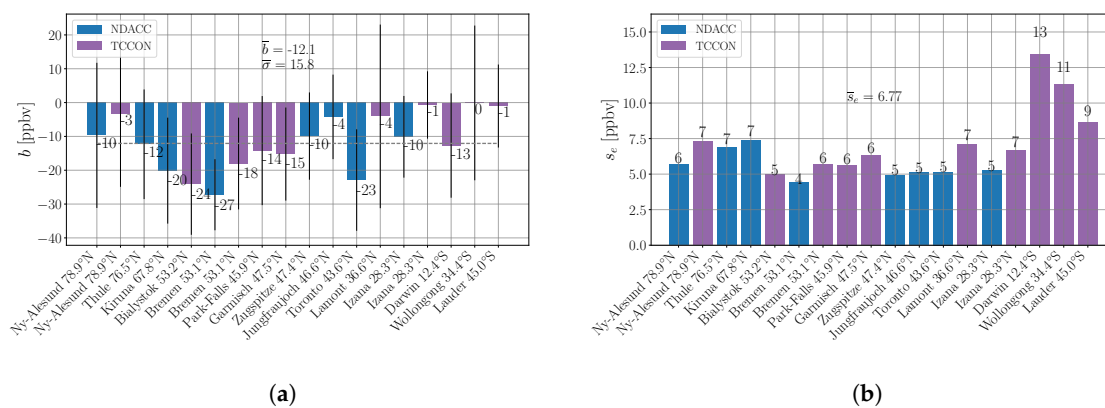
Figure 7. Mean bias of NDACC (blue) and TCCON (purple) stations with co-located SCIAMACHY CO retrievals from 2003 to 2011. Bias is the average of the monthly-mean differences weighted by the standard deviations. The total (2003–2011) bias per site was subsequently calculated using the annual weighted biases. (**a**) Global bias $\bar{b}$ and respective global standard error $\bar{\sigma}$. The dashed gray line indicates $\bar{b}$; (**b**) The standard error of the means s_e.

3.2.2. Distance Weighted Bias

The impact on xCO averages was studied using linear (r^{-1}) and quadratic (r^{-2}) inverse distance weighting of spatially close measurements (not to be confused with the error weighting of the bias also used in Section 3.2.1). In general, at most sites, the biases are decreased in the case of inverse distance weighting. This effect can be observed for both the linear inverse and quadratic inverse weighting (Figures 8 and 9). In the case of linear weighting only a few sites exhibit slightly larger biases, such as Kiruna (NDACC), Bialystok, Wollongong, and Lauder (TCCON). However, the overall bias was reduced to -11.2 ppbv while the average standard deviation was increased to 16.6 ppbv. Furthermore, most sites show that quadratic weighting (with respect to inverse distance of SCIAMACHY measurements) increases the agreement of the compared averaged values. The majority of biases observed in Figure 9 are reduced even more compared to the linear approach and the distribution throughout the sites from north to south is significantly smoothed. This could indicate that the results in the linear and quadratic weighted cases are better suited for the estimation of the instrument and retrieval deficiencies compared to the unweighted results which might include larger fractions of representation errors. In other words, it is likely that a fraction of the revealed offsets are attributed to mismatch artifacts and not only incorporate instrument or retrieval issues.

Favorable examples to study this effect more closely are g-b sites located in polluted areas, especially large cities embedded in rural areas such as Toronto. While the SCIAMACHY seems to significantly underestimate CO at Toronto in the unweighted example (Figure 7), the quadratic weighting reveals that the offset is not significant but rather suggests to be introduced by less representative observations. However, discrepancies at Kiruna, Bialystok, and Bremen in the northern mid-latitudes still seem to be significant. In the case of Bremen, the negative offset is significant with respect to both NDACC and TCCON observations, which underlines the finding of a negative bias of the SCIAMACHY SWIR observations. The global bias was reduced to 11.9 ppbv compared to the unweighted approach, but $\overline{\sigma}$ increased to 20.8 ppbv.

Table 3 reveals a discrepancy of the global bias between the NDACC and TCCON of 6.2 ppbv for the unweighted case and 5.3 ppbv in case of distance weighting (r^{-2}). In general the findings with respect to both networks are in accordance with the conclusions by Borsdorff et al. [9] in many aspects. The global $\overline{b}$ bias is larger for sites affiliated to the NDACC than those affilliated with the TCCON. With respect to the NDACC and TCCON, we found $\overline{b} \approx -14$ ppbv and $\overline{b} = -7.4$ up to -9.1, respectively, depending on the weighting while Borsdorff et al. [9] revealed values of -9.2 ppb and -1.2 ppb for the clear-sky retrievals for the NDACC and TCCON, respectively. In case of cloudy-sky retrievals, Borsdorff et al. [10] found a negative bias $\overline{b} = -6.0$ ppb which is similar to our results for the TCCON (see Table 3).

In order to underline the previous findings, it was important to confirm that the bias of the SCIAMACHY data with respect to the g-b reference did not exhibit a linear trend throughout the mission. Therefore, a bias trend t in ppbv/year was calculated for the annual biases b at each site using linear least-squares regression. The results are depicted in Figure 10. The global trend $\overline{t}$ is the average of the bias trends weighted by the standard error of the fit. In order to get a complete picture, initially all sites for all available years were investigated. It was found that the number of sites that reveal a negative bias trend exceeds the number of sites with a positive bias trend. Note that the trend is within ± 5 ppbv around most reference sites. In addition, for most sites no significant trend with respect to the accompanying standard error was identified. Most sites show a consistent standard error of about 5 ppbv/year in both the weighted and unweighted cases. Moreover, the magnitude of the standard error is in the order of (or even larger than) the calculated average trend at most sites (see Figure 10). The numbers in the figure indicate that the global trend of the bias is not significant for at least two cases. Hence, we decided that a linear trend correction of the results presented above was not required.
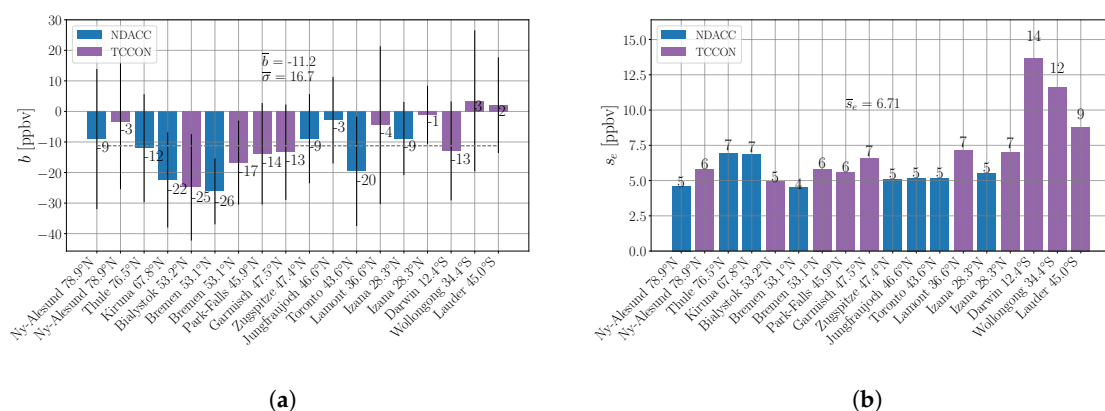

(**a**)　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 8. Mean bias, standard deviation, and standard error as in Figure 7 but with monthly-mean averages of SCIAMACHY CO weighted according to the inverse distance from the reference site. (**a**) Global monthly-mean bias $\overline{b}$ and respective standard error $\overline{\sigma}$; (**b**) The global standard error of the mean biases $\overline{s}_e$ is the average of all standard errors of the means.
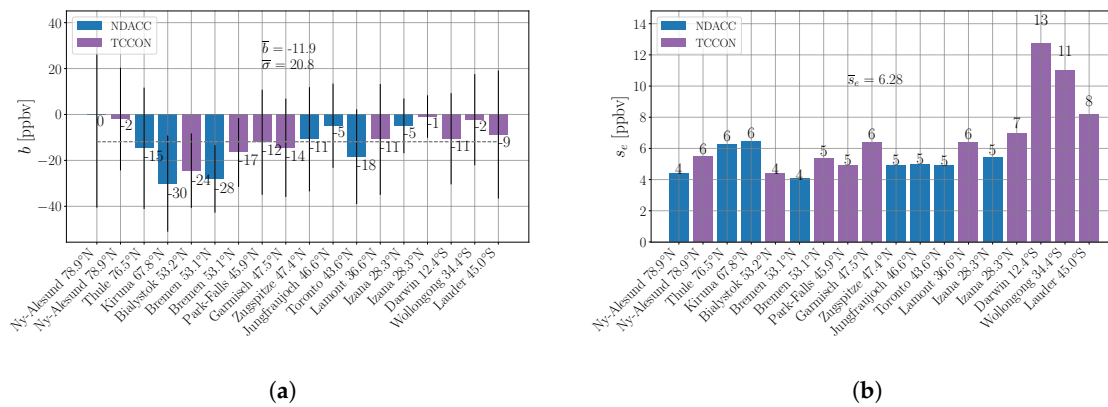
(**a**)

(**b**)

Figure 9. Same illustration as in Figure 7 and 8 but weighted inverse to the distance squared. (**a**) Global monthly-mean bias $\bar{b}$ and respective standard error $\bar{\sigma}$; (**b**) The global standard error of the mean biases $\bar{s}_e$ is the average of all standard errors of the means.

Table 3. The respective biases $\bar{b}$ (bold) and standard errors $\bar{s}_e$ separated for NDACC and TCCON. The NDACC reveals a significant bias and standard error of the means. This is also true if linear (r^{-1}) and quadratic (r^{-2}) inverse distance weighting is applied. With respect to the TCCON, the bias is less significant and closer to the magnitude of the standard error. ppbv: parts per billion in volume.

ppbv	1		r^{-1}		r^{-2}	
NDACC	**−14.3**	5.5	**−13.8**	5.4	**−14.4**	5.1
TCCON	**−8.1**	7.7	**−7.4**	7.9	**−9.1**	7.1
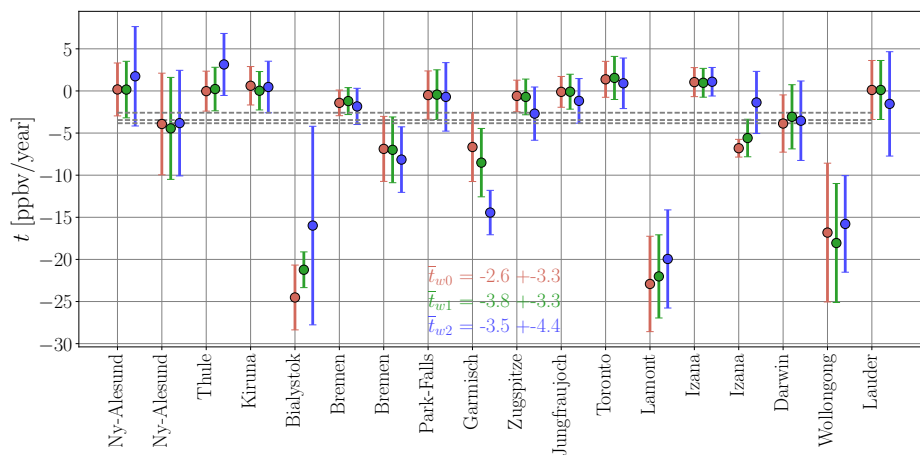
Figure 10. Linear bias trend t of the mean biases b for the classical (unweighted) and weighted approaches. No significant bias trend is identified on the global scale, however, a few sites do show a significant negative bias trend.

3.2.3. Spatio-Temporal Weighting

In general, the findings suggest that accounting for the spatial mismatch by inverse distance weighting does not necessarily require the introduction of a temporal penalty term (accounting for temporal mismatch) in order to obtain representative results for comparison. Figure 11 shows the effects of the spatio-temporal weighting on the 90-day averages for Toronto in 2003. Overall, the outcome demonstrates that temporal weighting has only minor effects on the accuracy and should in general

only be regarded as an option in areas where data quality is degraded, and therefore averaging over a larger number of observations in the space and/or time domain becomes necessary.
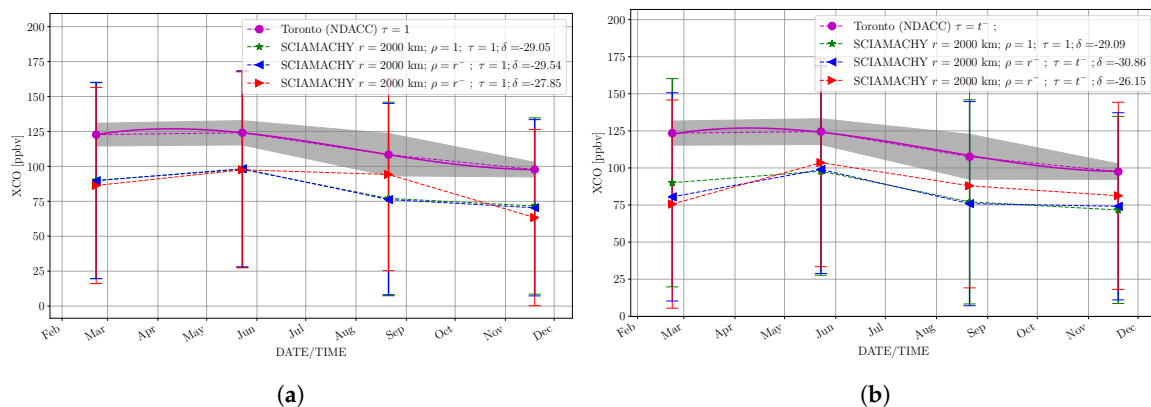

(**a**)   (**b**)

Figure 11. The 90-day averages within 2000 km from the Toronto truthing site in 2003. In (**a**) only spatial weighting was applied in the calculation of averages while in (**b**) both spatial and temporal weighting was used. δ designates the offset of the averages between xCO from SCIAMACHY and Toronto.

3.2.4. Multiannual Averages

Measurements in the SWIR channel suffered from decontamination procedures (see Section 3.1) often executed at the turn of the year. Fitting the intercept and slope based on least-squares is a process which is sensitive to outliers, and therefore, the months December and January were excluded in the calculation. In about 40% of the cases the intercept of the s-b retrievals matched within the standard deviation of the respective g-b intercept and in around 60% a positive slope of the g-b measurements matched with a positive slope of the s-b observations. Figure 12 shows two sites in the southern hemisphere where both intercept and slope coincide within the standard deviation of the g-b reference.

In Figure 13, s-b CO columns are compared to the Zugspitze (NDACC) and Garmisch (TCCON) reference sites. These g-b sites are distinct due to their close horizontal but great vertical distance (see Table 1). The permanently higher CO values at Garmisch might be an effect of the planetary boundary layer, since observations at Zugspitze are not influenced by this throughout most of the year. Notice that the 500-km circle (from where the SCIAMACHY measurements were taken) was centered at Zugspitze for both cases due to the proximity of both sites ($\approx$8 km) and the large sampling area.

Figure 13a shows the comparison for the unweighted and weighted xCO averages. The results demonstrate that some fraction of the SCIAMACHY xCO bias seen at Garmisch and Zugspitze is likely to be an effect of spatial mismatch, i.e., representation error, and should therefore not solely be considered as a retrieval or instrument flaw. Figure 13b examines how the squared inverse distance weighting is affected by variations in the size of the sampling area. It clearly reveals that the offset becomes smaller the closer the measurements are taken from the g-b site. In this case, the weighted average of SCIAMACHY observations from within 500 km shows the smallest offset with respect to the g-b sites. However, the CO annual cycle unveils in the 1000-km and 2000-km cases but is not visible in the 500-km circle.
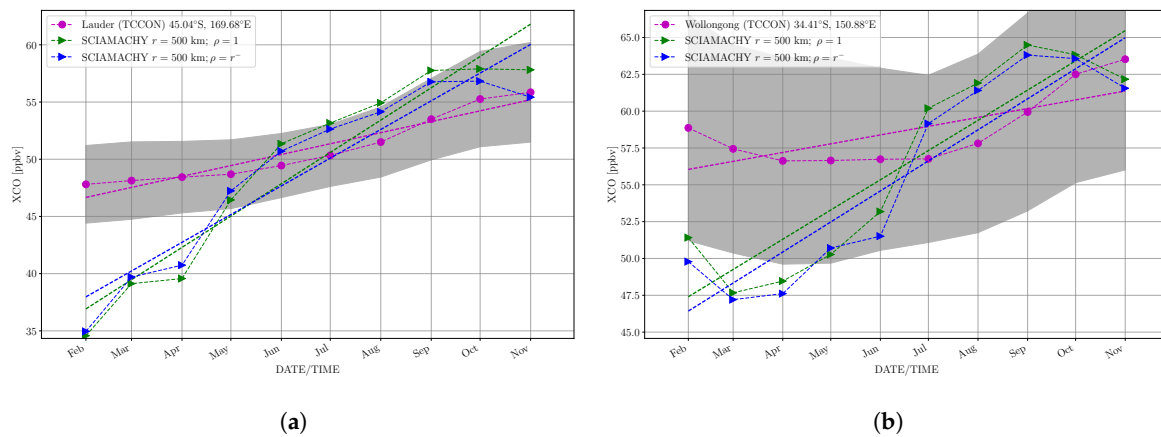
(**a**)

(**b**)

Figure 12. Comparison of intercept and slope of s-b and g-b xCO. The years 2004 through 2011 were included in the analysis for Lauder depicted in (**a**). Results for Wollongong that include data from 2008–2011 are shown in (**b**).
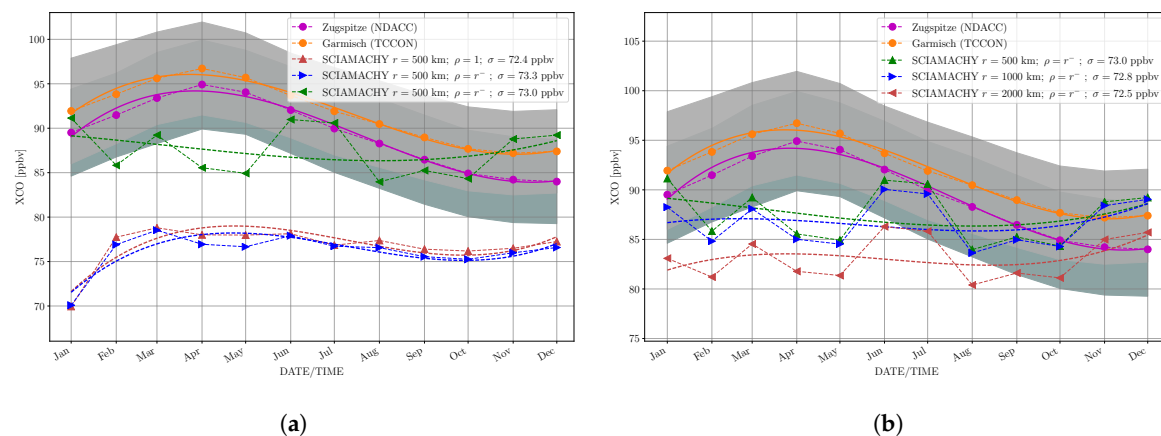

(**a**)

(**b**)

Figure 13. The effect of distance weighting on SCIAMACHY x̂CO with respect to Garmisch (2007–2011) and Zugspitze (2003–2011) reference sites. The solid and dashed lines show the three-degree polynomial fit on the individual observations and $\sigma \equiv \langle xCO_{sb} \rangle_e$. (**a**) Weighting of measurements within a given distance leads to better agreement with g-b references; (**b**) The effect of weighting partly reveals the annual cycle at the reference sites. However, the amplitude is low and a slightly negative offset remains. Note that the standard deviation of the s-b measurements is included in the legend.

3.2.5. Time Series

For the NDACC and TCCON stations listed in Table 1, Figures 14 and 15 show the time series of dry air quadratic inverse distance-weighted monthly-mean CO CMRs. Both figures demonstrate that the large scatter of the BIRRA retrievals is mainly caused by the noise in the SCIAMACHY observations. The standard deviation of individual months sometimes even exceeds a typical mean CO concentration. The results underline that solar radiation reflected at the Earth surface or clouds reduces scattering due to a higher SNR.

3.3. Mission Averaged Global CO

The 2003–2011 averaged CO product is shown in Figure 16 with the corresponding errors in Figure 17. In order to acquire representative CO mixing ratios on a regular grid globally, averaging over the complete validation period was required. Beside increased CO concentrations over China, parts of India, and Indonesia due to pollution and wildfires, Figure 16 also reveals valuable information

of the global transport of pollution in the Earth's atmosphere. For example, the outflow of CO over the Atlantic Ocean due to biomass burning in central Africa becomes clearly visible. However, the gradients in the CO columns appear to be less pronounced due to several reasons. One is that the averaging smoothes peaks of CO that occur during the wildfire seasons over central Africa. This shows that this effect is less pronounced in regions where high concentrations of CO prevail throughout the year (see eastern parts of China). Another aspect is that the filter criteria applied also accept scenes where optically thick clouds extend into the mid-troposphere, therefore providing good SNRs but obscuring the high CO concentrations in the lower troposphere and boundary layer. This reveals that the effect is most pronounced in the region of the Inter Tropic Convergence (ITC) zone.
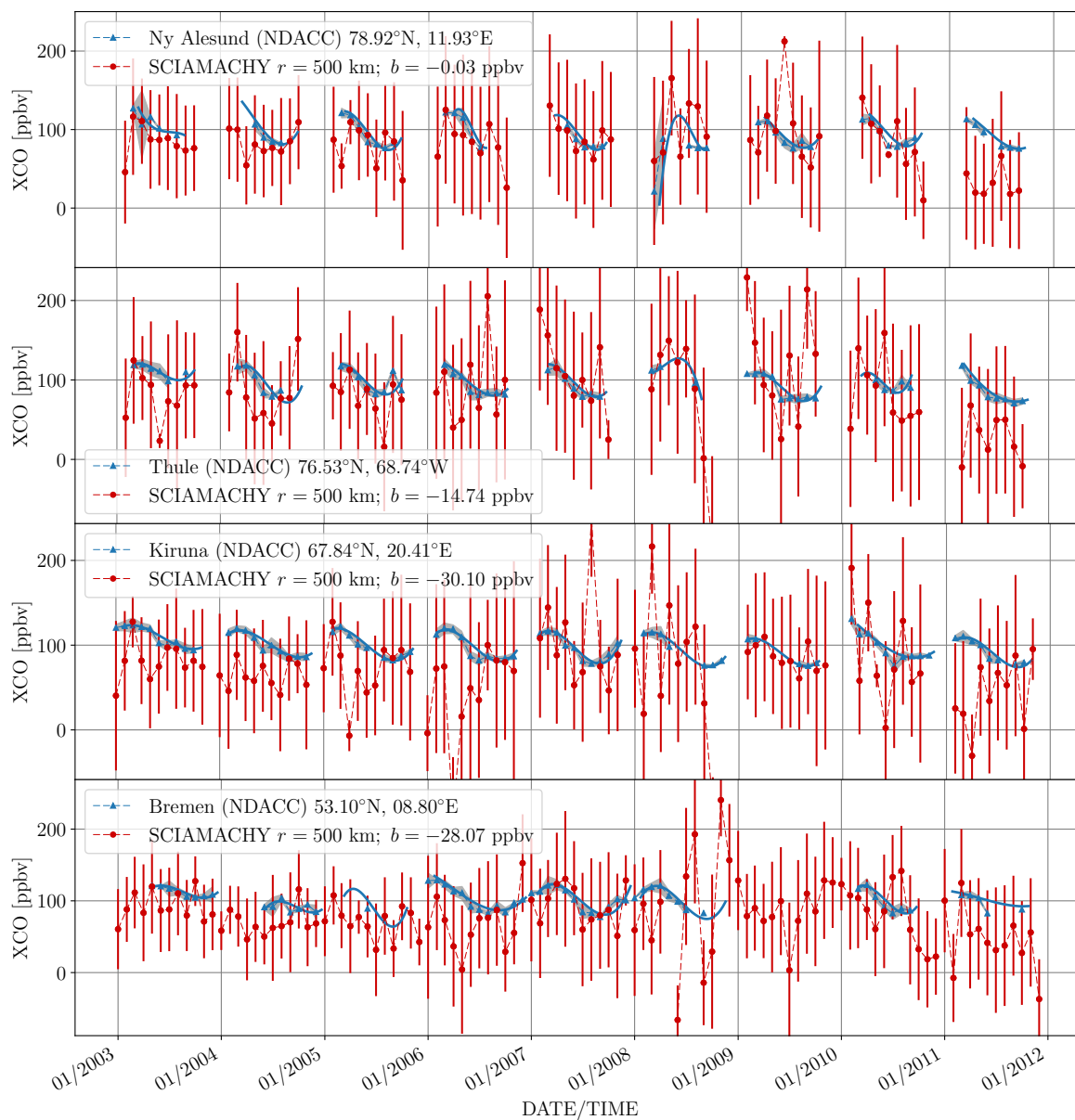
Figure 14. *Cont.*
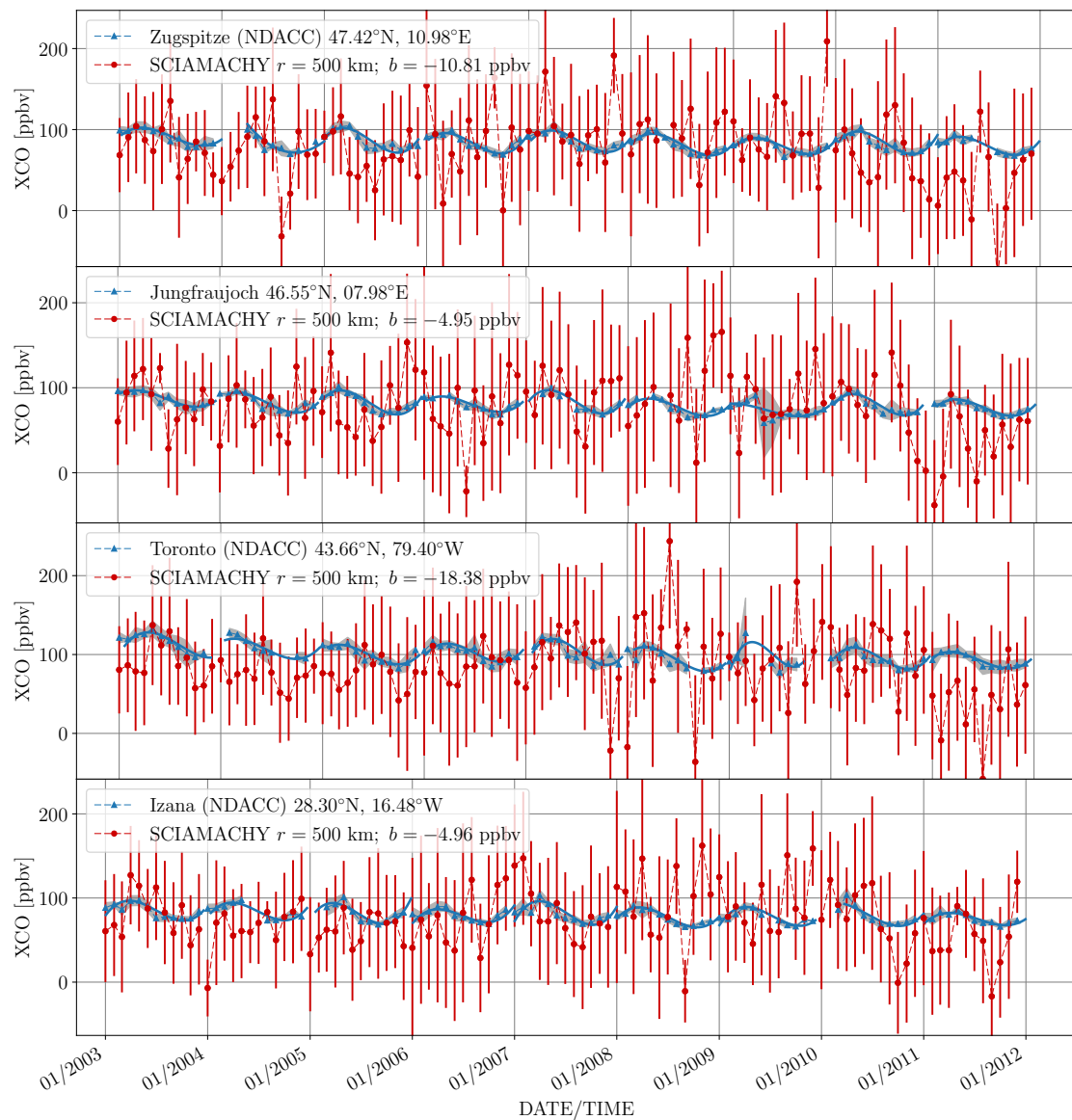
Figure 14. Time series of r^{-2} distance weighted monthly-mean xCO dry air column mixing ratios measured by the SCIAMACHY and NDACC stations. The CO mixing ratios were calculated using the surface pressure information per station according to the description in Section 2.1.1. Sites where no such information was available were skipped.
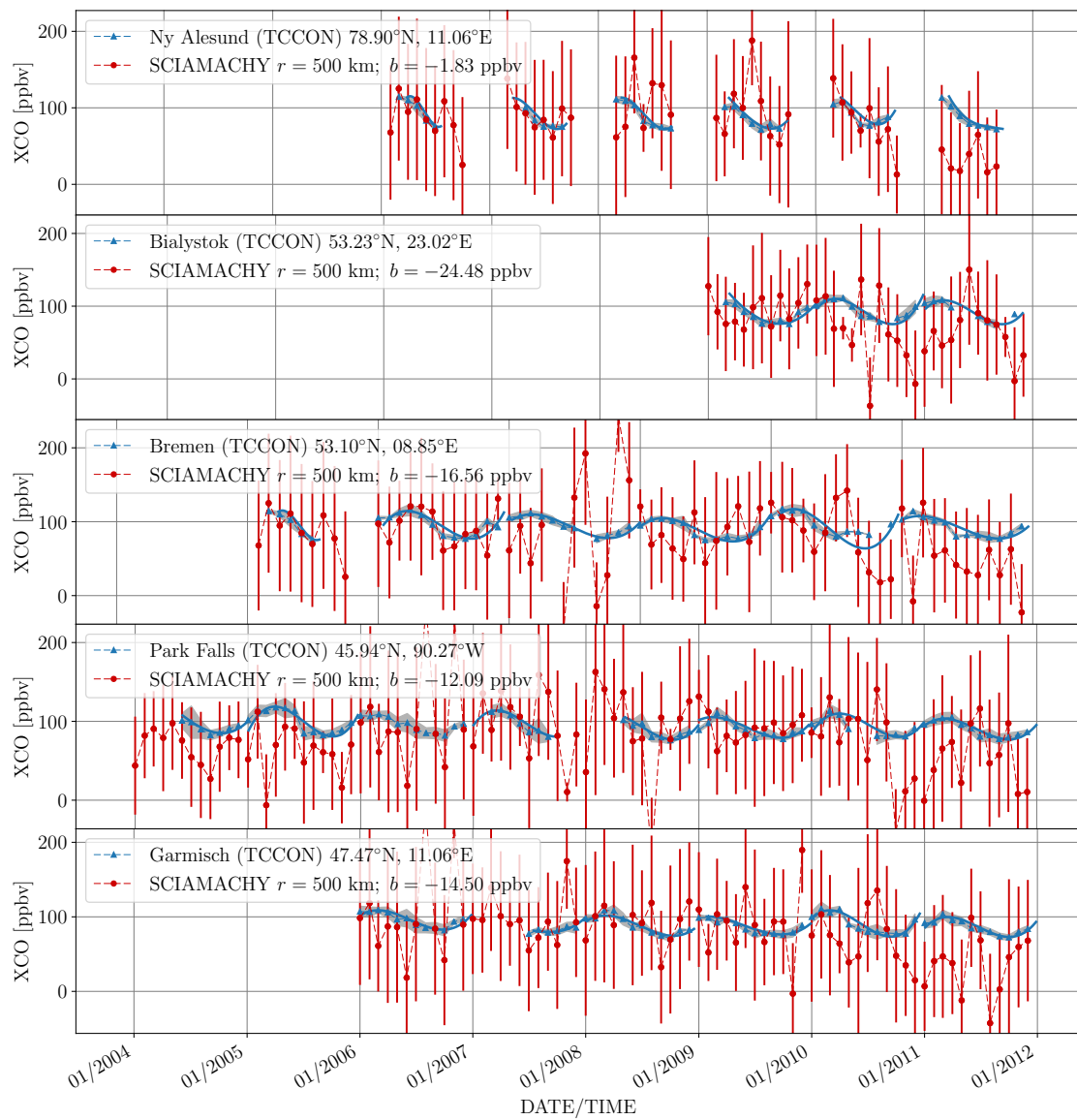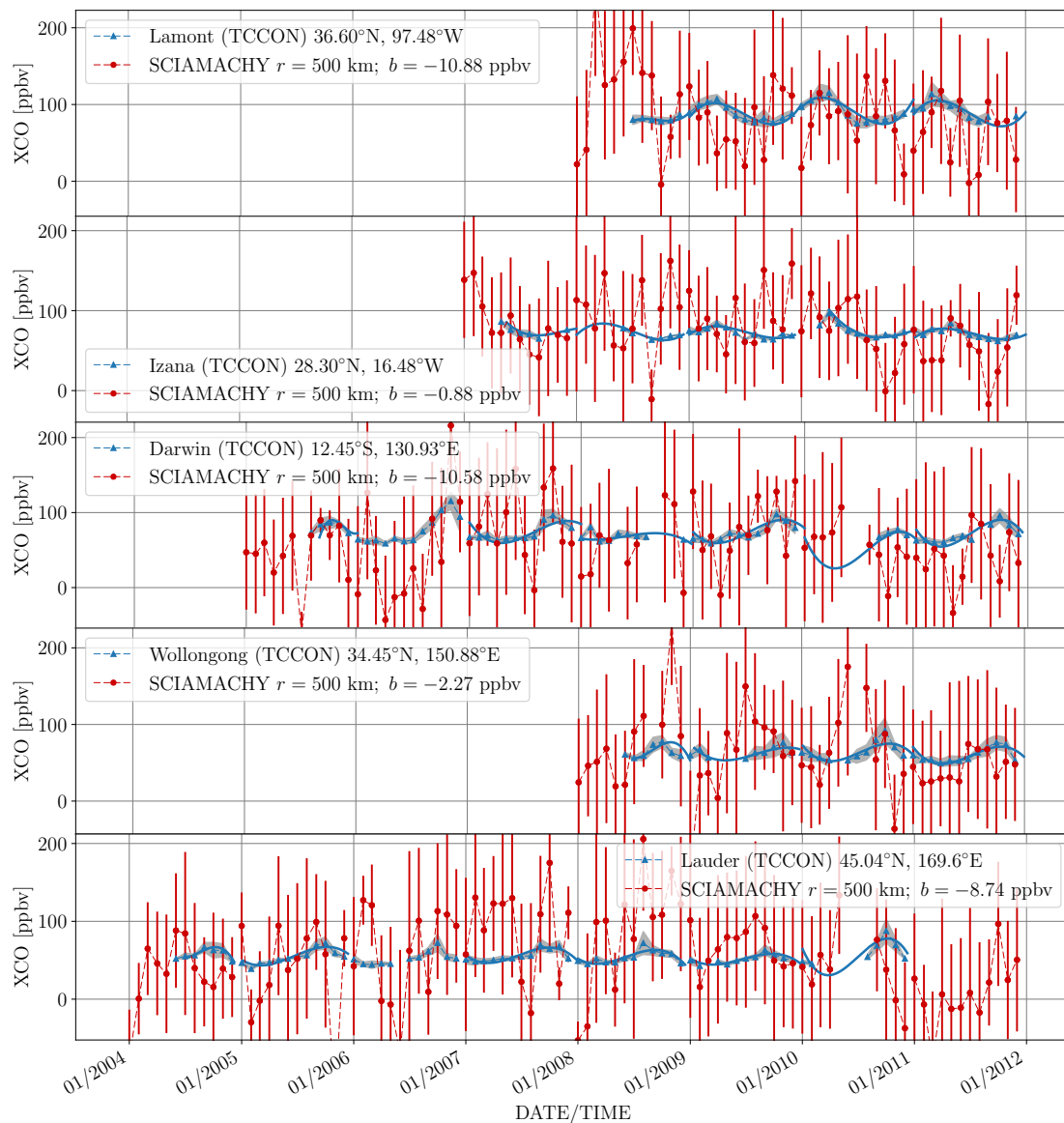
Figure 15. *Cont.*

Figure 15. Time series as in Figure 14 but for the TCCON stations. Sites that provided retrievals according to the GGG2014 data release of TCCON were used for the comparison to the BIRRA columns.
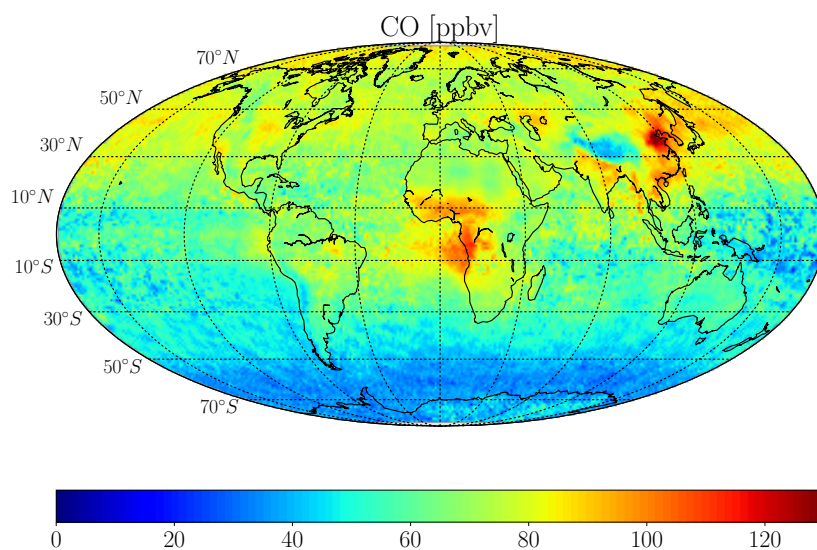
Figure 16. Dry air CO column-averaged mixing ratios in ppbv over land and the oceans. The values are averaged from January 2003 to December 2011 on a [1×1]° latitude/longitude grid. Measurements of clear-sky scenes and optically thick cloud conditions are included.
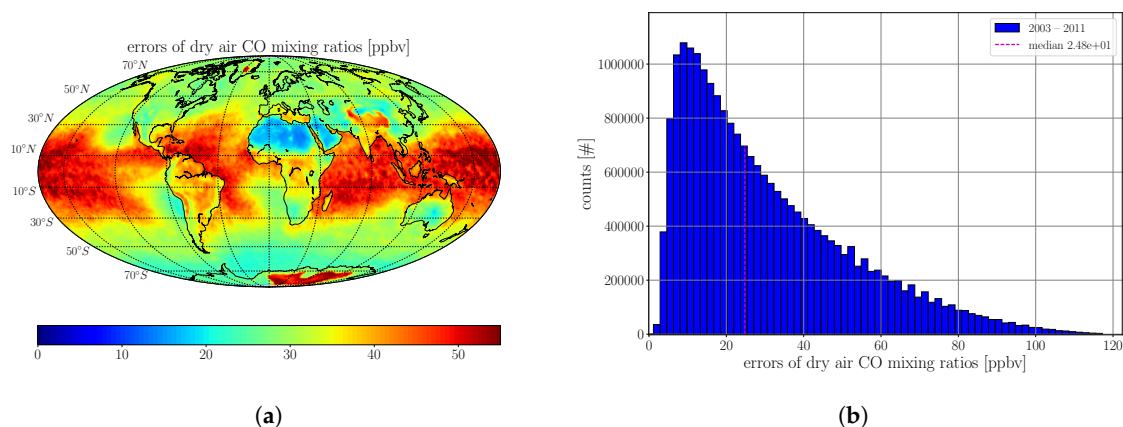

(a)

(b)

Figure 17. (**a**) Errors of the CO total columns for the observations presented in Figure 16. It clearly shows that errors are smallest in dry regions with low cloud coverage and high surface reflectivity (see the Sahara Desert, the Arabian Peninsula, Figure 1); (**b**) Frequency distribution of the errors including years from 2003–2011 (some individual years are shown in Figure 3).

4. Discussion

The BIRRA CO total column product of the SCIAMACHY Channel 8 spectra was validated for the entire mission. In general, the results suggest that weighting is capable of mitigating some incorporated representation errors that come into play when comparing satellite-averaged CO data to averages of g-b reference sites.

A global negative bias of about −10 ppbv and a standard deviation of 15–20 ppbv (depending on the approach used for averaging and the g-b reference network used, see Table 3) was found for the SCIAMACHY xCO. Since both networks perform direct sunlight observations using FTS, the fact that SCIAMACHY measurements acquired in remote areas within the subtropical bands tend to exhibit smaller negative biases compared to observations in the mid-latitudes suggests that part of the s-b bias is caused by cloud shielding of the lower troposphere and an insufficient representation of the CO a priori profile in those regions (see Figures 7–9). In the case of Izana, a site located at about 2.4 km above mean sea level, the effect of cloud shielding is small, hence the cloud induced bias is reduced.

The full-mission analysis from Borsdorff et al. [10] also included cloudy-sky retrievals and found similar results for the bias which was estimated to be $\bar{b} = -6$ ppb. Another common feature our results is the larger negative bias for the NDACC as compared to TCCON (see Table 3). The differences might arise from the fact that both networks observe in different spectral regions and are using different retrieval approaches.

4.1. Weighted Averages

In the validation of the s-b to g-b reference, it is necessary to find representative averages of both data sets with respect to one another. Early validation studies (e.g., Sussmann and Buchwitz [23]) pointed out that even without correction, the statistical effect of reducing scatter by increasing the ensemble of averaged pixels surpasses the effect of including less representative pixels. However, a thorough validation effort should also take into account the incorporating representation issue. In particular, the variability of CO induces gradients on a regional basis which the validation methodology should account for when comparing two independent systems.

The results demonstrate that assuming xCO to be less representative with increasing distance from the reference is reasonable and that distance-weighted averages reduce the negative bias for the 2003–2011 period at most reference sites (see Figures 7 and 9). The different averaging techniques agree upon the fact that the global bias is not significant with respect to the standard deviation at most g-b locations for the 2003–2011 period (see Figures 7–9). Moreover, the methods agree upon the larger standard error of the mean bias in the Southern Hemisphere. It also became evident that the selection of larger sampling areas requires smaller temporal averages to be chosen for a given noise level (see Figure 4). The findings give confidence that the BIRRA retrievals are consistent on the global scale and for the time period of the SCIAMACHY mission.

4.2. Temporal Averages

Firstly, it is important to note that with respect to time, close temporal measurements of two observation sites do not necessarily mean similar concentrations since spatial gradients are much more pronounced compared to temporal ones (i.e., show smoother gradients). The concentration of CO in rural as well as urban areas will exhibit only weak variations throughout a monthly or seasonal period. Due to this weak fluctuation, the temporal domain, including temporal weighting in the analysis of the averages, requires utilization of the spatial distance of the measurements from the independent observing systems as a penalty term. This penalty term accounts for the decreasing representativeness of the measurements due to the spatial mismatch—although they may have been observed at the same time. Therefore, including a stronger penalty term for a spatial mismatch as compared to a temporal mismatches is regarded as a reasonable approach. It is also important to note that since spatial weighting only considers measurements with respect to their distance from the reference without taking into account directional gradients, spatial weighting is expected to deliver the best results when the CO concentration is isotropically distributed around a g-b site. However, even in cases of non-perfect isotropic distribution from the reference, the method of weighted averages seems to deliver more representative results than simple averaging without weighting (see Figure 13).

4.3. Multiannual Averages

Figures 2 and 3 clearly demonstrates that later years (2006 onward) in the SCIAMACHY dataset suffered from low SNRs, and may therefore require a modified approach for comparison. The multiannual averages, presented in Section 3.2.4, are demonstated as useful and capable in further reducing random errors. Although the multiannual averages cannot eliminate the negative bias completely (see Figure 13) and reveal disagreements in slope at most sites, the results improve for the majority of sites compared to the single-year analysis.

4.4. Spatio-Temporal Averages:

Section 3.2.3 shows that incorporating temporal weighting, in addition to distance weighting, is not crucial since CO columns are quite constant within a month. Hence, closer, mutually temporal observations do not necessarily mean improved representativeness (in contrast to the spatially persistent CO gradients). However, if stronger weighting on spatial mismatch is deemed useful (e.g., in areas of strong spatial gradients) one might consider including a (less pronounced) penalty term that accounts for the temporal mismatch. Quadratic spatial weighting, applied in conjunction with temporal averages beyond 30 days, should perhaps be considered (see Figure 11).

4.5. CO from TIR vs SWIR Observations

As an alternative to the SWIR range, atmospheric remote sensing of CO can exploit the rotational band in the microwave [38–40] or the fundamental band in the TIR range. Several s-b nadir viewing instruments observe CO in the TIR: the Atmospheric Infrared Sounder (AIRS) [41], the Cross-track Infrared Sounder (CrIS) [42,43], the Infrared Atmospheric Sounding Interferometer (IASI) [44–46], the Measurement of Pollution in the Troposphere (MOPITT) instrument [47], and the Tropospheric Emission Spectrometer (TES) [48]. Initially MOPITT retrievals utilized the TIR range only, but Worden et al. [49] demonstrated that multispectral retrievals exploiting both the near infrared (NIR, 0.8–3 μm) and TIR channels increase the sensitivity to CO.

A distinct advantage of the SWIR band over the TIR band is its almost uniform sensitivity down to Earth's surface. On the other hand, the CO absorption in the TIR is significantly stronger, i.e., the vertical optical depth τ_{CO} of the CO fundamental band around 4.6 μm is about a factor of 100 larger than the optical depth of the first overtone band around 2.3 μm (see, e.g., the Atmospheric Infrared Spectrum Atlas, http://eodg.atm.ox.ac.uk/ATLAS/). Moreover, the methane optical depth is larger than the carbon monoxide optical depth in the SWIR range, whereas $\tau_{CH_4} < \tau_{CO}$ in the TIR range, at around 2100 cm^{-1}.

The accuracy of CO columns estimated from TIR observations is typically 10% (see e.g, [41,50]), i.e., much better than for SCIAMACHY: de Laat et al. [51], and Kopacz et al. [52] found that the random instrument noise related error of a single CO total column measurement is large, typically 10–100% or larger [9], and Tangborn et al. [53] stated that total error values generally range from 20 to 100% for the observations tagged as 'good'. The smaller noise related error of TIR retrievals has been confirmed by the analysis of AIRS observations using Column EstimatoR Vertical Infrared Sounding of the Atmosphere (CERVISA), a variant of the BIRRA based on Schwarzschild's equation appropriate for the thermal infrared range [54].

5. Conclusions

Dry air CO total column estimates from 2.3 μm reflectance measurements using SCIAMACHY from 2003–2011 have been validated against 18 stations of the g-b NDACC and TCCON. While observations from the NDACC covered the entire SCIAMACHY mission period, the TCCON observations could only be used to validate the CO product in the later phase of the mission. It was found that after postprocessing of the BIRRA retrieval output, the actual xCO noise still varied significantly between sites, ranging from around 100 ppbv up to 200 ppbv. This was similar to findings from other authors cited in Section 1.5.2 and therefore, averaging individual s-b CO observations was an essential task in order to validate the product. Here, a distance-weighted approach was chosen, to our knowledge for the first time, to compare the CO mixing ratios from spatially distributed satellite observations to point-like g-b measurements. The approach was demonstrated reasonable in validating CO from s-b sensors. The results also suggest that part of the discrepancy between the BIRRA-retrieved columns and g-b observations arises from representation errors, inadequate description of the a priori CO profile, and cloud shielding. In an upcoming study we aim to use model data as a reference in

addition to g-b data and examine the approach of weighted averages for other s-b platforms such as GOSAT-2 and the Sentinel-5 Precursor (S5P) [55].

Abbreviations

The following abbreviations are used in this manuscript:

AIRS	Atmospheric Infrared Sounder
BIRRA	Beer InfraRed Retrieval Algorithm
CH4	Methane
CMR	Column mixing ratio
CO	Carbon monoxide
CrIS	Cross-track Infrared Sounder
DOAS	Differential Optical Absorption Spectroscopy
ENVISAT	Environmental Satellite
FTIR	Fourier transform infrared
FTS	Fourier transform spectroscopy
G-B	Ground-based
GOMOS	Global Ozone Monitoring by Occultation of Stars
IASI	Infrared Atmospheric Sounding Interferometer
IMAP	Iterative Maximum A Posteriori
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MOPITT	Measurement of Pollution in the Troposphere
NDACC	Network for the Detection of Atmospheric Composition Change
O2	Oxygen
PPBV	Parts per billion in volume
S-B	Satellite-based
S5P	Sentinel-5 Precursor
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
SICOR	Shortwave infrared CO retrieval
SWIR	Short-wave infrared
TCCON	Total Carbon Column Observing Network
TES	Tropospheric emission spectrometer
TIR	Thermal infrared
TROPOMI	Tropospheric Monitoring Instrument
VMR	Volume mixing ratio
WFM	Weighting function modified
XCO	Dry-air average column mixing ratio

Appendix A. NDACC Data Providers

The NDACC data in this publication were obtained from sites listed in Table A1. The data are publicly available (see http://www.ndacc.org).

Table A1. NDACC sites used in this publication and the coorperating institutions.

NDACC Site	Coorperating Institutions
Bremen	Institute of Environmental Physics, University of Bremen, Germany
Garmisch	Institute for Meteorology and Climate Research, Karlsruhe Institute of Technology, Gemany
Izana	Izana Atmospheric Research Center, AEMET- Meteorological State Agency, Spain
Jungfraujoch	Universtiy of Liége, Belgium
Kiruna	Institute of Space Physics, Sweden
Ny-Alesund	Institute of Environmental Physics, University of Bremen, Germany
	Alfred Wegener Institute for Polar and Marine Research, Germany
Thule	Danish Climate Center, Danish Meteorological Institute Copenhagen, Denmark
Toronto	Department of Physics, University of Toronto, Canada

Appendix B. TCCON Data Providers

The TCCON data were obtained from sites listed in Table B1. The TCCON Data Archive is hosted by CaltechDATA, California Institute of Technology, CA (US), doi:https://doi.org/10.14291/tccon.archive/1348407.

Table B1. TCCON sites with references to the data used in this publication.

TCCON Site	Reference
Bialystok	Deutscher et al. [56]
Bremen	Notholt et al. [57]
Darwin	Griffith et al. [58]
Garmisch	Sussmann and Rettinger [59]
Izana	Blumenstock et al. [60]
Lamont	Wennberg et al. [61]
Lauder	Sherlock et al. [62]
Ny-Alesund	Notholt et al. [63]
Park-Falls	Wennberg et al. [64]
Wollongong	Griffith et al. [65]

References

1. ESA. ESA Declares End of Mission for ENVISAT. 2012. Available online: https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/envisat (accessed on 15 January 2018).
2. Gottwald, M.; Bovensmann, H. (Eds.) *SCIAMACHY—Exploring the Changing Earth's Atmosphere*; Springer: Dordrecht, The Netherlands, 2011.
3. Gloudemans, A.; Schrijver, H.; Kleipool, Q.; van den Broek, M.; Straume, A.; Lichtenberg, G.; van Hees, R.; Aben, I.; Meirink, J. The impact of SCIAMACHY near-infrared instrument calibration on CH_4 and CO total columns. *Atmos. Chem. Phys.* **2005**, *5*, 2369–2383.
4. Lichtenberg, G.; Gimeno Garcia, S.; Schreier, F.; Slijkhuis, S.; Snel, R.; Bovensmann, H. Impact of Level 1 Quality on SCIAMACHY Level 2 Retrieval. In Proceedings of the 38th COSPAR Scientific Assembly, Bremen, Germany, 18–25 July 2010.
5. Buchwitz, M.; de Beek, R.; Bramstedt, K.; Noël, S.; Bovensmann, H.; Burrows, J.P. Global carbon monoxide as retrieved from SCIAMACHY by WFM-DOAS. *Atmos. Chem. Phys.* **2004**, *4*, 1945–1960.
6. Buchwitz, M.; de Beek, R.; Noël, S.; Burrows, J.P.; Bovensmann, H.; Bremer, H.; Bergamaschi, P.; Körner, S.; Heimann, M. Carbon monoxide, methane and carbon dioxide columns retrieved from SCIAMACHY by WFM-DOAS: year 2003 initial data set. *Atmos. Chem. Phys.* **2005**, *5*, 3313–3329.

7. Frankenberg, C.; Platt, U.; Wagner, T. Retrieval of CO from SCIAMACHY onboard ENVISAT: Detection of strongly polluted areas and seasonal patterns in global CO abundances. *Atmos. Chem. Phys.* **2005**, *5*, 1639–1644.

8. Gimeno García, S.; Schreier, F.; Lichtenberg, G.; Slijkhuis, S. Near infrared nadir retrieval of vertical column densities: methodology and application to SCIAMACHY. *Atmos. Meas. Tech.* **2011**, *4*, 2633–2657.

9. Borsdorff, T.; Tol, P.; Williams, J.E.; de Laat, J.; aan de Brugh, J.; Nédélec, P.; Aben, I.; Landgraf, J. Carbon monoxide total columns from SCIAMACHY 2.3 µm atmospheric reflectance measurements: Towards a full-mission data product (2003–2012). *Atmos. Meas. Tech.* **2016**, *9*, 227–248.

10. Borsdorff, T.; aan de Brugh, J.; Hu, H.; Nédélec, P.; Aben, I.; Landgraf, J. Carbon monoxide column retrieval for clear-sky and cloudy atmospheres: a full-mission data set from SCIAMACHY 2.3 µm reflectance measurements. *Atmos. Meas. Tech.* **2017**, *10*, 1769–1782.

11. Clerbaux, C.; Boynard, A.; Clarisse, L.; George, M.; Hadji-Lazaro, J.; Herbin, H.; Hurtmans, D.; Pommier, M.; Razavi, A.; Turquety, S.; et al. Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmos. Chem. Phys.* **2009**, *9*, 6041–6054.

12. Fischer, H.; Birk, M.; Blom, C.; Carli, B.; Carlotti, M.; von Clarmann, T.; Delbouille, L.; Dudhia, A.; Ehhalt, D.; Endemann, M.; et al. MIPAS: an instrument for atmospheric and climate research. *Atmos. Chem. Phys.* **2008**, *8*, 2151–2188.

13. Bernath, P.; McElroy, C.T.; Abrams, M.C.; Boone, C.D.; Butler, M.; Camy-Peyret, C.; Carleer, M.; Clerbaux, C.; Coheur, P.F.; Colin, R.; et al. Atmospheric Chemistry Experiment (ACE): Mission overview. *Geophys. Res. Lett.* **2005**, *32*, L15S01, doi:10.1029/2005GL022386.

14. Kuze, A.; Suto, H.; Nakajima, M.; Hamazaki, T. Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.* **2009**, *48*, 6716–6733.

15. Wunch, D.; Toon, G.C.; Blavier, J.F.L.; Washenfelder, R.A.; Notholt, J.; Connor, B.J.; Griffith, D.W.T.; Sherlock, V.; Wennberg, P.O. The Total Carbon Column Observing Network. *Philos. Trans. R. Soc. A* **2011**, *369*, 2087–2112.

16. Feist, D. Global FTS-Network. 2015. Available online: https://oc.bgc-jena.mpg.de (accessed on 20 June 2016).

17. Warnek, T.; Notholt, J.; Chen, H. *Validation and Calibration of Greenhouse-Gas Satellite Observations by Ground-Based Remote Sensing Measurements*; University of Bremen: Bibliothekstraße, Bremen, 2015.

18. TCCON. Total Carbon Column Observing Network—Wiki. 2018. Available online: https://tccon-wiki.caltech.edu/Network_Policy/Data_Use_Policy/Data_Description#Measurement_precision_(repeatability) (accessed on 11 January 2018).

19. Einarsson, B. (Ed.) *Accuracy and Reliability in Scientific Computing*; SIAM: Philadelphia, PA, USA, 2005.

20. Verhoelst, T.; Granville, J.; Hendrick, F.; Köhler, U.; Lerot, C.; Pommereau, J.P.; Redondas, A.; Van Roozendael, M.; Lambert, J.C. Metrology of ground-based satellite validation: co-location mismatch and smoothing issues of total ozone comparisons. *Atmos. Meas. Tech.* **2015**, *8*, 5039–5062.

21. De Laat, A.T.J.; Gloudemans, A.M.S.; Schrijver, H.; Aben, I.; Nagahama, Y.; Suzuki, K.; Mahieu, E.; Jones, N.B.; Paton-Walsh, C.; Deutscher, N.M.; et al. Validation of five years (2003–2007) of SCIAMACHY CO total column measurements using ground-based spectrometer observations. *Atmos. Meas. Tech.* **2010**, *3*, 1457–1471.

22. De Laat, A.; Gloudemans, A.; Schrijver, H.; van den Broek, M.; Meirink, J.; Aben, I.; Krol, M. Quantitative analysis of SCIAMACHY carbon monoxide total column measurements. *Geophys. Res. Lett.* **2006**, *33*, L07807.

23. Sussmann, R.; Buchwitz, M. Initial validation of ENVISAT/SCIAMACHY columnar CO by FTIR profile retrievals at the Ground-Truthing Station Zugspitze. *Atmos. Chem. Phys.* **2005**, *5*, 1497–1503.

24. Sussmann, R.; Stremme, W.; Buchwitz, M.; de Beek, R. Validation of ENVISAT/SCIAMACHY columnar methane by solar FTIR spectrometry at the Ground-Truthing Station Zugspitze. *Atmos. Chem. Phys.* **2005**, *5*, 2419–2429.

25. Dils, B.; De Maziere, M.; Muller, J.; Blumenstock, T.; Buchwitz, M.; de Beek, R.; Demoulin, P.; Duchatelet, P.; Fast, H.; Frankenberg, C.; et al. Comparisons between SCIAMACHY and ground-based FTIR data for total columns of CO, CH_4, CO_2 and N_2O. *Atmos. Chem. Phys.* **2006**, *6*, 1953–1976.

26. Schneising, O.; Bergamaschi, P.; Bovensmann, H.; Buchwitz, M.; Burrows, J.P.; Deutscher, N.M.; Griffith, D.W.T.; Heymann, J.; Macatangay, R.; Messerschmidt, J.; et al. Atmospheric greenhouse gases retrieved from SCIAMACHY: Comparison to ground-based FTS measurements and model results. *Atmos. Chem. Phys.* **2012**, *12*, 1527–1540.

27. Meier, A. *Spectroscopic Atlas of Atmospheric Microwindows in the Middle Infrared*; IRF Technical Report; IRF Institute för Rymdfysik: Kiruna, Sweden, 2004.

28. Kiel, M.; Wunch, D.; Wennberg, P.O.; Toon, G.C.; Hase, F.; Blumenstock, T. Improved retrieval of gas abundances from near-infrared solar FTIR spectra measured at the Karlsruhe TCCON station. *Atmos. Meas. Tech.* **2016**, *9*, 669–682.

29. Wunch, D.; Toon, G.C.; Sherlock, V.; Deutscher, N.M.; Liu, C.; Feist, D.G.; Wennberg, P.O. *The Total Carbon Column Observing Network's GGG2014 Data Version*; Technical Report; Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2015.

30. Schreier, F.; Gimeno García, S.; Hedelt, P.; Hess, M.; Mendrok, J.; Vasquez, M.; Xu, J. GARLIC—A General Purpose Atmospheric Radiative Transfer Line-by-Line Infrared-Microwave Code: Implementation and Evaluation. *J. Quant. Spectrosc. Radiat. Transf.* **2014**, *137*, 29–50.

31. Schreier, F.; Gimeno García, S.; Vasquez, M.; Xu, J. Algorithmic vs. finite difference Jacobians for infrared atmospheric radiative transfer. *J. Quant. Spectrosc. Radiat. Transf.* **2015**, *164*, 147–160.

32. Dennis, J., Jr.; Gay, D.; Welsch, R. Algorithm 573: NL2SOL—An Adaptive Nonlinear Least–Squares Algorithm. *ACM Trans. Math. Softw.* **1981**, *7*, 369–383.

33. Golub, G.; Pereyra, V. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Probl.* **2003**, *19*, R1–R26.

34. Anderson, G.; Clough, S.; Kneizys, F.; Chetwynd, J.; Shettle, E. *AFGL Atmospheric Constituent Profiles (0–120 km)*; Technical Report TR-86-0110; AFGL: Hanscom AFB, MA, USA, 1986.

35. Gloudemans, A.; de Laat, A.; Schrijver, H.; Aben, I.; Meirink, J.; van der Werf, G. SCIAMACHY CO over land and oceans: 2003–2007 interannual variability. *Atmos. Chem. Phys.* **2009**, *9*, 3799–3813.

36. Buschmann, M.; Deutscher, N.M.; Sherlock, V.; Palm, M.; Warneke, T.; Notholt, J. Retrieval of xCO$_2$ from ground-based mid-infrared (NDACC) solar absorption spectra and comparison to TCCON. *Atmos. Meas. Tech.* **2016**, *9*, 577–585.

37. Ostler, A.; Sussmann, R.; Rettinger, M.; Deutscher, N.M.; Dohe, S.; Hase, F.; Jones, N.; Palm, M.; Sinnhuber, B.M. Multistation intercomparison of column-averaged methane from NDACC and TCCON: Impact of dynamical variability. *Atmos. Meas. Tech.* **2014**, *7*, 4081–4101.

38. Filipiak, M.; Harwood, R.; Jiang, J.; Li, Q.; Livesey, N.; Manney, G.; Read, W.; Schwartz, M.; Waters, J.; Wu, D. Carbon Monoxide Measured by the EOS Microwave Limb Sounder on Aura: First Results. *Geophys. Res. Lett.* **2005**, *32*, L14825.

39. Pumphrey, H.; Filipiak, M.; Livesey, N.; Schwartz, M.; Boone, C.; Walker, K.; Bernath, P.; Ricaud, P.; Barret, B.; Clerbaux, C.; et al. Validation of middle-atmosphere carbon monoxide retrievals from the Microwave Limb Sounder on Aura. *J. Geophys. Res.* **2007**, *112*, D24S38.

40. Dupuy, E.; Urban, J.; Ricaud, P.; Le Flochmoën, E.; Lautié, N.; Murtagh, D.; De La Noë, J.; El Amraoui, L.; Eriksson, P.; Forkman, P.; et al. Strato-mesospheric measurements of carbon monoxide with the Odin Sub-Millimetre Radiometer: Retrieval and first results. *Geophys. Res. Lett.* **2004**, *31*, L20101.

41. McMillan, W.W.; Barnet, C.; Strow, L.; Chahine, M.T.; McCourt, M.L.; Warner, J.X.; Novelli, P.C.; Korontzi, S.; Maddy, E.S.; Datta, S. Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. *Geophys. Res. Lett.* **2005**, *32*, L11801.

42. Gambacorta, A.; Barnet, C.; Wolf, W.; King, T.; Maddy, E.; Strow, L.; Xiong, X.; Nalli, N.; Goldberg, M. An Experiment Using High Spectral Resolution CrIS Measurements for Atmospheric Trace Gases: Carbon Monoxide Retrieval Impact Study. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1639–1643.

43. Fu, D.; Bowman, K.W.; Worden, H.M.; Natraj, V.; Worden, J.R.; Yu, S.; Veefkind, P.; Aben, I.; Landgraf, J.; Strow, L.; et al. High-resolution tropospheric carbon monoxide profiles retrieved from CrIS and TROPOMI. *Atmos. Meas. Tech.* **2016**, *9*, 2567–2579.

44. George, M.; Clerbaux, C.; Hurtmans, D.; Turquety, S.; Coheur, P.F.; Pommier, M.; Hadji-Lazaro, J.; Edwards, D.P.; Worden, H.; Luo, M.; et al. Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors. *Atmos. Chem. Phys.* **2009**, *9*, 8317–8330.

45. Fortems-Cheiney, A.; Chevallier, F.; Pison, I.; Bousquet, P.; Carouge, C.; Clerbaux, C.; Coheur, P.F.; George, M.; Hurtmans, D.; Szopa, S. On the capability of IASI measurements to inform about CO surface emissions. *Atmos. Chem. Phys.* **2009**, *9*, 8735–8743.

46. Illingworth, S.; Remedios, J.; Boesch, H.; Moore, D.; Sembhi, H.; Dudhia, A.; Walker, J. ULIRS, an optimal estimation retrieval scheme for carbon monoxide using IASI spectral radiances: sensitivity analysis, error budget and simulations. *Atmos. Meas. Tech.* **2011**, *4*, 269–288.

47. Deeter, M.N.; Emmons, L.K.; Francis, G.L.; Edwards, D.P.; Gille, J.C.; Warner, J.X.; Khattatov, B.; Ziskin, D.; Lamarque, J.F.; Ho, S.P.; et al. Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument. *J. Geophys. Res.* **2003**, *108*, doi:10.1029/2002JD003186.

48. Rinsland, C.; Luo, M.; Logan, J.; Beer, R.; Worden, H.; Worden, J.; Bowman, K.; Kulawik, S.; Rider, D.; Osterman, G.; et al. Nadir Measurements of Carbon Monoxide Distributions by the Tropospheric Emission Spectrometer onboard the Aura Spacecraft: Overview of Analysis Approach and Examples of Initial Results. *Geophys. Res. Lett.* **2006**, *33*, L22806.

49. Worden, H.; Deeter, M.; Edwards, D.; Gille, J.; Drummond, J.; Nédélec, P. Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals. *J. Geophys. Res.* **2010**, *115*, D16314.

50. Emmons, L.; Pfister, G.; Edwards, D.; Gille, J.; Sachse, G.; Blake, D.; Wofsy, S.; Gerbig, C.; Matross, D.; Nedelec, P. Measurements of Pollution in the Troposphere (MOPITT) validation exercises during summer 2004 field campaigns over North America. *J. Geophys. Res.* **2007**, *112*, D12S02.

51. De Laat, A.; Gloudemans, A.; Aben, I.; Krol, M.; Meirink, J.; van der Werf, G.; Schrijver, H. Scanning Imaging Absorption Spectrometer for Atmospheric Chartography carbon monoxide total columns: Statistical evaluation and comparison with chemistry transport model results. *J. Geophys. Res.* **2007**, *112*, D12310.

52. Kopacz, M.; Jacob, D.J.; Fisher, J.A.; Logan, J.A.; Zhang, L.; Megretskaia, I.A.; Yantosca, R.M.; Singh, K.; Henze, D.K.; Burrows, J.P.; et al. Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES). *Atmos. Chem. Phys.* **2010**, *10*, 855–876.

53. Tangborn, A.; Stajner, I.; Buchwitz, M.; Khlystova, I.; Pawson, S.; Burrows, J.; Hudman, R.; Nedelec, P. Assimilation of SCIAMACHY total column CO observations: Global and regional analysis of data impact. *J. Geophys. Res.* **2009**, *114*, D07307.

54. Schreier, F.; Gimeno García, S.; Lichtenberg, G.; Hoffmann, P. Intercomparison of Carbon Monoxide Retrievals from SCIAMACHY and AIRS Nadir Observations. In Proceedings of the Atmospheric Science Conference, Barcelona, Spain, 7–11 September 2009; Volume SP-676.

55. Landgraf, J.; aan de Brugh, J.; Scheepmaker, R.; Borsdorff, T.; Hu, H.; Houweling, S.; Butz, A.; Aben, I.; Hasekamp, O. Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements. *Atmos. Meas. Tech.* **2016**, *9*, 4955–4975.

56. Deutscher, N.M.; Notholt, J.; Messerschmidt, J.; Weinzierl, C.; Warneke, T.; Petri, C.; Grupe, P.; Katrynski, K. *TCCON Data from Bialystok (PL), Release GGG2014.R1*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2015; doi:10.14291/tccon.ggg2014.bialystok01.R1/1183984.

57. Notholt, J.; Petri, C.; Warneke, T.; Deutscher, N.M.; Buschmann, M.; Weinzierl, C.; Macatangay, R.C.; Grupe, P. *TCCON Data from Bremen (DE), Release GGG2014.R0*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2014; doi:10.14291/tccon.ggg2014.bremen01.R0/1149275.

58. Griffith, D.W.T.; Deutscher, N.M.; Velazco, V.A.; Wennberg, P.O.; Yavin, Y.; Keppel-Aleks, G.; Washenfelder, R.; Toon, G.C.; Blavier, J.F.; Paton-Walsh, C.; et al. *TCCON Data from Darwin (AU), Release GGG2014.R0*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2014; doi:10.14291/tccon.ggg2014.darwin01.R0/1149290.

59. Sussmann, R.; Rettinger, M. *TCCON Data from Garmisch (DE), Release GGG2014.R0*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2014; doi:10.14291/tccon.ggg2014.garmisch01.R0/1149299.

60. Blumenstock, T.; Hase, F.; Schneider, M.; Garcia, O.E.; Sepulveda, E. *TCCON Data from Izana (ES), Release GGG2014.R0*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2014; doi:10.14291/tccon.ggg2014.izana01.R0/1149295.

61. Wennberg, P.O.; Wunch, D.; Roehl, C.; Blavier, J.F.; Toon, G.C.; Allen, N. *TCCON Data from Lamont (US), Release GGG2014.R1*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2016; doi:10.14291/tccon.ggg2014.lamont01.R1/1255070.

62. Sherlock, V.; Connor, B.; Robinson, J.; Shiona, H.; Smale, D.; Pollard, D. *TCCON Data from Lauder (NZ), 120HR, Release GGG2014.R0*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2014; doi:10.14291/tccon.ggg2014.lauder01.R0/1149293.

63. Notholt, J.; Warneke, T.; Petri, C.; Deutscher, N.M.; Weinzierl, C.; Palm, M.; Buschmann, M. *TCCON Data from Ny Ålesund, Spitsbergen (NO), Release GGG2014.R0*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2017; doi:10.14291/tccon.ggg2014.nyalesund01.R0.

64. Wennberg, P.O.; Roehl, C.; Wunch, D.; Toon, G.C.; Blavier, J.F.; Washenfelder, R.; Keppel-Aleks, G.; Allen, N.; Ayers, J. *TCCON Data from Park Falls (US), Release GGG2014.R0*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2014; doi:10.14291/tccon.ggg2014.parkfalls01.R0/1149161.

65. Griffith, D.W.T.; Velazco, V.A.; Deutscher, N.M.; Paton-Walsh, C.; Jones, N.B.; Wilson, S.R.; Macatangay, R.C.; Kettlewell, G.C.; Buchholz, R.R.; Riggenbach, M. *TCCON Data from Wollongong (AU), Release GGG2014.R0*; TCCON Data Archive; CaltechDATA, California Institute of Technology: Pasadena, CA, USA, 2014; doi:10.14291/tccon.ggg2014.wollongong01.R0/1149291.

Publication II

remote sensing

MDPI

Impact of Molecular Spectroscopy on Carbon Monoxide Abundances from SCIAMACHY

Philipp Hochstaffl *[D] and **Franz Schreier**[D]

DLR—German Aerospace Center, Remote Sensing Technology Institute, 82234 Oberpfaffenhofen, Germany;
franz.schreier@dlr.de
* Correspondence: philipp.hochstaffl@dlr.de

check for updates

Abstract: High-quality observations have indicated the need for improved molecular spectroscopy for accurate atmospheric characterization. Line data provided by the new SEOM-IAS (Scientific Exploitation of Operational Missions—Improved Atmospheric Spectroscopy) database in the shortwave infrared (SWIR) region were used to retrieve CO total vertical columns from a selected set of nadir SCIAMACHY (SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY) observations. In order to assess the quality of the retrieval results, differences in the spectral fitting residuals with respect to the HITRAN 2016 (High-resolution TRANsmission molecular absorption) and GEISA 2015 (Gestion et Etude des Informations Spectroscopiques Atmosphériques) line lists were quantified and column-averaged dry-air CO mole fractions were compared to NDACC (Network for the Detection of Atmospheric Composition Change) and TCCON (Total Carbon Column Observing Network) ground-based measurements. In general, it was found that using SEOM-IAS line data with corresponding line models improve the spectral quality of the retrieval (smaller residuals) and increase the fitted CO columns, thereby reducing the bias to both ground-based networks.

1. Introduction

Remote sensing is an important asset in monitoring the state of Earth's atmosphere. Only space-borne instruments are providing continuous global coverage and are now invaluable for long-term atmospheric characterization and measurements in the context of climate change, stratospheric ozone depletion, and regional air quality studies. Low Earth-orbiting sensors are complemented by numerous balloon- or airborne sensors as well as ground-based instrumentation that is important for the validation of satellite observations as well as for dedicated studies. In all cases a rigorous assessment of the quality of the remote-sensing products is essential and necessary.

Quantification of atmospheric state variables from remote-sensing instruments constitutes an inverse problem that is usually solved by the iterative solution of an optimization problem with repeated calls to a forward model. Among the numerous parameters required by the forward model (essentially comprising atmospheric radiative transfer and an instrument model) and the inversion algorithm, the spectroscopic data characterizing the atmospheric species play a central role [1–8]. More specifically, a thorough knowledge of spectral line characteristics is indispensable for line-by-line modeling of absorption through the atmosphere and therefore has a critical impact on the estimation of the atmospheric state [9–11]. Accordingly, a considerable effort has been devoted to collect, expand and improve line data, and the most recent releases of the most widely used HITRAN (High Resolution Transmission) and GEISA (Gestion et Etude des Informations Spectroscopiques Atmosphériques) databases comprise several million lines for some 50 molecular species [12,13]. Beside these "general purpose" spectroscopic databases several more specialized compilations exist, e.g., the Jet Propulsion

Laboratory (JPL) catalogue [14] and the Cologne Database for Molecular Spectroscopy (CDMS) [15] for the microwave or databases for a specific molecule or mission (e.g., [16–19]).

In contrast to the microwave, thermal infrared, visible or ultraviolet spectral range comparatively few space missions carry sensors working in the shortwave and near infrared (SWIR, NIR) regions, such as the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) aboard the ENVIronmental SATellite (ENVISAT) [20], the Measurements Of Pollution In The Troposphere (MOPITT) aboard the Terra satellite [21], the Thermal And Near infrared Sensor for carbon Observations – Fourier Transform Spectrometer (TANSO-FTS) aboard the Greenhouse Gases Observing Satellite (GOSAT) [22], the Orbiting Carbon Observatory (OCO-2) satellite [23], the TanSat minisatellite [24] and the TROPOspheric Monitoring Instrument (TROPOMI) aboard Sentinel-5p (S5p) [25].

Numerous studies have been devoted to assessing the impact of spectroscopic data on retrievals for past and present missions and to estimate the quality (i.e., accuracy and precision) requirements for the various line parameters needed to meet the mission objectives. Frankenberg et al. [26] exploited new laboratory spectra to fit line parameters of methane in the 1.6 µm region and found that SCIAMACHY retrievals using an updated dataset are systematically different from those using HITRAN 2004, thus reducing a seasonal and latitudinal bias and leading to better consistency with atmospheric models. Furthermore, Frankenberg et al. [27] showed that inaccuracies in water spectroscopic data cause a substantial overestimate of methane correlated with high water abundances. Scheepmaker et al. [10] found that improved water vapor spectroscopy in the 2.3 µm range is beneficial for SCIAMACHY HDO/H_2O retrievals as well as ground-based Fourier transform spectra. Gloudemans et al. [28] concluded that "spectroscopic uncertainties are mostly negligible except for uncertainties in the CH_4 line intensities."

Beside these SCIAMACHY related studies Oyafuso et al. [29] examined the updated carbon dioxide cross sections in the 1.6 µm and 2.3 µm region for the OCO-2 mission and concluded that "further work is needed to eliminate systematic residuals in atmospheric spectra". For the GOSAT mission, regular updates in the methane spectral line list were discussed by Nikitin et al. [16,18] with improvements mainly involving line positions and intensities. In preparation for ESA's S5p mission, Galli et al. [9] and Checa-García et al. [11] investigated the impact of spectroscopic uncertainties on an S5p-like observer. They found that spectroscopic errors in the 2.3 µm band can induce regionally correlated errors that exceed TROPOMI/S5p's CH_4 error budget and that further efforts from the spectroscopy community should be directed to the H_2O and CH_4 spectroscopy in this regime.

The Scientific Exploitation of Operational Missions (SEOM)—Improved Atmospheric Spectroscopy (IAS) [19], henceforth designated as SEOM, was an ESA funded study to improve spectroscopic data. Databases in the 2.3 µm region, covering most of SCIAMACHY's channel 8, contain molecular absorption line parameters for CO, CH_4 and H_2O according to the needs of the TROPOMI instrument. A first assessment of the impact of these new datasets has recently been published by Borsdorff et al. [30].

In recent decades several retrieval codes have been developed for the analysis of SWIR nadir spectra at different institutes, e.g., the Weighted Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS) algorithm [31,32], the Iterative Maximum A Posteriori (IMAP)-DOAS [33] method, the Iterative Maximum Likelihood Method (IMLM) [34], the Shortwave Infrared CO Retrieval (SICOR) algorithm [35,36], and the Beer InfraRed Retrieval Algorithm (BIRRA) [37]. Developed at the DLR for the retrieval of vertical column densities (VCD) from space-borne SWIR nadir observations BIRRA serves as ESA's operational SCIAMACHY processor for the CO and CH_4 products.

The objective of this work is an assessment of the impact of molecular spectroscopy on the SCIAMACHY retrievals of CO. Despite the well-known problems of SCIAMACHY's channel 8 measurements [34,38,39] (considerable noise, dead&bad pixels, ice layer contamination of the detector) this study is appropriate for several reasons. First, SCIAMACHY and TROPOMI feature very similar spectral characteristics in the 2.3 µm channels—the nominal spectral resolution is 0.26 nm and 0.25 nm,

respectively. Second, BIRRA has been validated in terms of accuracy and precision using SCIAMACHY observations regarding NDACC (Network for the Detection of Atmospheric Composition Change) and TCCON (Total Carbon Column Observing Network) [40] and it was found to be largely consistent with findings by Borsdorff et al. [35,36]. Furthermore, a coherent time series of CO comprising measurements from different instruments (e.g., TROPOMI is regarded as SCIAMACHY's successor) requires a harmonized and consistent description of physical processes, hence, improved molecular spectroscopy is relevant for SCIAMACHY, too. Please note that ESA plans another reprocessing of SCIAMACHY Level 2 data within the next two years (G. Lichtenberg, personal communication).

This paper is organized in four sections: The methodology is described in the following Section 2. It includes a brief description of infrared radiative transfer with high spectral resolution in Section 2.1, a short review of the retrieval algorithm BIRRA and its updates in Section 2.2 and a few aspects on pre- and postprocessing in Section 2.3. The main results are discussed in Section 3 (some supplementary material is provided in the Appendix). Finally, the study concludes and summarizes its findings in Section 4.

2. Methodology: Forward Modeling and Inversion

2.1. The Forward Model: SWIR Radiative Transfer

The quantitative retrieval of atmospheric constituents requires an accurate description of the radiative transfer [41–43]. In the SWIR region, the transfer of radiation is simplified because thermal emission of Earth's atmosphere and surface is negligible during daytime compared to reflected and scattered solar radiation. Hence, the radiation seen by a space-borne observer is assumed to be essentially the downwelling solar radiation I_{sun} reflected at the surface or clouds and traveling back to space,

$$I(\nu) \;=\; \frac{r(\nu)}{\pi}\, \mu_\odot\, I_{sun}(\nu)\, \mathcal{T}_\uparrow \mathcal{T}_\downarrow \otimes \mathcal{S} \tag{1}$$

$$\mathcal{T}_\uparrow \mathcal{T}_\downarrow \;=\; \exp\left(-\sum_m \tau_m(\nu, s', s'')\right) \;=\; \exp\left(-\sum_m \int\limits_{\substack{\text{double}\\\text{path}}} n_m(s)\, k_m(\nu, s)\, \mathrm{d}s\right), \tag{2}$$

where r is the albedo depending on wavenumber ν. $\mathcal{T}_\uparrow \mathcal{T}_\downarrow$ describes the monochromatic transmission for two path segments s' and s'' of the up- and downwelling radiation (Sun to Earth and Earth to satellite) according to Beer's law and $\otimes$ denotes the convolution of that spectrum with the spectral response function $\mathcal{S}$ (SRF) modeling instrumental effects. The attenuation is determined by the molecular optical depth τ_m of molecule m described by the double path integral of the volume absorption coefficient, i.e., the absorption cross section k_m scaled by the molecular number density n_m.

2.1.1. Absorption Cross Section and Line Profiles

In high resolution line–by–line models, line position $\hat{\nu}$, line intensity S, line width γ (air- and self-broadening), temperature exponent n and lower-state energy E are mandatory line parameters for the determination of the absorption cross section k at different pressure p and temperature T levels. The cross section of a molecule k_m is calculated by the superposition of many lines l with line center positions $\hat{\nu}_l^{(m)}$ determined by the difference of initial and final state energies E_i and E_f, according to $\hat{\nu} = \frac{1}{hc}(E_f - E_i)$. Each line is described by the product of a temperature dependent line strength $S_l^{(m)}$ and a normalized line shape function $\int\limits_{-\infty}^{+\infty} g(\nu)\, \mathrm{d}\nu = 1$ modeling the broadening mechanisms (for conciseness the index m is omitted),

$$k(\nu, p(s), T(s)) \;=\; \sum_l S_l(T(s))\, g\left(\nu; \hat{\nu}_l, \gamma_l(p(s), T(s))\right). \tag{3}$$

For a long time, the Voigt profile g_V has been the standard for line-by-line modeling of infrared and microwave radiative transfer. It represents the combined effect of pressure and Doppler broadening [44]. The Voigt function is identical to the real part of the complex error function and can be numerically calculated by e.g., using the rational approximations of Humlíček [45] or Weideman [46] (also see Schreier [47]).

However, the increasing quality in atmospheric spectroscopy observations has indicated that physical processes beyond pressure and Doppler broadening should be treated. The assumptions underlying the Voigt profile may break down when modeling highly resolved spectra from latest sensors [48]. To compute line shapes beyond Voigt the imaginary component of the complex error function can be employed to model higher-order effects in molecular absorption such as the speed-dependence of air broadening [49–52], collisional narrowing [53] or Rosenkranz line-mixing [51,54,55]. Depending on the line profile an additional set of line parameters is required.

Velocity changes due to collisions lead to the "Dicke narrowing" of the line shape, described by e.g., the Rautian (RTN) profile g_R with an extra parameter ν_{vc} for the frequency of velocity changing collisions [48,53,56].

The effect of line-mixing arises for lines which are close together in wavenumber [57]. According to Boone et al. [51], Ngo et al. [58] the effect on the line profile can be modeled by the Rosenkranz approximation [54] which includes a coupling coefficient Y (Rosenkranz parameter).

The speed-dependent Voigt (SDV) profile g_{SDV} refines the pressure broadening component of the Voigt profile. It introduces two extra parameters that represent the speed-dependence of the pressure broadening γ_2 and line shift δ_2 [48]. Finally, in order to calculate the combined effect of speed-dependence and line-mixing (SDVM) a simple empirical extension of the first order Rosenkranz approximation suggested by Boone et al. [51], Ngo et al. [58] was used.

The SEOM line parameters [19] have been obtained by nonlinear least squares fitting using the partially Correlated quadratic Speed-Dependent Hard Collision (pCqSDHC) model including line-mixing. The pCqSDHC model, generally called the Hartmann-Tran (HT) profile g_{HT} [48,58,59] basically models the combined effects of speed-dependence and narrowing. Note, however, that the parameter η quantifying the partial correlation between velocity and rotational state changes has not been fitted in SEOM ($\eta = 0$), hence the HT profile is equivalent to the speed-dependent Rautian (SDR). With line-mixing included the speed-dependent Rautian (SDRM) model includes seven parameters in total.

2.1.2. Spectroscopic Line Data

In recent decades new releases of the HITRAN and GEISA database were made available every few years. The latest versions, namely HITRAN 2016 and GEISA 2015, provide updated line data for the computation of the Voigt profile and are summarized in Table 1. It should be noted that beside the classical HITRAN 2016 line list several extended formats are available to include line parameters beyond Voigt. In the 2.3 µm region, so far, only some CO lines have parameters for the SDVM and SDRM profiles while CH_4 and H_2O do not (details see [60], ([61] Section 2.5, 2.6), [62], ([12] Section 2.1, 2.5, 2.6]). In contrast, in this SWIR region the SEOM database provides an extended set of line parameters for all three molecules. Hence, it was decided to compare the classical variants of HITRAN and GEISA with the more advanced SEOM line list (also see Appendix A.3).

For the retrieval of CO in the SWIR the strongest transitions of the R-branch are exploited. The interval between 4277.20–4302.90 cm^{-1} (containing the R04-R11 branch lines) was chosen to exclude the interference of strong H_2O lines above 4303 cm^{-1} (see ([37] Figure 5)). Since there is still a strong interference with CH_4 and H_2O in the chosen spectral range both species must be incorporated in the retrieval. In this respect the significant updates of CH_4 and H_2O line data in SEOM play an important role for the purpose of the CO retrieval examined in this study. Please note that lbl modeling within a given interval (e.g., 4277.20–4302.90 cm^{-1}) requires a symmetric extension of $\pm$ 25 cm^{-1} in order to account for line wings and for consistency with the Mlawer–Tobin–Clough–Kneizys–Davies (MT-CKD) continuum [63]. A survey of line parameters in this extended interval is given in Table 1.

Table 1. A comparison of GEISA 2015, HITRAN 2016, and SEOM line data for the 4252.20–4327.90 cm^{-1} region. The number of spectral lines along with the corresponding range of intensities S, air broadening HWHM γ_{air} and temperature exponents n are shown. In the last two columns the mean of the non-zero air-broadening speed-dependence parameters $\gamma_{air}^{(2)}$ and the Dicke narrowing parameters ν_{vc} are given, the number of non-zero values is indicated in the parentheses. In GEISA HDO is treated as an individual species, in the table here these 660 lines are included in the H_2O entry. Please note that in HITRAN 2016 the CH_4 line at $\hat{\nu} = 4270.377259\,cm^{-1}$ has an incorrect temperature dependence of $n = 7.7$ (correct $n = 0.7$) and lines at $\hat{\nu} = 4255.703126\,cm^{-1}$, $4291.829949\,cm^{-1}$, and $4291.840848\,cm^{-1}$ list $n = 0$, respectively (I.E., Gordon, personal communication). Furthermore, the lower-state energy of the $\hat{\nu} = 4270.377259\,cm^{-1}$ line is incorrectly set to $E = 293.1266\,cm^{-1}$ in GEISA 2015, the correct value is $E = 1556.481\,cm^{-1}$ as in HITRAN 2016.

Gas	Data	# Lines	$S\left[cm^{-1}/molec\,cm^2\right]$	$\gamma_{air}^{(0)}\left[cm^{-1}\right]$	n	$\gamma_{air}^{(2)}\left[cm^{-1}\right]$	$\nu_{vc}\left[cm^{-1}\right]$
CO	G15	160	$1.1 \cdot 10^{-36} - 3.6 \cdot 10^{-21}$	$0.040 - 0.079$	0.69		
	H16	110	$1.0 \cdot 10^{-31} - 3.5 \cdot 10^{-21}$	$0.042 - 0.081$	$0.67 - 0.79$		
	SEOM	110	$1.0 \cdot 10^{-31} - 3.5 \cdot 10^{-21}$	$0.042 - 0.081$	$0.67 - 0.79$	$0.00607(10)$	$0.0047(10)$
CH_4	G15	7213	$9.5 \cdot 10^{-28} - 5.4 \cdot 10^{-21}$	$0.034 - 0.077$	$0.46 - 0.97$		
	H16	8375	$1.0 \cdot 10^{-29} - 5.5 \cdot 10^{-21}$	$0.034 - 0.077$	$0.67 - 0.77$		
	SEOM	6205	$6.7 \cdot 10^{-27} - 5.5 \cdot 10^{-21}$	$0.019 - 0.182$	$0.19 - 1.82$	$0.008405(610)$	$0.00911(65)$
H_2O	G15	1101	$8.4 \cdot 10^{-30} - 2.2 \cdot 10^{-23}$	$0.004 - 0.096$	$0.32 - 0.69$		
	H16	1197	$1.0 \cdot 10^{-32} - 2.2 \cdot 10^{-23}$	$0.004 - 0.109$	$0.32 - 0.73$		
	SEOM	1177	$1.4 \cdot 10^{-30} - 2.2 \cdot 10^{-23}$	$0.004 - 0.141$	$0.31 - 1.02$	$0.007197(35)$	$0.01083(12)$

Regarding CO, the number of lines in HITRAN were reduced from 2012 to 2016 by about one third, whereas in GEISA 2015 the number of lines has approximately doubled regarding its predecessor [12,13,61,64]. In addition, about 20% more methane lines are included in GEISA 2015 due to the inclusion of numerous weak lines. In contrast, the number decreased for the latest version of HITRAN despite new lines being added for $^{12}CH_4$ and $^{13}CH_4$. The number of water lines relevant in the spectral window has roughly doubled in the latest version of HITRAN and increased by a factor of four in GEISA. Also note that the exponent n characterizing the temperature dependence of the Lorentz width is a constant for CO in GEISA 2015, but variable in HITRAN 2016 and SEOM.

Figure 1 shows the spectral line intensities S and cross sections k in the selected retrieval window. In case of CO, line data agree well for the various databases but in case of CH_4 and H_2O line strengths show less agreement, esp. for weak lines. Since the radiance is depending only on the product $S \cdot n_m$ uncertainties in the line strength map into corresponding uncertainties in the molecule number densities to be retrieved.
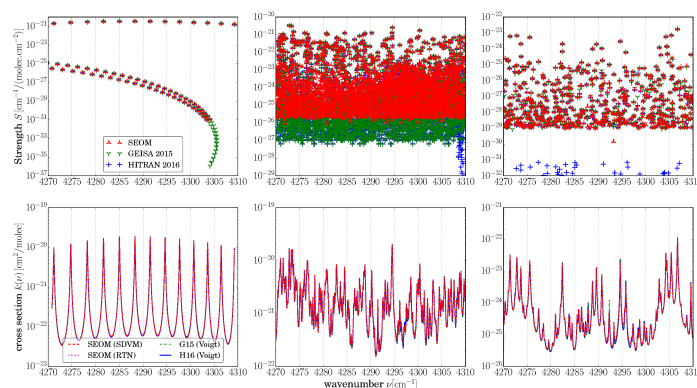

Figure 1. (Top) Line strengths S of CO (left), CH_4 (center), and H_2O (right) for the three databases in Table 1. While the strong lines are similar in all databases the values do not agree completely. (Bottom) Cross sections of the three molecules at reference pressure (1013.25 hPa) and temperature (296 K).

2.2. Inversion and Its Implementation—BIRRA

In principle, the retrieval of the vertical column density $N_m = \int n_m(z)\,dz$ of molecule m is equivalent to the problem of finding a scaling factor α_m that is relating a (climatological) reference profile (e.g., US-Standard) to the actual profile $n_m(z) = \alpha_m n_m^{(\text{ref})}(z)$.

The retrieval setup in BIRRA for this study comprises a state vector **x** that includes the scaling factors α_m of the reference optical depths of the molecules CO, CH_4 and H_2O, three coefficients for the second order polynomial representing the surface reflectivity and optionally a wavenumber shift and the half width of the Gaussian SRF [37,40]. Atmospheric data for pressure, temperature, and water vapor concentrations were taken from the NCEP reanalysis [65] which provides four profiles per day since 1948 with a 2.5° latitudinal and longitudinal resolution. Methane and carbon monoxide a priori profiles are taken from the AFGL dataset [66]. After convergence, BIRRA stores the fitted state vector elements (scaling factors and auxiliary parameters) along with their error estimates, the initial and final residual norms, number of iterations, a set of further quality indicators and a proxy normalized CO column densities ([37] Section 2.3.1) in a file.

The upgrades in the algorithm for this study primarily include enhancements in the GARLIC (Generic Atmospheric Radiation Line-by-line Infrared Code [67]) forward model, e.g., line profiles beyond Voigt which use additional line parameters given in the SEOM database.

2.3. Data Preparation and Postprocessing

A subset of SCIAMACHY nadir measurements in the years 2003, 2004 and 2005 were selected to study the impact of latest spectroscopic line data with corresponding models. Those years were chosen since both are within the period before instrument degradation dropped the performance of the SWIR channel significantly (see ([37] Figure 12)) and ground-based reference observations from NDACC and TCCON are available. To mitigate the impact of sensor deficiencies the creation of a consistent dead and bad pixel mask (DBPM) was important. In a preprocessing step, a DBPM was generated that only contains 'good' pixels for the entire analyzed period. Pixels in the Level 1 product that had been flagged 'bad' once within this time frame were ignored for the retrieval. It was found that $\approx$30–40% of the pixels are bad resulting in 80 good pixels available within the CO fitting window between $4277.20\,\text{cm}^{-1}$ and $4302.90\,\text{cm}^{-1}$ on October 27 2003. The number of good pixels dropped to 71 for the investigated periods in 2004 (April–June) and 2005 (August–October).

Another issue is the ice layer existing on SCIAMACHY's channel 8 detector which is affecting all pixels and leads to a change of the SRF [34,68]. To minimize its impact on the retrievals the slit function half width is treated as an additional unknown (see Section 2.2).

Postprocessing includes the conversion of total columns to column-averaged dry-air mole fractions, designated as xCO ([40] Section 2). Errors (obtained from the diagonal elements of the least squares covariance matrix ([37] Section 4.2.1)) of the molecular scaling factors were used to eliminate outliers. BIRRA flags non-converged retrievals and those were also filtered out. It was found that most of the outliers in the retrieved parameter arise from measurements with extremely small signal-to-noise (SNR) (see Table 4). Observations that indicate an enhancement in the light path by, for example, aerosols or optically thin clouds were rejected by using OCRA (Optical Cloud Recognition Algorithm) [69] cloud fractions and SACURA (Semi-Analytical Cloud Retrieval Algorithm) [70] cloud top heights. Please note that no bias correction was applied to the retrieval results.

3. Results

In this section, the quality of the retrieval output is assessed based on the analysis of the spectral fitting residuals and the impact on CO mole fractions. Both quantities were investigated for individual orbits, different climatological regions and globally, including a comparison to NDACC and TCCON ground-based measurements. An overview of the experiments conducted is given in Table 2.

Table 2. Outline of the experiments conducted in the respective sections.

	Quantity	Spatial Extend	Temporal Coverage
Section 3.1.1	spec. residuum	single orbits	overpasses in 2003 & 2004
Section 3.1.2	spec. residuum	regional scale	April 2004
Section 3.2	mole fraction	regional scale	April–June 2004
Section 3.3	mole fraction	within 1000 km from site	April–June 2004 & August–October 2005
Section Appendix A.1	mole fraction	global scale	April–June 2004

The classic Voigt line profile was used with the HITRAN 2016 and GEISA 2015 line lists, denoted as H16 and G15 subsequently. Initially, the Voigt and more sophisticated profiles, namely the SDRM, the speed-dependent Voigt with line-mixing (SDVM) and the Rautian (RTN) model were used with SEOM data as input. However, strictly speaking, the classic Voigt profile is not an adequate line model for the SEOM line list ([19], M. Birk, personal communication) but was initially also considered in order to discriminate the impact of data versus model. This discussion is deemed inevitable since SEOM and "beyond Voigt" is strongly linked.

3.1. Spectral Fitting Residuals

The relative change of the norm of the residuum vector

$$\rho = \mathbf{I}_{\text{obs}} - \mathbf{I}(\mathbf{x}) \tag{4}$$

is one of the criteria defining convergence in BIRRA ([37] [Section 3.9 and 4.2.1]). The better the forward model $\mathbf{I}(\mathbf{x})$ can mimic the measurements $\mathbf{I}_{\text{obs}}$ the smaller the scaled norm of the spectral fitting residual

$$\sigma^2 = ||\rho||^2 / (m - n) \tag{5}$$

becomes after successful iteration. It is, therefore, a suitable quantity to assess deficiencies in the forward model of the retrieval. In Equation (5) m designates the number of measurements, i.e., 80 and 71 spectral points for orbit 8663 (October 27, 2003) and 13212 (September 09, 2004), respectively, and n stands for the number of elements in the state vector $\mathbf{x}$, i.e., the number of unknown quantities.

It is important to note that in order to obtain the best estimate of the state vector in the fitting procedure the spectral residuals ρ are assumed to be random errors caused by measurement errors (instrument noise, etc.) following a normal distribution with expected value 0 and an $m \times m$ positive definite variance matrix.

3.1.1. Single Orbits

Observations in orbit 8663 and 13212 were used to examine the spectral residuals for different spectroscopic line data along with their corresponding profile(s). Orbit 8663 was chosen since it is mainly over land including Eastern Africa, the Arabian Peninsula and Russia while measurements in orbit 13212 covers parts of the Indian Ocean and the South China Sea as well as large fractions of the polluted areas in eastern China.

Table 3 shows the average of the scaled norm of the spectral fitting residual $E(\sigma^2)$ along with the median and standard deviation of the residuum vector ρ for both orbits and different combinations of line data and models. The residuals are similar; however, they vary across the different spectroscopic inputs. Especially observations in the early stages of the mission underline that to minimize the norm of the residuals the use of an appropriate line model is crucial. The scaled norm residuals for the SEOM-Voigt (VGT)-based retrievals are similar to the Rautian (RTN)-based fits but still larger compared to the SDRM and SDVM cases. This indicates that the updated SEOM line data is used optimally if an appropriate line model is chosen.

Table 3. Mean of the norm of the spectral fitting residuals along with the median and standard deviation of the residuum vector for various combinations of line data and models in orbit 8663 (left column) and 13212 (right column). Note that for both the speed-dependent Rautian and speed-dependent Voigt profiles, line-mixing was taken into account.

Data	Model	$E(\sigma^2)$		Median ρ		std. dev. ρ	
SEOM	SDRM	$1.90 \cdot 10^{-2}$	$1.71 \cdot 10^{-2}$	$3.19 \cdot 10^{-5}$	$1.97 \cdot 10^{-4}$	$2.86 \cdot 10^{-3}$	$2.72 \cdot 10^{-3}$
SEOM	SDVM	$1.92 \cdot 10^{-2}$	$1.70 \cdot 10^{-2}$	$3.49 \cdot 10^{-5}$	$2.00 \cdot 10^{-4}$	$2.87 \cdot 10^{-3}$	$2.72 \cdot 10^{-3}$
SEOM	Rautian	$1.92 \cdot 10^{-2}$	$1.73 \cdot 10^{-2}$	$4.78 \cdot 10^{-5}$	$1.94 \cdot 10^{-4}$	$2.88 \cdot 10^{-3}$	$2.74 \cdot 10^{-3}$
SEOM	Voigt	$1.93 \cdot 10^{-2}$	$1.73 \cdot 10^{-2}$	$4.32 \cdot 10^{-5}$	$2.03 \cdot 10^{-4}$	$2.88 \cdot 10^{-3}$	$2.73 \cdot 10^{-3}$
HITRAN 2016	Voigt	$1.96 \cdot 10^{-2}$	$1.75 \cdot 10^{-2}$	$3.23 \cdot 10^{-5}$	$1.94 \cdot 10^{-4}$	$2.91 \cdot 10^{-3}$	$2.75 \cdot 10^{-3}$
GEISA 2015	Voigt	$1.98 \cdot 10^{-2}$	$1.77 \cdot 10^{-2}$	$2.45 \cdot 10^{-5}$	$2.07 \cdot 10^{-4}$	$2.92 \cdot 10^{-3}$	$2.76 \cdot 10^{-3}$

The histograms of the spectral residuals for orbit 8663 are depicted in Figure 2 each contains $\approx 1.5 \cdot 10^5$ data points. To examine whether the individual SEOM distributions are drawn from the same distribution regarding H16 (null hypothesis) the non-parametric Kolmogorov—Smirnov test [71] was chosen since for large sample sizes even small values of skewness and kurtosis can compromise the analysis and results of a parametric statistical test. It was found that for observations in orbit 8663 the null hypothesis can be rejected for ordinary significance levels up to 1% with p-values of $3.50 \cdot 10^{-7}, 6.33 \cdot 10^{-6}, 3.33 \cdot 10^{-3}$ and $7.81 \cdot 10^{-4}$ for SDRM, SDVM, RTN and VGT, respectively. Please note that G15 is not depicted in Figure 2 since the results are similar to that of H16.
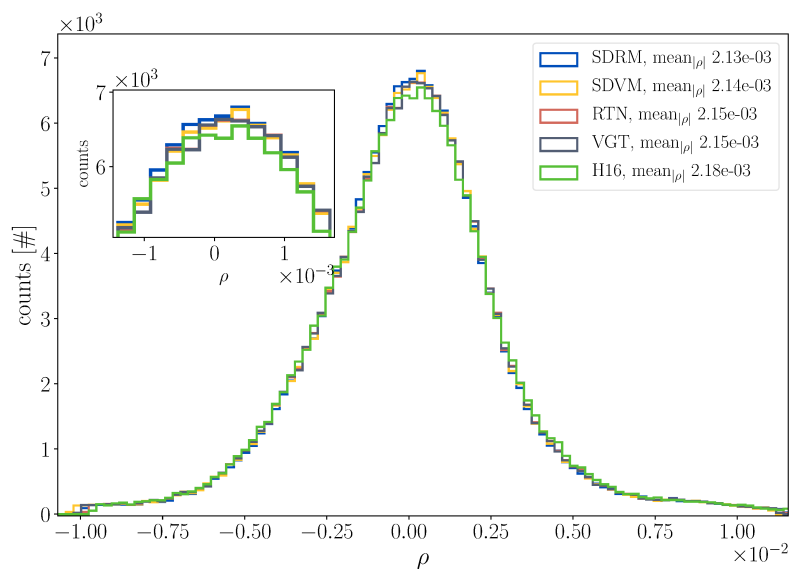
Figure 2. Histogram of the spectral residuals in orbit 8663 for SEOM and HITRAN 2016.

In Figure 3 the differences in the spectral fitting residuals for a set of line data and model combinations from Table 3 are depicted for orbit 13212. The y-axis corresponds to observations from southern to northern latitudes while the x-axis shows the 71 pixels (some with their wavenumber assignment) in the retrieval window. The figure indicates that the spectroscopic line data has a pixel-dependent impact in the forward model. Most pixels reveal rather small positive and negative differences; however, the patterns across the six cases differ. It shows that the SDVM and Rautian line shapes cause similar residuals and that individual pixels show significantly larger differences. A common feature is that positive and negative values are often preserved throughout the entire orbit (unveiled by the median and vertical regime of red or blue). This indicates that postprocessing has filtered out most of the clear-sky observations over water bodies as the significantly lower SNR of those measurements would be expected to reveal in the corresponding residuals.
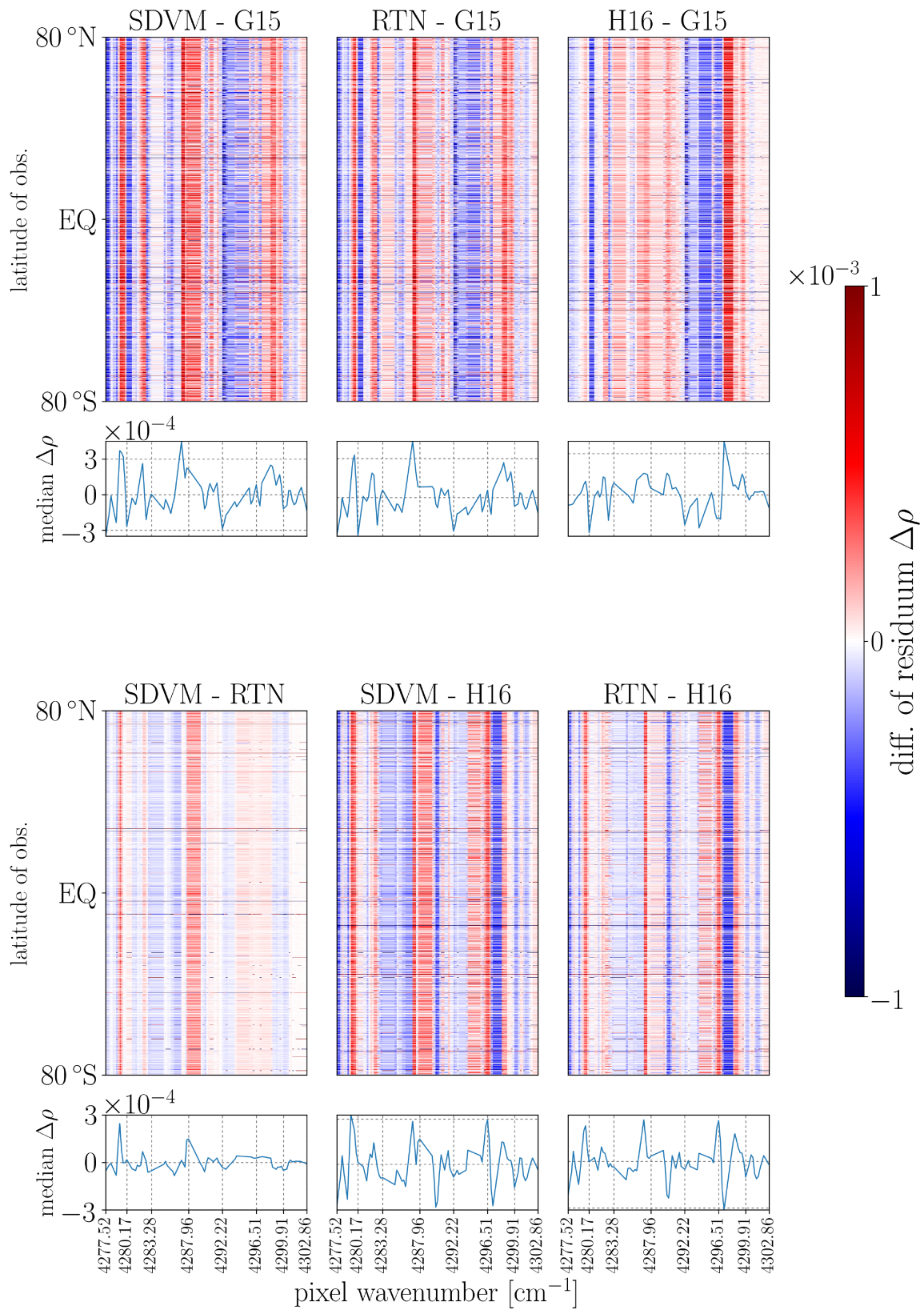
Figure 3. Difference in the spectral fitting residuals for various spectroscopic inputs and line models. The x-axis of each figure shows the 71 sensor pixels in the CO fitting window. The y-axis in the pseudocolor plots depicts 2090 observations in orbit 13212 from southern to northern latitudes and the lower panels show the median of the residuals for each pixel.

SCIAMACHY derived columns are greatly affected by a precision error due to instrumental noise which shows substantial variations as it strongly depends on the SNR of the recorded spectra and so on surface albedo [38,72,73]. The retrieved CO total columns depicted in Figure 4 show that using the SEOM line list increases the columns regarding the HITRAN database by $\approx$ 6–8%. The systematic rise of the individual columns is also indicated by the shift of the SEOM histograms towards larger CO values. The means of the five cases suggest that the increase is mainly caused by updated line parameters (line strength etc.), although an appropriate higher-order line shape model for the SEOM data additionally raises the columns. The standard deviations are very similar between 1.18–1.19 · 10^{18} molec cm^{-2} and the medians of the distributions are ranging from 1.71–1.84 · 10^{18} molec cm^{-2} with the upper and lower bounds attributed to H16 and SDRM, respectively.
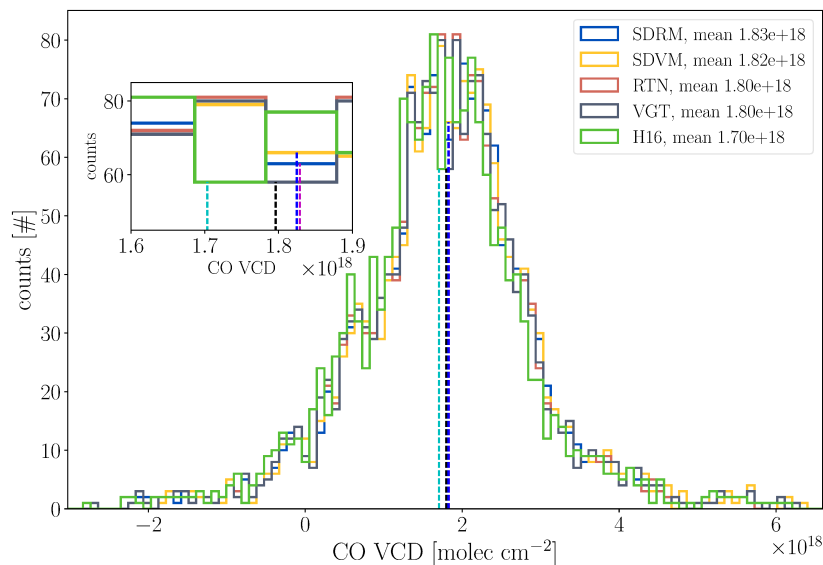
Figure 4. The CO total columns for different line shapes with SEOM input compared to HITRAN in orbit 8663. It shows that SEOM-based retrievals cause larger columns on average (dashed vertical lines).

The assessment of the single SCIAMACHY orbits indicates that including speed-dependence of the relaxation rates and line-mixing from the SEOM line data in the forward model have the largest effect on the retrieved CO columns. Hence, it was decided to apply the SDVM profile for the SEOM line data in the subsequent analysis of SCIAMACHY retrievals, i.e., SDVM $\Leftrightarrow$ SEOM. The HITRAN 2016 line list was chosen as a reference upon which updates in the SEOM line list were assessed.

3.1.2. Regional Scale

The analysis was extended to different climatological regions (see Table 4). Please note that in the assessment of the spectral fitting residuals it was deemed sufficient to include measurements only from April 2004 since every single SCIAMACHY observation contains 71 measurements that can be analyzed. The second quarter was chosen since small solar zenith angles (SZAs) were preferred for the analysis of retrievals in the (mostly) northern hemispheric regions. Four areas with different conditions in temperature and humidity were selected to discriminate the impact of H_2O from other molecules. For example, updates of the H_2O spectroscopy are expected to be more pronounced for measurements over humid areas and less in dry environments. Climatological records show that the attributes dry and hot are well fulfilled in the subtropical Sahara region and a dry and cold climate is prevailing in Siberia. Amazonia is dominated by wet and hot conditions while wet and cold weather is typical for the Canadian West Coast. Since the selected area of the latter region includes adjacent parts of the Pacific Ocean and parts of Alaska it was denoted as CaPaAl. Most ocean pixels are rejected during postprocessing due to high errors in the estimated column scaling factors.

Table 4. The examined regions along with their spatial extent, rate of filtered observations and number of measurements after postprocessing in April 2004. The relative change in the norm of the spectral fitting residual is given in the last column.

Region	Latitude	Longitude	Discarded Obs.	# Survivors	$\Delta\sigma^2/\sigma^2_{H16}$
Sahara	18°N–30°N	004°W–026°E	14%	1234	-6.89%
Siberia	55°N–66°N	066°E–095°E	60%	1104	-10.25%
Amazonia	08°S–03°N	078°W–051°W	54%	1084	0.14%
CaPaAl	49.5°N–58.5°N	117°W–140°W	49%	683	-14.71%

The residual vector in the forward model of the retrieval comprises 71 elements. Table 4 gives the number of observations and relative change in the residual norm regarding HITRAN 2016 for the respective region. The results in Figure 5 show smaller disagreements between the forward model and observations for the retrievals using SEOM data in three of the four examined regions (Amazonia in the second panel shows no significant difference). This finding is in good agreement with the outcome of the analysis for the two orbits presented in Table 3. Especially observations over Siberia and CaPaAl (but also Sahara) reveal that many elements (pixels) of the spectral residuum vector have smaller disagreements with the measured spectrum after convergence. By comparing the position of those pixels on the spectral axis to the lower three panels depicting the optical depths of CO, H_2O and CH_4 one can observe that some coincide with strong vibrational-rotational transitions of the molecules.
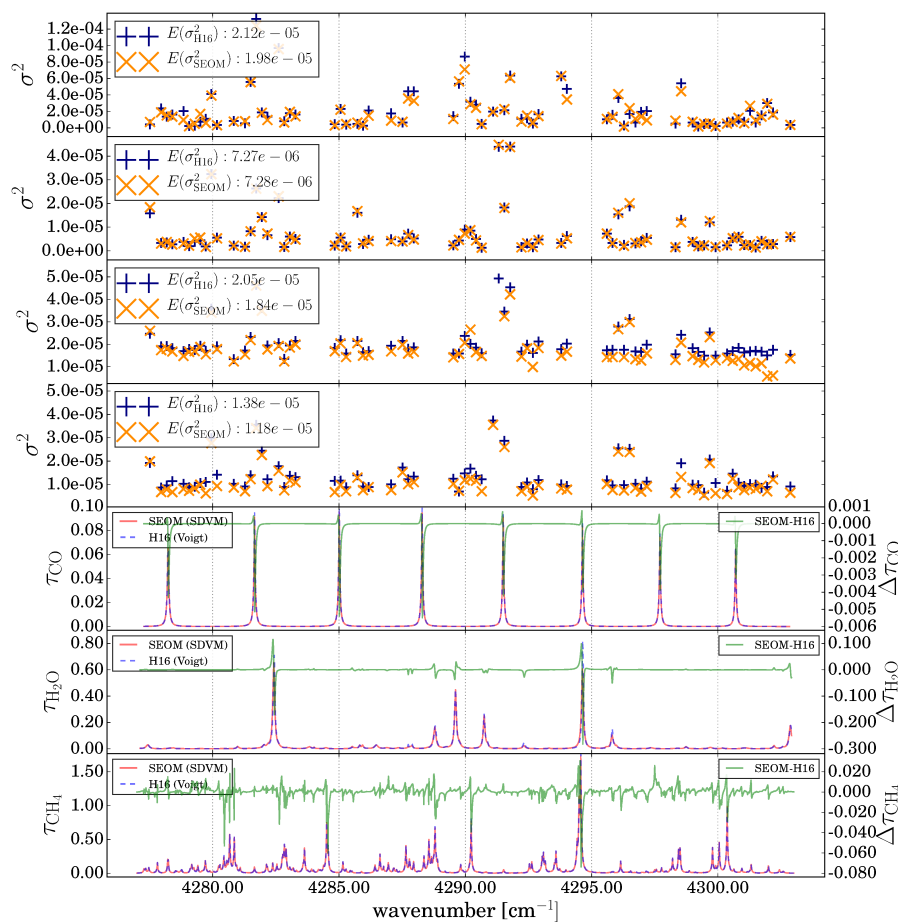
Figure 5. The upper four panels depict the average residual norm in April 2004 for each available SCIAMACHY pixel for the regions Sahara, Amazonia, Siberia and CaPaAl, respectively. Gaps indicate bad pixels according to the DBPM. The lower three panels show the optical depths of the molecules CO, H_2O and CH_4 for SEOM and H16 according to US-Standard along with their differences.

3.2. Differences in CO Mole Fractions

The histogram of xCO for each of the selected regions is depicted in Figure 6. Sahara (top-left) reveals a narrow distribution of CO, only spanning around 200 ppbv including only a few negative values (physically not meaningful on an individual basis). The high surface albedo, resp. reflectivity of this region enhances the SNR and improves the quality of the observed spectra. The narrow distribution of the retrieved xCO is expected to be primarily caused by well represented absorption features of CO, CH_4 (and H_2O) in the spectrum, allowing the retrieval to converge within reasonable values for most measurements. However, mean values in xCO differ by ≈ 8 ppbv across spectroscopic inputs and therefore spectroscopy clearly has an impact on the retrieval results. In contrast to Sahara the spread of CO mole fractions over Siberia (top-right) is larger but the mean values are similar. The SDVM based retrievals deliver the largest mean, i.e., a 2–8% increase. Amazonia (lower-left) shows a similar result, although medians are lower. xCO values also increase by about 7% in the CaPaAl region (lower-right) for SDVM.
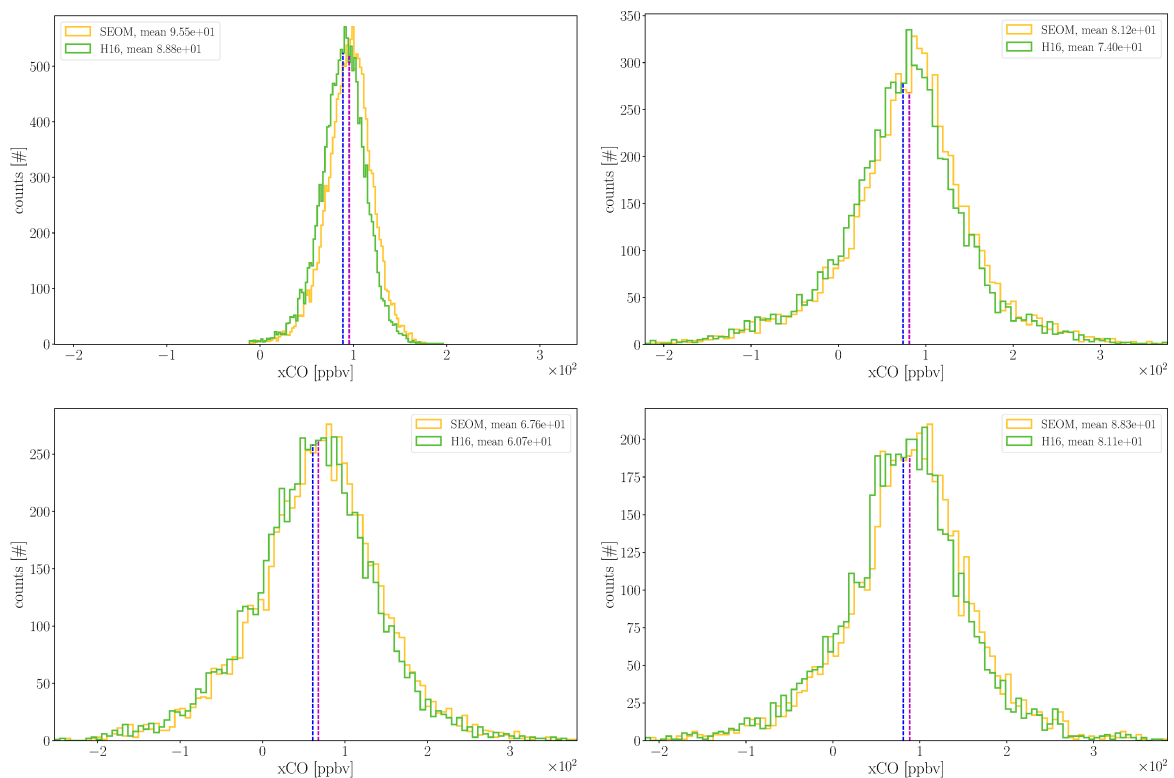
Figure 6. The distribution of xCO in the four analyzed regions Sahara (top-left), Siberia (top-right), Amazonia (lower-left) and CaPaAl for the April–June 2004 period. Please note that the retrieved mole fractions in the Sahara span a significant smaller interval.

The spatial distribution of xCO differences for SEOM and H16 is depicted in Figure 7. Recall that for the analysis of mole fractions the time period was extended to June 2004 (i.e., it includes three months of observations, April through June 2004). A rise in xCO ranging from around 4% to 11% is observed globally. Particularly in the Sahara Desert and the Arabic Peninsula the updated spectroscopy causes a uniform increase in CO mole fractions by around 7%.

The increase in the xCO is not homogeneous but larger towards higher and lower latitudes. Although the increase of xCO is similar towards the north and the south from the equator in absolute numbers the relative change is larger by around 3% in the southern hemisphere.
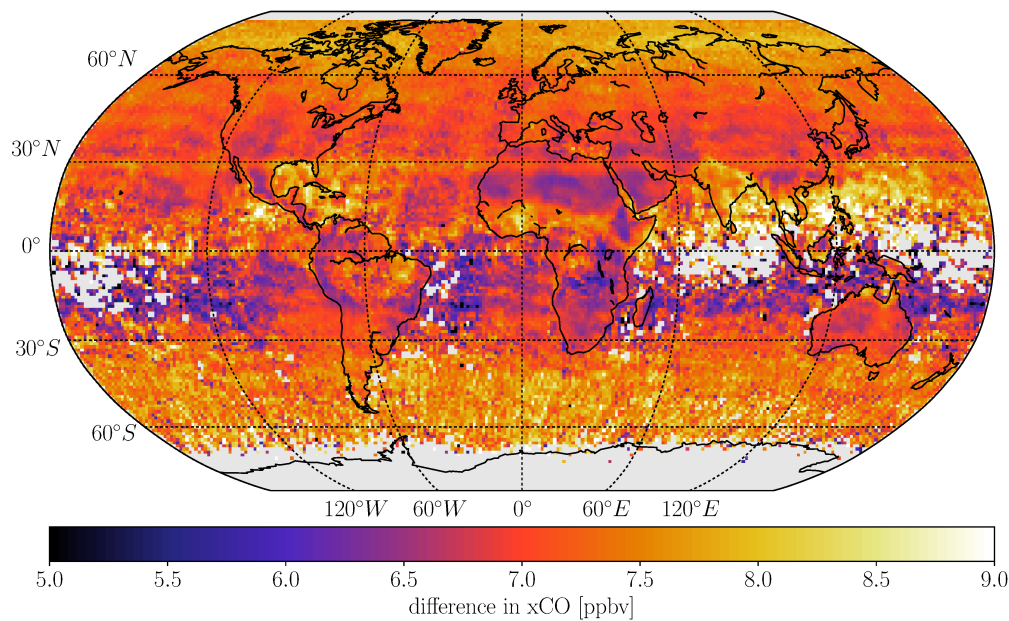
Figure 7. Spatial distribution of differences in CO mole fractions between SEOM and HITRAN 2016 for the April–June 2004 time period. Gray pixels indicate areas where no retrievals were left after postprocessing.

3.3. Comparison to NDACC and TCCON

To assess the quality of the updates on the product level a comparison to ground-based measurements was relevant. In our recent validation study [40] we concluded that the global mean bias of the BIRRA CO retrievals is around -14 parts per billion in volume (ppbv) regarding NDACC, and ≈ 8 ppbv when compared to TCCON. The discrepancy of the global bias between the two networks is likely caused by a correction factor (1.0672) in the GGG2014 dataset (GGG stands for the whole software package) that is applied to TCCON in order to compensate for spectroscopic uncertainties and ties their calibration to the currently accepted World Meteorological Organization (WMO) gas standard scale (D. G. Feist, personal communication and [74–77]). NDACC in contrast does not use such a scaling factor.

On one hand, data coverage for NDACC is guaranteed for many sites in 2004 and comparisons of SWIR satellite Level 2 data with MIR (mid infrared) ground-based products have been presented by several groups (e.g., [35,36,78–80]). However, it records the fundamental band of CO in the MIR while on the other hand, TCCON retrieves total columns from two broad spectral bands in the first overtone spectrum of CO (details see [76]) which overlaps with our fitting window. In order to compare retrievals with SEOM and HITRAN for at least two TCCON sites, namely Darwin (Australia) and Lauder (New Zealand), it was decided to analyze SCIAMACHY observations from August through October 2005, too. This period constitutes a tradeoff between available observations from reference sites and SCIAMACHY's SWIR channel performance, although it was already worse compared to previous years ([40] Figure 3).

According to Rodgers and Connor [81] averaging kernels characterize the altitude sensitivity of a retrieval by relating the true and estimated state vectors. With respect to NDACC and TCCON retrievals Zhou et al. ([77] Figure 3) showed that the column averaging kernels are different due to their different retrieval windows, spectral resolution and retrieval schemes. It was found that the TCCON retrieved CO total column underestimates a deviation from the a priori in the lower troposphere, and overestimates it at high altitudes and that NDACC CO retrievals have a good sensitivity to the whole troposphere and lower stratosphere.

The total column averaging kernels for the BIRRA CO retrieval are shown in Figure 8. In accordance with findings by Gloudemans et al. [28], Buchwitz et al. [31] the averaging kernels are depending on the SZA and, only to a small extent, on the observation angle, a priori trace gas profile and surface albedo. Moreover, as NDACC and TCCON use the HITRAN 2008 (HITRAN 2009 updates for H_2O and HDO) and HITRAN 2012 line list, respectively, the dependence on the spectroscopic input (SEOM and HITRAN 2016) was examined and found to be insignificant. However, for a thorough validation with ground-based networks the use of consistent spectroscopic data across platforms would be beneficial.
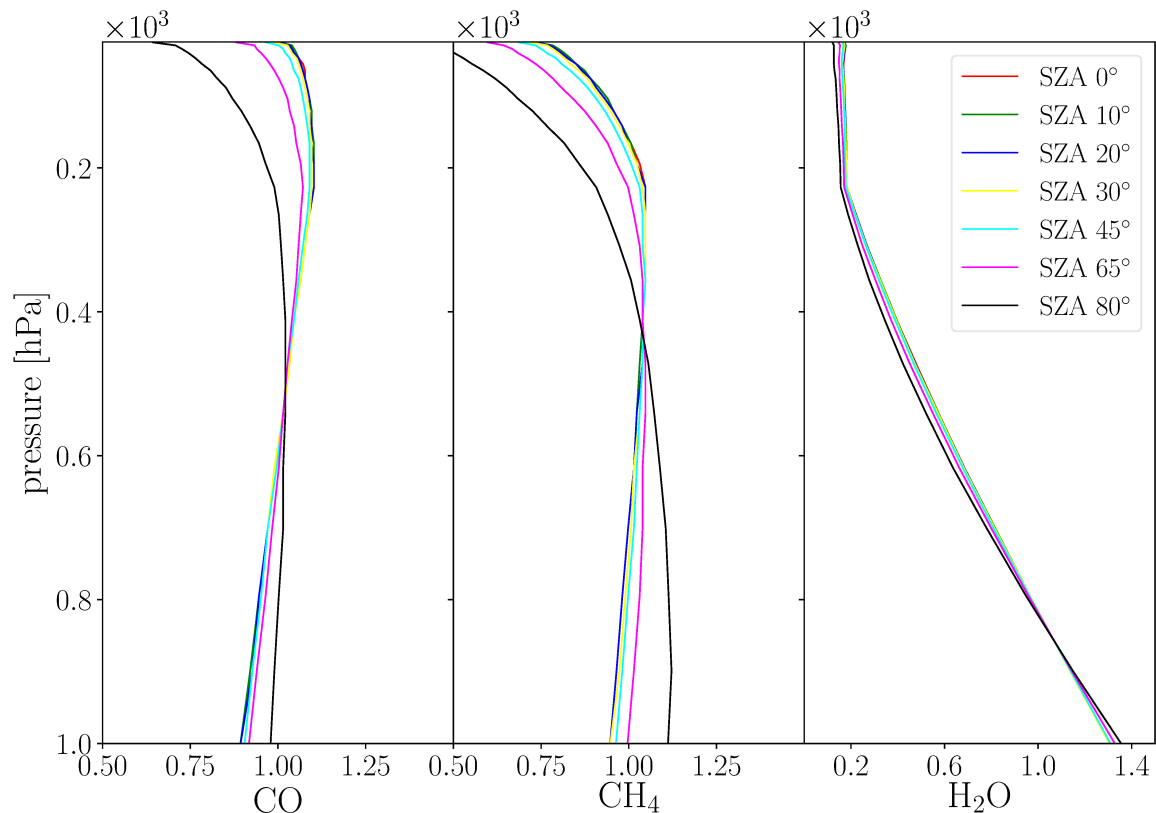
Figure 8. Total column averaging kernels for the BIRRA CO retrieval [4277.20–4302.90 cm^{-1}] in the nadir viewing geometry. The target molecule CO as well as CH_4 reveal a kernel close to unity up to 200 hPa. Measurements with an SZA > 80° are not considered in the retrieval.

Based upon these findings, it was concluded that the smoothing error is negligible when comparing strictly clear-sky SCIAMACHY CO total columns to NDACC and TCCON [81] and that a direct comparison is appropriate to estimate the accuracy of the product. This is in good agreement with Borsdorff et al. [35] who also found that for cloud filtered SCIAMACHY total columns the application of the total column averaging kernels is of little importance.

Figure 9 shows the comparison of averaged CO mole fractions retrieved by BIRRA to NDACC and TCCON ground-based reference sites (Appendix Tables A1 and A2). Systematic differences were calculated according to ([40] Section 2.4 and 3.2.1). A total bias $\bar{b}$ of -5.8 ppbv and -12.8 ppbv was found for the SEOM and HITRAN 2016-based retrievals, respectively.

The decreased quality of SCIAMACHY observations over time resulted in an increased number of rejected measurements during postprocessing for the August–October 2005 period thereby decreasing the ensemble of averaged columns. In addition, strict cloud filtering and large scatter of individual measurements are the main reasons for the large standard error of the biases and hence the bias is only significant for two sites when using SEOM compared to six stations when using HITRAN 2016.
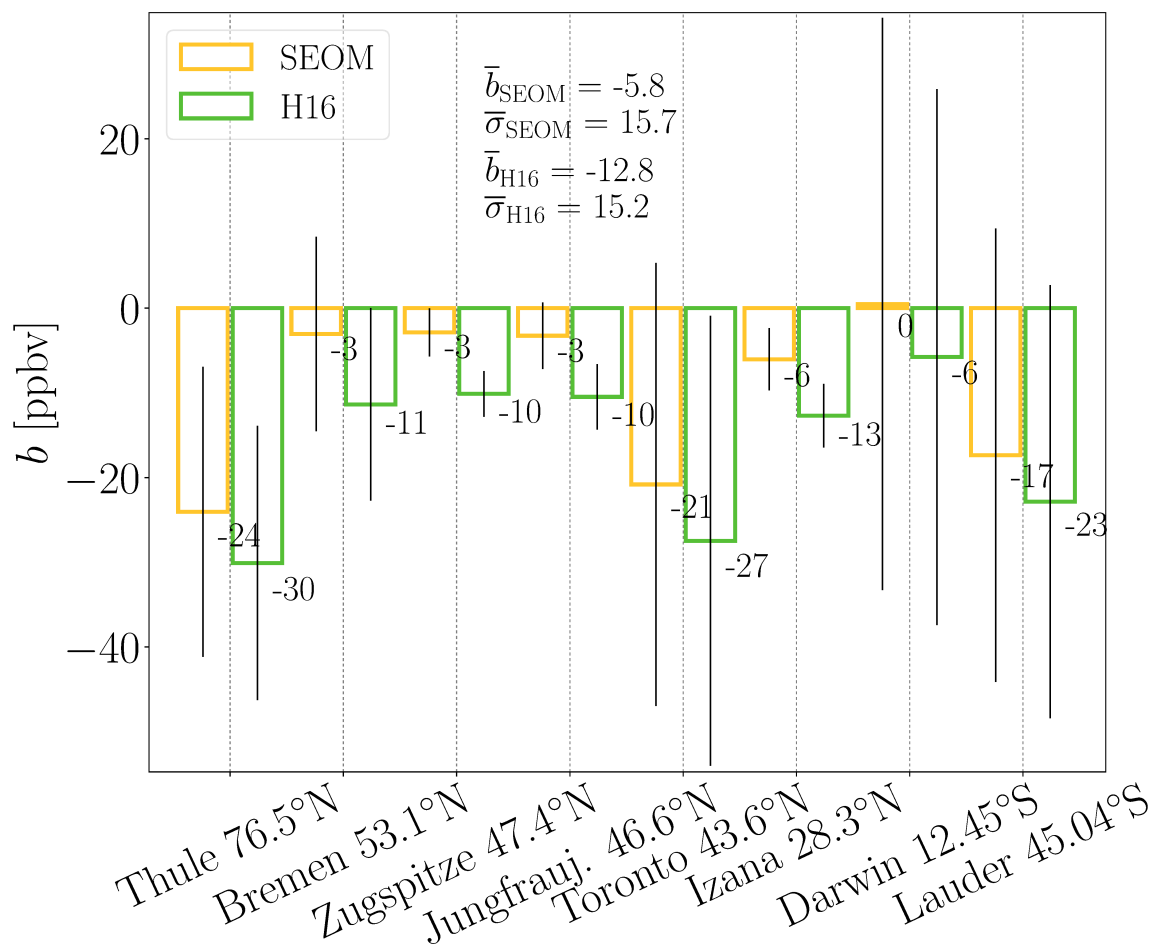
Figure 9. Bias between NDACC/TCCON and SCIAMACHY xCO retrievals within 1000 km (distance weighted) of the ground-based reference sites. Stations Darwin and Lauder belong to TCCON while the others are affiliated to NDACC.

4. Summary and Conclusions

In this study, CO total columns were retrieved in the SWIR from a subset of SCIAMACHY measurements in 2003, 2004 and 2005 using SEOM, HITRAN 2016 (and GEISA 2015) line data with corresponding line profiles. The impact of molecular spectroscopy on spectral fitting residuals and the retrieved quantity was examined and quantified.

This analysis lead to the conclusion that the spectroscopy of CO, CH_4 and H_2O in the 2.3 μm regime has significant impact on the retrieval of CO in the SWIR. The results show that the spectral residuals are reduced with the new line data and corresponding model. It was found that the impact on the fitted residuals is non-homogeneously distributed across the globe and residuals can be reduced up to $\approx 15\%$. The updates in the CH_4 and H_2O lines have a great impact because both molecules are strong absorbers and experienced the most significant updates in SEOM.

The CO mole fractions increased by about ≈ 4–11% reducing the bias to both NDACC and TCCON. It was shown that the largest fraction of the increase is due to updates in the SEOM line parameters and that the line profile only plays a minor role, yet SEOM line data is used optimally when an appropriate line model is chosen. The outcome also confirms recommendations from earlier investigations e.g., by Galli et al. [9] or Checa-García et al. [11], i.e., trace gas retrievals in the SWIR will benefit from improved molecular spectroscopy.

Overall, the findings suggest that the updated line data and models are beneficial for the retrieval of CO from SCIAMACHY in the 2.3 μm regime. The remaining significant bias at three sites might be reduced when averaging over longer time periods and smaller sampling areas.

Although SEOM has been compiled to meet the accuracy requirements of new operational missions such as TROPOMI/S5p, the results of this study suggest that the updated spectroscopy improves the SWIR Level 2 product of SCIAMACHY, too. These findings are important regarding reprocessing and long-term consistency of the CO product since the compilation of a homogeneous multi-mission time series requires consistent forward modeling and harmonized auxiliary data [82].

Author Contributions: P.H. developed the methods, tools, and strategy for this study and performed all retrievals. F.S. originally designed the forward model and retrieval algorithm. The original draft of the manuscript was prepared by P.H. and then reviewed and edited by both authors. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are frequently used in this manuscript:

BIRRA	Beer Infrared Retrieval Algorithm
GARLIC	Generic Atmospheric Radiation Line-by-line Infrared Code
G15	GEISA 2015 spectroscopic database
H16	HITRAN 2016 spectroscopic database
IR	infrared
lbl	line-by-line
MIR	mid infrared
Py4CAtS	PYthon scripts for Computational ATmospheric Spectroscopy
RTN	Rautian line profile
SCIAMACHY	Scanning Imaging Absorption SpectroMeter for Atmospheric CHartographY
SEOM	SEOM – Improved Atmospheric Spectroscopy (IAS) database
SDV	speed-dependent Voigt line profile
SDVM	speed-dependent Voigt line profile with line-mixing
SDRM	speed-dependent Rautian line profile with line-mixing
SWIR	shortwave infrared
VGT	Voigt line profile
xCO	carbon monoxide (CO) column-averaged dry-air mole fraction

Appendix A. Impact on Retrieved and Co-Retrieved Quantities

Subsequently, more details on the global variation of the CO mole fractions and individual molecular scaling factors for the different spectroscopic inputs are provided.

Appendix A.1. Differences in the Column Scaling Factors

In general, SEOM-based retrievals show similar differences in the inferred CO regarding both Voigt spectroscopies (see Figures A1 and A2). The smallest disagreements are found in the subtropical regions, especially the Sahara, the Arabic peninsula and some parts of India.

Another pattern unfolds for CH_4 (middle row) as the results indicate some latitudinal dependence in column differences. SDVM retrievals deliver smaller CH_4 scaling factors in the subtropics compared to Rautian retrievals in Appendix Figure A1.

The comparisons of the retrieved H_2O scaling factors (bottom row) show minor differences for the majority of observations, however, especially around the subpolar regions higher differences are observed.

Finally, the left column in Appendix Figure A3 underlines the fact that the line model has less impact on the scaling factors. The two Voigt spectroscopies in the right column also show the latitudinal pattern in the CH_4 difference.
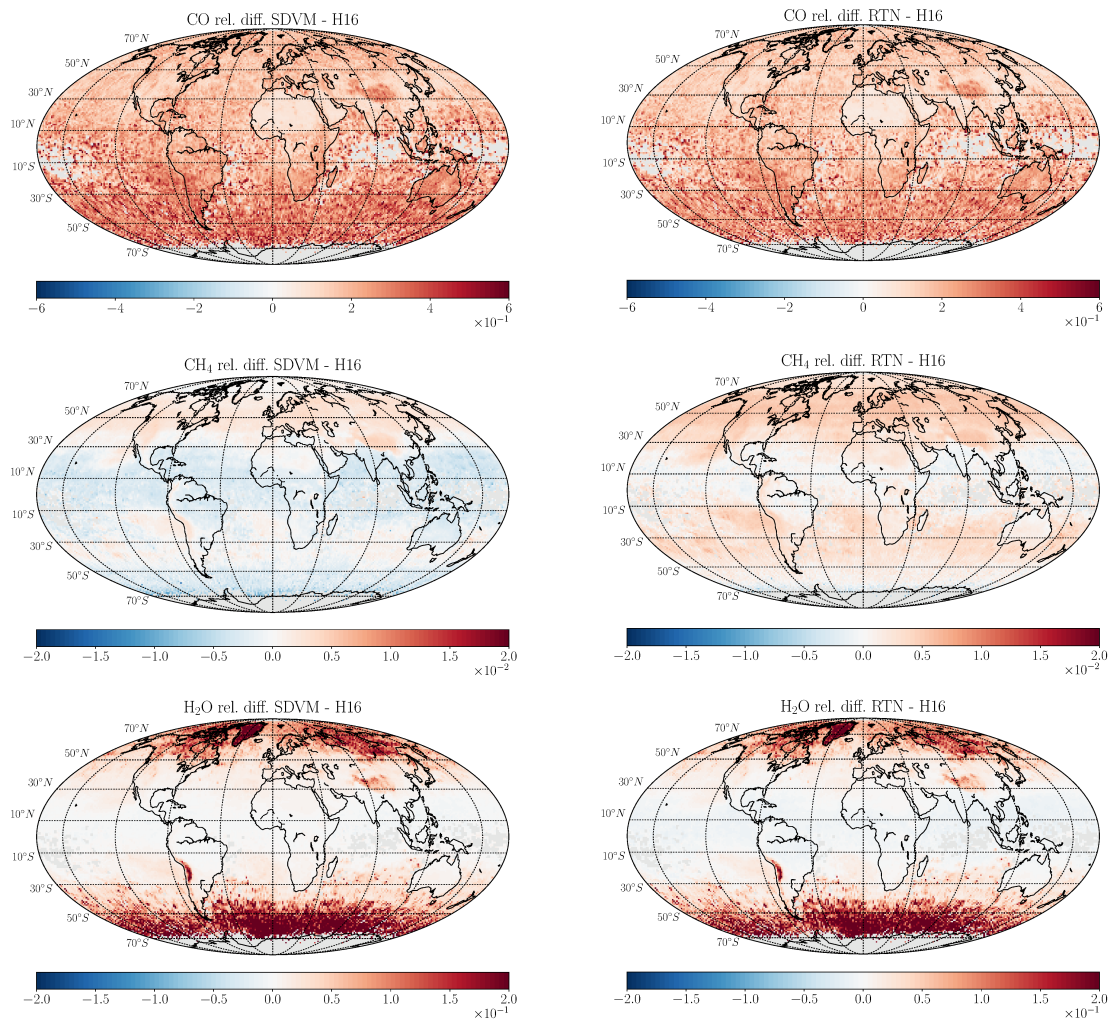
Figure A1. Spatial distribution of differences in the retrieved profile scaling factors on a 1×1 grid from April–June 2004. Left column: SDVM - H16. Right column: RTN - H16 relative difference.
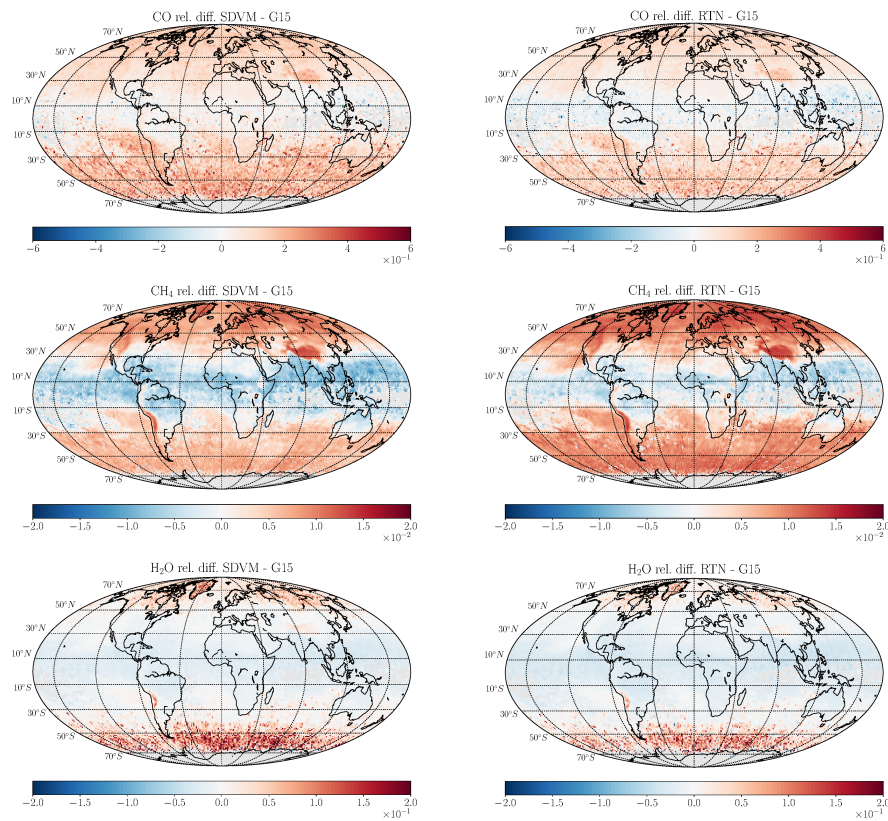
Figure A2. Left column: Relative difference of SDVM to G15. Right column: Rautian to G15 relative difference.
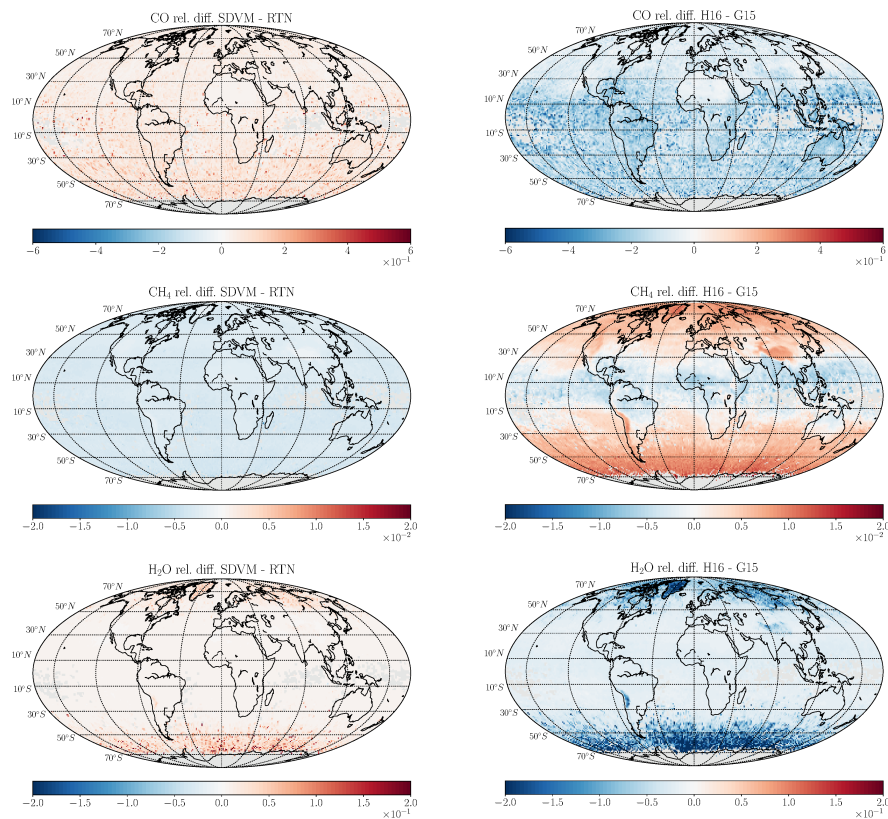

Figure A3. Left column: Global maps of SDVM to Rautian relative difference for CO, CH_4 and H_2O scaling factors. Right column: The relative difference of the two Voigt spectroscopies.

Appendix A.2. Global Map of CO Mole Fractions and Errors

The retrieved CO mole fractions for SDVM are depicted in Appendix Figure A4 (top panel) and reveal largest values in Eastern Asia, and parts of India and Central Africa. The corresponding errors (bottom panel) are smallest in dry regions with low cloud coverage and high surface reflectivity (the Sahara Desert, the Arabian Peninsula, parts of Australia, ...).
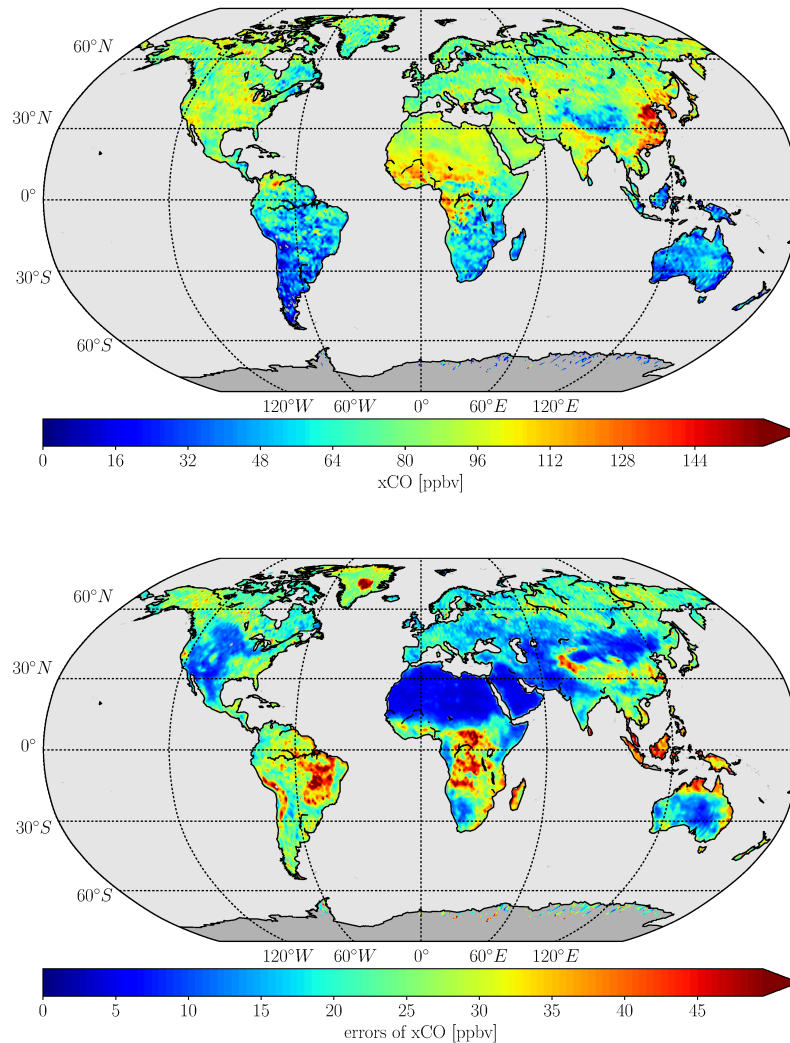

Figure A4. (Top) Global distribution of CO column-averaged dry-air mole fractions for the second quarter in 2004. (Bottom) The corresponding errors.

Appendix A.3. Line Parameters in SEOM and HITRAN 2016

To accommodate the additional line parameters for the molecules CO, CH_4 and H_2O in the SEOM database an extended HITRAN format was used (details see ([19] Section 4)). Beside the classical Voigt parameters entries exist for γ_2, δ_2, Y and ν_{vc} (Section 2.1.1). In case of CO these higher-order estimates are only available for the strongest lines so the (Voigt) parameters of weaker lines were taken from HITRAN 2016 (see Table 1 and Figure 1).

Although it is possible to retrieve speed-dependent parameters for the first overtone band of CO from HITRAN 2016, there are no entries for CH_4 and H_2O. More specifically 24 of the 110 CO lines in the 4252.27–4327.90 cm^{-1} spectral range include the additional speed-dependent Voigt data. However, since CO has a low optical depth and is responsible for only around one percent of absorption in this

spectral region the classical Voigt variant of the HITRAN 2016 database was chosen for this study. Also note that the additional parameters are not included in the standard HITRAN format (since HITRAN 2004 with 160 characters [83]) but one needs to create a user-defined output format (at https://hitran.org) specifying the individual parameters. A comparison of the cross-sections k_{CO} for two variants of the HITRAN 2016 line list and SEOM is shown in Appendix Figure A5.
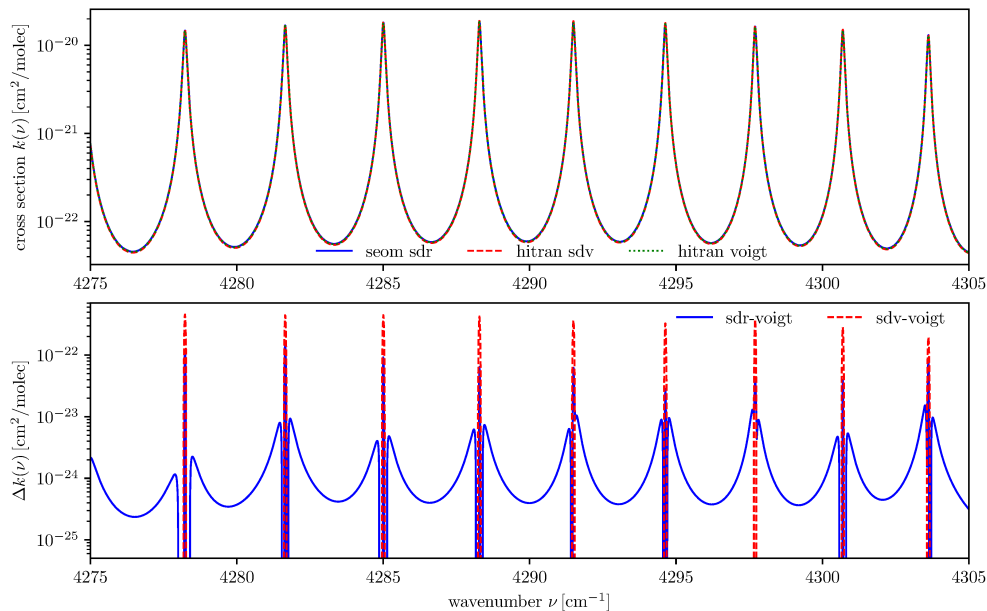

Figure A5. Absorption cross-sections of CO at reference pressure and temperature for SEOM and two variants of the HITRAN 2016 line data (upper panel). In the lower panel the classical HITRAN case (Voigt) is subtracted from the SDR and SDV cross-sections using SEOM and HITRAN, respectively. The SDR-Voigt difference (blue) is moderate in the line center and wings and small in the transition region. In contrast, the SDV-Voigt difference (red) is large only in the line center and drops sharply in the line wings.

Appendix B. NDACC Data Providers

The NDACC data in this publication were obtained from sites listed in Appendix Table A1. The data are publicly available via http://www.ndacc.org.

Table A1. NDACC sites used in this publication along with the stations' representative and cooperating institution.

NDACC Site	Coorperating Institutions
Bremen	Prof Dr Justus Notholt
	Institute of Environmental Physics; University of Bremen, Germany
Zugspitze	Dr Ralf Sussmann
	Institute for Meteorology and Climate Research; Karlsruhe Institute of Technology, Germany
Izana	Dr Emilio Cuevas Agulló
	Izana Atmospheric Research Center; AEMET – Meteorological State Agency, Spain
Jungfraujoch	Dr Emmanuel Mahieua
	University of Liége, Belgium
Kiruna	Dr Uwe Raffalski
	Institute of Space Physics, Sweden
Thule	Dr Niels Larsen
	Danish Climate Center; Danish Meteorological Institute, Denmark
Toronto	Dr Kimberly Strong
	Department of Physics; University of Toronto, Canada

Appendix C. TCCON Data Providers

The TCCON data were obtained from sites listed in Appendix Table A2. The TCCON Data Archive is hosted by CaltechDATA, California Institute of Technology, CA (US), doi: https://doi.org/10.14291/tccon.archive/1348407. The Darwin TCCON site is supported by ARC grants DP160101598, DP140101552, DP110103118 and DP0879468, and NASA grants NAG5-12247 and NNG05-GD07G.

Table A2. Total Carbon Column Observing Network (TCCON) sites with the references to the data used in this publication.

TCCON Site	Reference
Darwin	Griffith et al. [84]
Lauder	Sherlock et al. [85]

References

1. Perrin, A. Review on the existing spectroscopic databases for atmospheric applications. In *Spectroscopy from Space*; NATO Science Series II; Demaison, J., Sarka, K., Cohen, E., Eds.; Springer Netherlands: Dordrecht, The Netherlands, 2001; Volume 20, pp. 235–258.

2. Flaud, J.; Oelhaf, H. Infrared spectroscopy and the terrestrial atmosphere. *C. R. Phys.* **2004**, *5*, 259–271. [CrossRef]

3. Payan, S.; de La Noë, J.; Hauchecorne, A.; Camy-Peyret, C. A review of remote sensing techniques and related spectroscopy problems. *C. R. Phys.* **2005**, *6*, 825–835. [CrossRef]

4. Chance, K. Ultraviolet and visible spectroscopy and spaceborne remote sensing of the Earth's atmosphere. *Comp. R. Phys.* **2006**, *6*, 836–847. [CrossRef]

5. Flaud, J.; Perrin, A.; Picquet-Varrault, B.; Gratien, A.; Orphal, J.; Doussin, J. Quantitative Spectroscopy and Atmospheric Measurements. In *Remote Sensing of the Atmosphere for Environmental Security*; NATO Security through Science Series; Perrin, A., Ben Sari-Zizi, N., Demaison, J., Eds.; Springer: Dordrecht, The Netherlands, 2006; Volume 14, pp. 107–121. [CrossRef]

6. Flaud, J.; Picquet-Varrault, B.; Gratien, A.; Orphal, J.; Doussin, J. Synergistic Use of Different Atmospheric Instruments: What about the Spectral Parameters? In *Proceedings of the First Atmospheric Science Conference*; Lacoste, H., Ed.; ESA: Frascati, Italy, 2006; Volume SP-628.

7. Feng, X.; Zhao, F.S. Effect of changes of the HITRAN database on transmittance calculations in the near-infrared region. *J. Quant. Spectrosc. Radiat. Transf.* **2009**, *110*, 247–255. [CrossRef]

8. Kratz, D.P. The sensitivity of radiative transfer calculations to the changes in the HITRAN database from 1982 to 2004. *J. Quant. Spectrosc. Radiat. Transf.* **2008**, *109*, 1060–1080. [CrossRef]

9. Galli, A.; Butz, A.; Scheepmaker, R.A.; Hasekamp, O.; Landgraf, J.; Tol, P.; Wunch, D.; Deutscher, N.M.; Toon, G.C.; Wennberg, P.O.; et al. CH_4, CO, and H_2O spectroscopy for the Sentinel-5 Precursor mission: An assessment with the Total Carbon Column Observing Network measurements. *Atmos. Meas. Tech.* **2012**, *5*, 1387–1398. doi:10.5194/amt-5-1387-2012. [CrossRef]

10. Scheepmaker, R.A.; Frankenberg, C.; Galli, A.; Butz, A.; Schrijver, H.; Deutscher, N.M.; Wunch, D.; Warneke, T.; Fally, S.; Aben, I. Improved water vapour spectroscopy in the 4174–4300 cm^{-1} region and its impact on SCIAMACHY HDO/H_2O measurements. *Atmos. Meas. Tech.* **2013**, *6*, 879–894. [CrossRef]

11. Checa-García, R.; Landgraf, J.; Galli, A.; Hase, F.; Velazco, V.; Tran, H.; Boudon, V.; Alkemade, F.; Butz, A. Mapping spectroscopic uncertainties into prospective methane retrieval errors from Sentinel-5 and its precursor. *Atmos. Meas. Tech.* **2015**, *8*, 3617–3629. [CrossRef]

12. Gordon, I.; Rothman, L.; Hill, C.; Kochanov, R.; Tan, Y.; Bernath, P.; Birk, M.; Boudon, V.; Campargue, A.; Chance, K.; et al. The HITRAN2016 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **2017**, *203*, 3–69. [CrossRef]

13. Jacquinet-Husson, N.; Armante, R.; Scott, N.; Chédin, A.; Crépeau, L.; Boutammine, C.; Bouhdaoui, A.; Crevoisier, C.; Capelle, V.; Boonne, C.; et al. The 2015 edition of the GEISA spectroscopic database. *J. Mol. Spectrosc.* **2016**, *327*, 31–72. New Visions of Spectroscopic Databases, Volume II. [CrossRef]

14. Pickett, H.; Poynter, R.; Cohen, E.; Delitsky, M.; Pearson, J.; Müller, H. Submillimeter, millimeter, and microwave spectral line catalog. *J. Quant. Spectrosc. Radiat. Transf.* **1998**, *60*, 883–890. [CrossRef]

15. Endres, C.P.; Schlemmer, S.; Schilke, P.; Stutzki, J.; Müller, H.S. The Cologne Database for Molecular Spectroscopy, CDMS, in the Virtual Atomic and Molecular Data Centre, VAMDC. *J. Mol. Spectrosc.* **2016**, *327*, 95–104. New Visions of Spectroscopic Databases, Volume II. [CrossRef]

16. Nikitin, A.; Lyulin, O.; Mikhailenko, S.; Perevalov, V.; Filippov, N.; Grigoriev, I.; Morino, I.; Yokota, T.; Kumazawa, R.; Watanabe, T. GOSAT-2009 methane spectral line list in the 5550–6236 cm^{-1} range. *J. Quant. Spectrosc. Radiat. Transf.* **2010**, *111*, 2211–2224. [CrossRef]

17. Campargue, A.; Leshchishina, O.; Wang, L.; Mondelain, D.; Kassi, S.; Nikitin, A. Refinements of the WKMC empirical line lists (5852–7919 cm^{-1}) for methane between 80 K and 296 K. *J. Quant. Spectrosc. Radiat. Transf.* **2012**, *113*, 1855–1873. [CrossRef]

18. Nikitin, A.; Lyulin, O.; Mikhailenko, S.; Perevalov, V.; Filippov, N.; Grigoriev, I.; Morino, I.; Yoshida, Y.; Matsunaga, T. GOSAT-2014 methane spectral line list. *J. Quant. Spectrosc. Radiat. Transf.* **2015**, *154*, 63–71. [CrossRef]

19. Birk, M.; Wagner, G.; Loos, J.; Mondelain, D.; Campargue, A. ESA SEOM–IAS—Spectroscopic Parameters Database 2.3 µm Region [Data set]. Zenodo. 2017. Available online: https://zenodo.org/record/1009126#.XnyN7q19iXl (accessed on 8 March 2019).

20. Buchwitz, M.; Burrows, J. Retrieval of CH$_4$, CO, and CO$_2$ total column amounts from SCIAMACHY near-infrared nadir spectra: Retrieval algorithm and first results. In Proceedings of the 10th International Symposium Remote Sensing—Remote Sensing of Clouds and the Atmosphere VIII, Barcelona, Spain, 8–12 September 2003.

21. Pan, L.; Gille, J.; Edwards, D.; Bailey, P.; Rodgers, C. Retrieval of tropospheric carbon monoxide for the MOPITT experiment. *J. Geophys. Res.* **1998**, *103*, 32277–32290. [CrossRef]

22. Kuze, A.; Suto, H.; Nakajima, M.; Hamazaki, T. Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.* **2009**, *48*, 6716–6733. [CrossRef]

23. Crisp, D.; Atlas, R.; Breon, F.M.; Brown, L.; Burrows, J.; Ciais, P.; Connor, B.; Doney, S.; Fung, I.; Jacob, D.; et al. The Orbiting Carbon Observatory (OCO) mission. *Adv. Space Res.* **2004**, *34*, 700–709. [CrossRef]

24. Yang, D.; Liu, Y.; Cai, Z.; Chen, X.; Yao, L.; Lu, D. First Global Carbon Dioxide Maps Produced from TanSat Measurements. *Adv. Atmos. Sci.* **2018**, *35*, 621–623. [CrossRef]

25. Landgraf, J.; aan de Brugh, J.; Scheepmaker, R.; Borsdorff, T.; Hu, H.; Houweling, S.; Butz, A.; Aben, I.; Hasekamp, O. Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements. *Atmos. Meas. Tech.* **2016**, *9*, 4955–4975. [CrossRef]

26. Frankenberg, C.; Warneke, T.; Butz, A.; Aben, I.; Hase, F.; Spietz, P.; Brown, L.R. Pressure broadening in the 2v_3 band of methane and its implication on atmospheric retrievals. *Atm. Chem. Phys.* **2008**, *8*, 5061–5075. [CrossRef]

27. Frankenberg, C.; Bergamaschi, P.; Butz, A.; Houweling, S.; Meirink, J.F.; Notholt, J.; Petersen, A.K.; Schrijver, H.; Warneke, T.; Aben, I. Tropical methane emissions: A revised view from SCIAMACHY onboard ENVISAT. *Geophys. Res. Letters* **2008**, *35*, L15811. [CrossRef]

28. Gloudemans, A.; Schrijver, H.; Hasekamp, O.; Aben, I. Error analysis for CO and CH$_4$ total column retrievals from SCIAMACHY 2.3 µm spectra. *Atm. Chem. Phys.* **2008**, *8*, 3999–4017. [CrossRef]

29. Oyafuso, F.; Payne, V.; Drouin, B.; Devi, V.; Benner, D.; Sung, K.; Yu, S.; Gordon, I.; Kochanov, R.; Tan, Y.; et al. High accuracy absorption coefficients for the Orbiting Carbon Observatory-2 (OCO-2) mission: Validation of updated carbon dioxide cross-sections using atmospheric spectra. *J. Quant. Spectrosc. Radiat. Transf.* **2017**, *203*, 213–223. HITRAN2016 Special Issue. [CrossRef]

30. Borsdorff, T.; aan de Brugh, J.; Schneider, A.; Lorente, A.; Birk, M.; Wagner, G.; Kivi, R.; Hase, F.; Feist, D.G.; Sussmann, R.; et al. Improving the TROPOMI CO data product: update of the spectroscopic database and destriping of single orbits. *Atmos. Meas. Tech.* **2019**, *12*, 5443–5455. [CrossRef]

31. Buchwitz, M.; de Beek, R.; Bramstedt, K.; Noël, S.; Bovensmann, H.; Burrows, J.P. Global carbon monoxide as retrieved from SCIAMACHY by WFM-DOAS. *Atm. Chem. Phys.* **2004**, *4*, 1945–1960. [CrossRef]

32. Buchwitz, M.; de Beek, R.; Noël, S.; Burrows, J.P.; Bovensmann, H.; Bremer, H.; Bergamaschi, P.; Körner, S.; Heimann, M. Carbon monoxide, methane and carbon dioxide columns retrieved from SCIAMACHY by WFM-DOAS: year 2003 initial data set. *Atm. Chem. Phys.* **2005**, *5*, 3313–3329. [CrossRef]

33. Frankenberg, C.; Platt, U.; Wagner, T. Retrieval of CO from SCIAMACHY onboard ENVISAT: detection of strongly polluted areas and seasonal patterns in global CO abundances. *Atm. Chem. Phys.* **2005**, *5*, 1639–1644. [CrossRef]

34. Gloudemans, A.; Schrijver, H.; Kleipool, Q.; van den Broek, M.; Straume, A.; Lichtenberg, G.; van Hees, R.; Aben, I.; Meirink, J. The impact of SCIAMACHY near-infrared instrument calibration on CH_4 and CO total columns. *Atm. Chem. Phys.* **2005**, *5*, 2369–2383. [CrossRef]

35. Borsdorff, T.; Tol, P.; Williams, J.E.; de Laat, J.; aan de Brugh, J.; Nédélec, P.; Aben, I.; Landgraf, J. Carbon monoxide total columns from SCIAMACHY 2.3 μm atmospheric reflectance measurements: Towards a full-mission data product (2003–2012). *Atmos. Meas. Tech.* **2016**, *9*, 227–248. [CrossRef]

36. Borsdorff, T.; aan de Brugh, J.; Hu, H.; Nédélec, P.; Aben, I.; Landgraf, J. Carbon monoxide column retrieval for clear-sky and cloudy atmospheres: A full-mission data set from SCIAMACHY 2.3μm reflectance measurements. *Atmos. Meas. Tech.* **2017**, *10*, 1769–1782. [CrossRef]

37. Gimeno García, S.; Schreier, F.; Lichtenberg, G.; Slijkhuis, S. Near infrared nadir retrieval of vertical column densities: methodology and application to SCIAMACHY. *Atmos. Meas. Tech.* **2011**, *4*, 2633–2657. [CrossRef]

38. de Laat, A.; Gloudemans, A.; Schrijver, H.; van den Broek, M.; Meirink, J.; Aben, I.; Krol, M. Quantitative analysis of SCIAMACHY carbon monoxide total column measurements. *Geophys. Res. Lett.* **2006**, *33*, L07807. [CrossRef]

39. Lichtenberg, G.; Gimeno García, S.; Schreier, F.; Slijkhuis, S.; Snel, R.; Bovensmann, H. Impact of Level 1 Quality on SCIAMACHY Level 2 Retrieval. In Proceedings of the 38th COSPAR Scientific Assembly, Bremen, Germany, 18–25 July 2010.

40. Hochstaffl, P.; Schreier, F.; Lichtenberg, G.; Gimeno García, S. Validation of Carbon Monoxide Total Column Retrievals from SCIAMACHY Observations with NDACC/TCCON Ground-Based Measurements. *Remote Sens.* **2018**, *10*, 223. [CrossRef]

41. Goody, R.; Yung, Y. *Atmospheric Radiation—Theoretical Basis*, 2nd ed.; Oxford University Press: Oxford, UK, 1989.

42. Liou, K.N. *An Introduction to Atmospheric Radiation*, 2nd ed.; Academic Press: Cambridge, MA, USA, 2002; p. 583.

43. Zdunkowski, W.; Trautmann, T.; Bott, A. *Radiation in the Atmosphere—A Course in Theoretical Meteorology*; Cambridge University Press: Cambridge, UK, 2007.

44. Armstrong, B. Spectrum Line Profiles: The Voigt Function. *J. Quant. Spectrosc. Radiat. Transf.* **1967**, *7*, 61–88. [CrossRef]

45. Humlíček, J. Optimized computation of the Voigt and complex probability function. *J. Quant. Spectrosc. Radiat. Transf.* **1982**, *27*, 437–444. [CrossRef]

46. Weideman, J. Computation of the Complex Error Function. *SIAM J. Num. Anal.* **1994**, *31*, 1497–1518. [CrossRef]

47. Schreier, F. Optimized Implementations of Rational Approximations for the Voigt and Complex Error Function. *J. Quant. Spectrosc. Radiat. Transf.* **2011**, *112*, 1010–1025. [CrossRef]

48. Tennyson, J.; Bernath, P.; Campargue, A.; Császár, A.; Daumont, L.; Gamache, R.; Hodges, J.; Lisak, D.; Naumenko, O.; Rothman, L.; et al. Recommended isolated-line profile for representing high-resolution spectroscopic transitions (IUPAC Technical Report). *Pure Appl. Chem.* **2014**, *86*, 1931–1943. [CrossRef]

49. Berman, P. Speed-dependent collisional width and shift parameters in spectral profiles. *J. Quant. Spectrosc. Radiat. Transf.* **1972**, *12*, 1321–1342. [CrossRef]

50. Boone, C.; Walker, K.; Bernath, P. Speed–dependent Voigt profile for water vapor in infrared remote sensing applications. *J. Quant. Spectrosc. Radiat. Transf.* **2007**, *105*, 525–532. [CrossRef]

51. Boone, C.; Walker, K.; Bernath, P. An efficient analytical approach for calculating line mixing in atmospheric remote sensing applications. *J. Quant. Spectrosc. Radiat. Transf.* **2011**, *112*, 980–989. [CrossRef]

52. Kochanov, V.P. Speed-dependent spectral line profile including line narrowing and mixing. *J. Quant. Spectrosc. Radiat. Transf.* **2016**, *177*, 261–268. [CrossRef]

53. Varghese, P.; Hanson, R. Collisional narrowing effects on spectral line shapes measured at high resolution. *Appl. Opt.* **1984**, *23*, 2376–2385. [CrossRef]

54. Rosenkranz, P. Shape of the 5 mm oxygen band in the atmosphere. *IEEE Trans. Antennas Propag.* **1975**, *23*, 498–506. [CrossRef]

55. Strow, L.L.; Tobin, D.C.; Hannon, S.E. A compilation of first-order line-mixing coefficients for CO$_2$ Q-branches. *J. Quant. Spectrosc. Radiat. Transf.* **1994**, *52*, 281–294. [CrossRef]

56. Dicke, R.H. The Effect of Collisions upon the Doppler Width of Spectral Lines. *Phys. Rev.* **1953**, *89*, 472–473. [CrossRef]

57. Lévy, A.; Lacome, N.; Chackerian, C., Jr. Collisional Line Mixing. In *Spectroscopy of the Earth's Atmosphere and Interstellar Medium*; Rao, K., Weber, A., Eds.; Academic Press: Cambridge, MA, USA, 1992; pp. 261–338.

58. Ngo, N.; Lisak, D.; Tran, H.; Hartmann, J.M. An isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes. *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *129*, 89–100. Erratum: JQSRT 134, 105 (2014). [CrossRef]

59. Tran, H.; Ngo, N.; Hartmann, J.M. Efficient computation of some speed-dependent isolated line profiles. *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *129*, 199–203. Erratum: JQSRT 134, 104 (2014). [CrossRef]

60. Malathy Devi, V.; Benner, D.C.; Smith, M.; Mantz, A.; Sung, K.; Brown, L.; Predoi-Cross, A. Spectral line parameters including temperature dependences of self- and air-broadening in the 2-0 band of CO at 2.3 µm. *J. Quant. Spectrosc. Radiat. Transf.* **2012**, *113*, 1013–1033. [CrossRef]

61. Rothman, L.; Gordon, I.; Babikov, Y.; Barbe, A.; Benner, D.C.; Bernath, P.; Birk, M.; Bizzocchi, L.; Boudon, V.; Brown, L.; et al. The HITRAN2012 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **2013**, *130*, 4–50. [CrossRef]

62. Li, G.; Gordon, I.E.; Rothman, L.S.; Tan, Y.; Hu, S.M.; Kassi, S.; Campargue, A.; Medvedev, E.S. Rovibrational line lists for nine isotopologues of the CO molecule in the X$^1\Sigma^+$ ground electronic state. *Astrophys. J. Supp. S.* **2015**, *216*, 15. [CrossRef]

63. Mlawer, E.; Payne, V.; Moncet, J.L.; Delamere, J.; Alvarado, M.; Tobin, D. Development and recent evaluation of the MT-CKD model of continuum absorption. *Philos. Trans. Roy. Soc. Lond. Ser. A* **2012**, *370*, 2520–2556. [CrossRef] [PubMed]

64. Jacquinet-Husson, N.; Crepeau, L.; Armante, R.; Boutammine, C.; Chedin, A.; Scott, N.; Crevoisier, C.; Capelle, V.; Boone, C.; Poulet-Crovisier, N.; et al. The 2009 edition of the GEISA spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **2011**, *112*, 2395–2445. [CrossRef]

65. Kistler, R.; Collins, W.; Saha, S.; White, G.; Woollen, J.; Kalnay, E.; Chelliah, M.; Ebisuzaki, W.; Kanamitsu, M.; Kousky, V.; et al. The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bull. Am. Met. Soc.* **2001**, *82*, 247–267. [CrossRef]

66. Anderson, G.; Clough, S.; Kneizys, F.; Chetwynd, J.; Shettle, E. *AFGL Atmospheric Constituent Profiles (0–120 km)*; Technical Report TR-86-0110; Air Force Geophysics Laboratory (AFGL): Hanscom AFB, MA, USA, 1986.

67. Schreier, F.; Gimeno García, S.; Hedelt, P.; Hess, M.; Mendrok, J.; Vasquez, M.; Xu, J. GARLIC—A General Purpose Atmospheric Radiative Transfer Line-by-Line Infrared-Microwave Code: Implementation and Evaluation. *J. Quant. Spectrosc. Radiat. Transf.* **2014**, *137*, 29–50. [CrossRef]

68. Hamidouche, M.; Lichtenberg, G. In-Flight Retrieval of SCIAMACHY Instrument Spectral Response Function. *Remote Sens.* **2018**, *10*, 401. [CrossRef]

69. Loyola, D. A new cloud recognition algorithm for optical sensors. In Proceedings of the IGARSS '98. Sensing and Managing the Environment. 1998 IEEE International Geoscience and Remote Sensing. Symposium Proceedings, Seattle, WA, USA, 6–10 July 1998; pp. 572–574. [CrossRef]

70. Kokhanovsky, A.A.; von Hoyningen-Huene, W.; Rozanov, V.V.; Noël, S.; Gerilowski, K.; Bovensmann, H.; Bramstedt, K.; Buchwitz, M.; Burrows, J.P. The semianalytical cloud retrieval algorithm for SCIAMACHY II. The application to MERIS and SCIAMACHY data. *Atm. Chem. Phys.* **2006**, *6*, 4129–4136. [CrossRef]

71. Hodges, J.L. The significance probability of the Smirnov two-sample test. *Ark. Mat.* **1958**, *3*, 469–486. [CrossRef]

72. Frankenberg, C.; Meirink, J.F.; Bergamaschi, P.; Goede, A.P.H.; Heimann, M.; Körner, S.; Platt, U.; van Weele, M.; Wagner, T. Satellite chartography of atmospheric methane from SCIAMACHY on board ENVISAT: Analysis of the years 2003 and 2004. *J. Geophys. Res.* **2006**, *111*, D07303. [CrossRef]

73. Hochstaffl, P.; Gimeno García, S.; Schreier, F.; Hamidouche, M.; Lichtenberg, G. Validation of Carbon Monoxide Vertical Column Densities Retrieved from SCIAMACHY Infrared Nadir Observations. In Proceedings of the Living Planet Symposium, Prague, Czech Republic, 9–13 May 2016.

74. Wunch, D.; Toon, G.C.; Sherlock, V.; Deutscher, N.M.; Liu, C.; Feist, D.G.; Wennberg, P.O. Documentation for the 2014 TCCON Data Release (Version GGG2014.R0). *CaltechDATA* **2015**. [CrossRef]

75. Sussmann, R.; Ostler, A.; Forster, F.; Rettinger, M.; Deutscher, N.M.; Griffith, D.W.T.; Hannigan, J.W.; Jones, N.; Patra, P.K. First intercalibration of column-averaged methane from the Total Carbon Column Observing Network and the Network for the Detection of Atmospheric Composition Change. *Atmos. Meas. Tech.* **2013**, *6*, 397–418. [CrossRef]

76. Kiel, M.; Hase, F.; Blumenstock, T.; Kirner, O. Comparison of XCO abundances from the Total Carbon Column Observing Network and the Network for the Detection of Atmospheric Composition Change measured in Karlsruhe. *Atmos. Meas. Tech.* **2016**, *9*, 2223–2239. [CrossRef]

77. Zhou, M.; Langerock, B.; Vigouroux, C.; Sha, M.K.; Hermans, C.; Metzger, J.M.; Chen, H.; Ramonet, M.; Kivi, R.; Heikkinen, P.; et al. TCCON and NDACC XCO measurements: difference, discussion and application. *Atmos. Meas. Tech.* **2019**, *12*, 5979–5995. [CrossRef]

78. Sussmann, R.; Buchwitz, M. Initial validation of ENVISAT/SCIAMACHY columnar CO by FTIR profile retrievals at the Ground-Truthing Station Zugspitze. *Atm. Chem. Phys.* **2005**, *5*, 1497–1503. [CrossRef]

79. Sussmann, R.; Stremme, W.; Buchwitz, M.; de Beek, R. Validation of ENVISAT/SCIAMACHY columnar methane by solar FTIR spectrometry at the Ground-Truthing Station Zugspitze. *Atm. Chem. Phys.* **2005**, *5*, 2419–2429. [CrossRef]

80. Buchholz, R.; Deeter, M.; Worden, H.; Gille, J.; Edwards, D.; Hannigan, J.; Jones, N.; Paton-Walsh, C.; Griffith, D.; Smale, D.; et al. Validation of MOPITT carbon monoxide using ground-based Fourier transform infrared spectrometer data from NDACC. *Atmos. Meas. Tech.* **2017**, *10*, 1927–1956. [CrossRef]

81. Rodgers, C.; Connor, B. Intercomparison of remote sounding instruments. *J. Geophys. Res.* **2003**, *108*, 4116. [CrossRef]

82. Coldewey-Egbers, M.; Loyola, D.G.; Koukouli, M.; Balis, D.; Lambert, J.C.; Verhoelst, T.; Granville, J.; van Roozendael, M.; Lerot, C.; Spurr, R.; et al. The GOME-type Total Ozone Essential Climate Variable (GTO-ECV) data record from the ESA Climate Change Initiative. *Atmos. Meas. Tech.* **2015**, *8*, 3923–3940. [CrossRef]

83. Rothman, L.; Jacquemart, D.; Barbe, A.; Benner, D.C.; Birk, M.; Brown, L.; Carleer, M.; Chackerian, C., Jr.; Chance, K.; Coudert, L.; et al. The HITRAN 2004 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **2005**, *96*, 139–204. [CrossRef]

84. Griffith, D.W.T.; Deutscher, N.M.; Velazco, V.A.; Wennberg, P.O.; Yavin, Y.; Keppel-Aleks, G.; Washenfelder, R.; Toon, G.C.; Blavier, J.F.; Paton-Walsh, C.; et al. TCCON data from Darwin (AU), Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA. 2014. Available online: https://data.caltech.edu/records/269l (accessed on 8 March 2019).

85. Sherlock, V.; Connor, B.; Robinson, J.; Shiona, H.; Smale, D.; Pollard, D. TCCON data from Lauder (NZ), 120HR, Release GGG2014.R0. TCCON Data Archive, hosted by CaltechDATA. 2014. Available online: https://data.caltech.edu/records/281l (accessed on 8 March 2019).

Publication III

remote sensing

Article

Impact of Molecular Spectroscopy on Carbon Monoxide Abundances from TROPOMI

Philipp Hochstaffl [1,*], **Franz Schreier [1]**, **Manfred Birk [1]**, **Georg Wagner [1]**,
Dietrich G. Feist [2,3], **Justus Notholt [4]**, **Ralf Sussmann [5]** and **Yao Té [6]**

[1] Deutsches Zentrum für Luft- und Raumfahrt, Institut für Methodik der Fernerkundung,
82234 Oberpfaffenhofen, Germany; franz.schreier@dlr.de (F.S.); manfred.birk@dlr.de (M.B.);
georg.wagner@dlr.de (G.W.)

[2] Ludwig-Maximilians-Universität München, Lehrstuhl für Physik der Atmosphäre, 80333 Munich, Germany;
d.feist@lmu.de

[3] Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre,
82234 Oberpfaffenhofen, Germany

[4] Institute of Environmental Physics (IUP), University of Bremen, 28359 Bremen, Germany;
jnotholt@iup.physik.uni-bremen.de

[5] Karlsruhe Institute of Technology, IMK-IFU, 82467 Garmisch-Partenkirchen, Germany;
ralf.sussmann@kit.edu

[6] Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique et Atmosphères (LERMA-IPSL),
Sorbonne Université, CNRS, Observatoire de Paris, PSL Université, 75005 Paris, France;
yao-veng.te@sorbonne-universite.fr

* Correspondence: philipp.hochstaffl@dlr.de

check for updates

Abstract: The impact of SEOM–IAS (Scientific Exploitation of Operational Missions–Improved Atmospheric Spectroscopy) spectroscopic information on CO columns from TROPOMI (Tropospheric Monitoring Instrument) shortwave infrared (SWIR) observations was examined. HITRAN 2016 (High Resolution Transmission) and GEISA 2015 (Gestion et Etude des Informations Spectroscopiques Atmosphériques 2015) were used as a reference upon which the spectral fitting residuals, retrieval errors and inferred quantities were assessed. It was found that SEOM–IAS significantly improves the quality of the CO retrieval by reducing the residuals to TROPOMI observations. The magnitude of the impact is dependent on the climatological region and spectroscopic reference used. The difference in the CO columns was found to be rather small, although discrepancies reveal, for selected scenes, in particular, for observations with elevated molecular concentrations. A brief comparison to Total Column Carbon Observing Network (TCCON) and Network for the Detection of Atmospheric Composition Change (NDACC) also demonstrated that both spectroscopies cause similar columns; however, the smaller retrieval errors in the SEOM with Speed-Dependent Rautian and line-Mixing (SDRM) inferred CO turned out to be beneficial in the comparison of post-processed mole fractions with ground-based references.

1. Introduction

Many species present in the atmosphere influence Earth's radiative transfer by absorbing, emitting, and scattering electromagnetic energy at certain wavelengths [1]. The interaction of radiation with matter makes molecular spectroscopy a powerful tool in investigating the composition, distribution, and evolution of atmospheric constituents. Key elements in the global monitoring of relevant

molecules are state-of-the-art passive remote sensors on satellites, which provide valuable spectroscopic measurements for many tropospheric and stratospheric constituents in near real time [2,3].

1.1. TROPOMI aboard S5P

The Sentinel-5 Precursor (S5P) is the first satellite mission within the European Union's (EU) Earth observation program Copernicus that is dedicated to the monitoring of atmospheric chemistry. Successfully launched on 13 October 2017, the satellite was put into a circular near-polar sun-synchronous low Earth orbit of 824 km with an ascending node equatorial crossing at 13:30 h mean local solar time [4]. S5P's payload is the Tropospheric Monitoring Instrument (TROPOMI) was jointly developed by the Netherlands and the European Space Agency (ESA) and is building upon the heritage of its predecessors GOME (Global Ozone Monitoring Experiment; Burrows et al. [5]), GOME-2 (Munro et al. [6]), SCIAMACHY (SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY Bovensmann et al. [7], Gottwald and Bovensmann [8]), and OMI (Ozone Monitoring Instrument Levelt et al. [9]).

The TROPOMI instrument consists of a grating spectrometer with spectral bands in the ultraviolet, visible (UVIS, 270–500 nm), near infrared (NIR, 675–775 nm), and shortwave infrared (SWIR, 2305–2385 nm) spectral range and observes sunlight reflected by the Earth's atmosphere and surface with a spectral resolution of 0.25 nm and a spectral sampling interval of 0.1 nm in the SWIR bands [10]. It operates in a push-broom configuration with a $108°$ field of view in the across-track direction and provides almost daily global coverage at a spatial resolution of about 7×7 km in the sub satellite point in bands 7 and 8 [4].

1.2. Carbon Monoxide

The absorption of infrared radiation by carbon monoxide (CO) allows space-borne sensors to record its infrared spectrum [11–15]. Although carbon monoxide (CO) is not a greenhouse gas, such as water (H_2O), methane (CH_4), or carbon dioxide (CO_2), it plays a role in climate change by influencing the tropospheric hydroxyl radical (OH) and, thereby, the CH_4 and CO_2 concentrations [16–18]. CO's main sources are combustion in oxygen-poor conditions, as well as atmospheric oxidation of CH_4 and other hydrocarbons. Oxidation by the OH radical is CO's primary loss mechanism, which essentially determines its one month lifetime in the troposphere [19–21]. Moreover, CO is a major atmospheric pollutant that can affect human health, and it indirectly plays a role in the catalytic ozone production and destruction.

1.3. Absorption in the SWIR and Retrieval

In the SWIR, the fraction of photons generated by thermal emission is well below 1% and can hence be neglected [22]. Furthermore, clear sky measurements are subject to little atmospheric scattering, and most of the measured light is reflected by the Earth's surface. Those characteristics and TROPOMI's high signal-to-noise ratio (SNR) in the 2.3 μm region over land surfaces allows the instrument to measure clear-sky CO columns with high sensitivity throughout the vertical extent of the atmosphere [23–29].

Our retrieval relies on measurements of the first overtone 2-0 absorption band of CO situated between 3780–4355 cm^{-1} from which the vertically integrated CO column density is inferred. Since the absorption of CO is weak in that spectral region (Figure 1), no vertical profile of CO can be retrieved in the selected fitting window between 4277.20–4302.90 cm^{-1}. Furthermore, the least squares fit of the forward model parameters to the measurement depends on initial guess data in order to get an appropriate linearization point for the unconstrained fit.
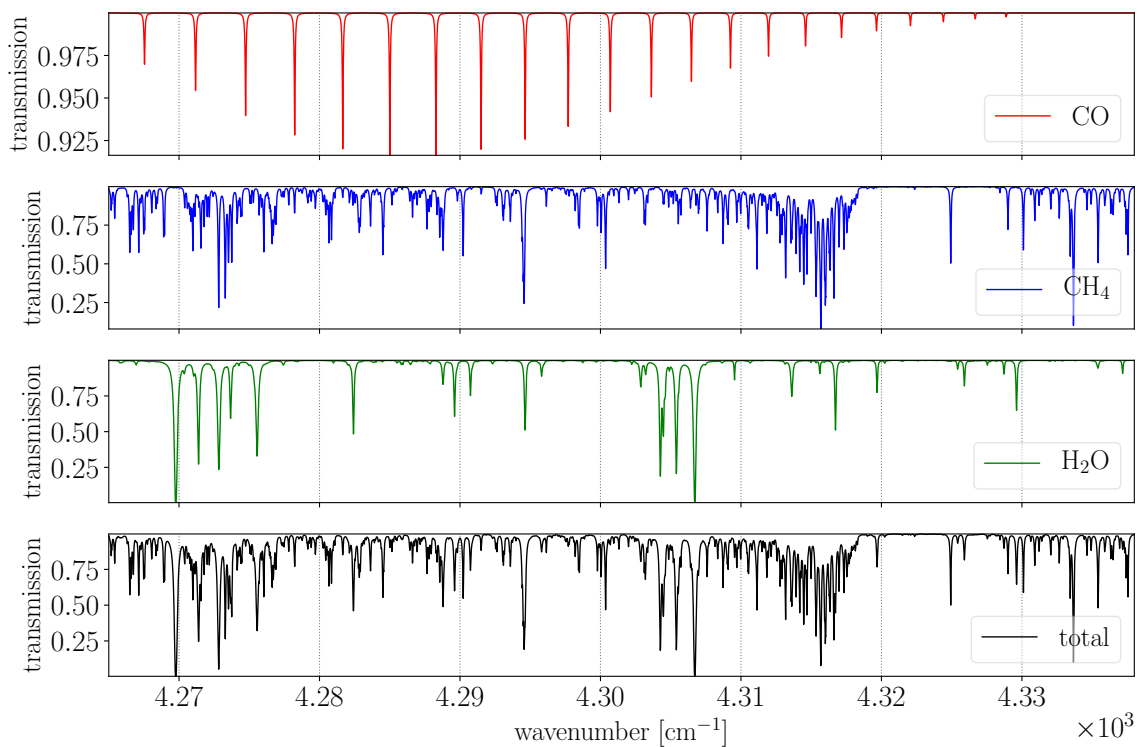
Figure 1. Spectral transmittance of CO, CH$_4$ and H$_2$O (upper three panels) according to Beer-Lambert's law in spectral range of TROPOMI's band 7 for a single path through the US-Standard atmosphere using Scientific Exploitation of Operational Missions (SEOM) line data. The total transmission for the three molecules is depicted in the lower panel. The Python tool Py4CAtS (Python for Computational ATmospheric Spectroscopy [30]) was used to calculate the absorption cross sections of the individual molecules. Note that CO is only responsible for ≈1% of the total optical thickness τ in that spectral range.

In order to infer the amount of atmospheric constituents from an observed spectrum, an accurate description of molecular absorption at different pressure p and temperature T levels is mandatory. In high resolution line–by–line (lbl) models, the cross section of a molecule k_m is calculated by the superposition of many lines l, where each line is the product of a temperature dependent line strength $S_l^{(m)}$ and a normalized line shape function $\int_{-\infty}^{+\infty} g(\nu)\, d\nu = 1$ that is describing mechanisms, such as pressure and Doppler broadening [31]. Therefore, the best possible knowledge of the spectral parameters, such as line position $\hat{\nu}$, line intensity S, line width γ (air- and self-broadening), temperature exponent n, lower-state energy E, and their variation with T and p, is required. However, advances in high resolution absorption spectroscopy and the advent of sensors, such as TROPOMI, with wide spectral ranges at rather high spectral resolutions and excellent SNR ratios have indicated discrepancies between spectroscopic models and observations [32–37]. It was found that physical processes beyond broadening mechanisms described by the Voigt function should be taken into account for accurate atmospheric characterization. Moreover, studies [24,38–42] also indicate that improved molecular spectroscopy is crucial to eliminate systematic residuals in atmospheric spectra and that trace gas retrievals in the SWIR will benefit accordingly.

1.4. Spectroscopic Line Data and Line Profiles

The SEOM–IAS (SEOM, Scientific Exploitation of Operational Missions–Improved Atmospheric Spectroscopy) is an improved line parameter database of H$_2$O, CH$_4$, and CO (available on Zenodo [43,44]) compiled within the framework of an ESA project according to the needs of

the TROPOMI instrument. Fourier transform spectrometer (FTS) and continuous wave cavity ring-down spectroscopy (CRDS) measurements (performed at the German Aerospace Center (DLR) and at Université Grenoble Alpes, respectively) were analyzed in the 4190–4340 cm^{-1} spectral range. The spectroscopic database was obtained from high resolution FTS measurements employing a multireflection cell with absorption path lengths from 14.4–168 m and a temperature range 198–361 K (12 pure and 32 air-broadened CH$_4$ measurements, 1 pure and 4 air-broadened CO measurements, 7 pure and 23 air-broadened H$_2$O measurements, and 4 pure HDO measurements). A multispectrum fitting software [45] developed at DLR was used for the analysis of the measured spectra. For modeling of absorption lines in the multispectrum fitting, a quadratic speed-dependent hard collision model based on the implementation of the Hartmann-Tran (HT) profile was used [35,46–49]. In order to account for line-mixing, the profile was extended using the first and second order perturbation approximation by Rosenkranz [50] and Smith [51]. The CRDS measurements served as validation.

The HT profile with vanishing correlation ($\eta = 0$) reduces to the speed-dependent Rautian (SDR [52–57]) profile. The transmissions in Figure 1 were computed with line-mixing included, i.e., the SEOM with Speed-Dependent Rautian and line-Mixing (SDRM) profile. Figure 2 shows a close up view of the molecular cross sections for CO, CH$_4$, and H$_2$O near 4295 cm^{-1}, where all three molecules possess a fairly strong and almost co-located transition. For each molecule, two cross sections were calculated, one with the SDRM profile (including the extended set of line data) and another with the classic Voigt model, while both use SEOM line data as input. The differences between the two turned out to be around two orders of magnitude smaller than the actual cross sections itself with the maximum disagreement located close to the line center positions.
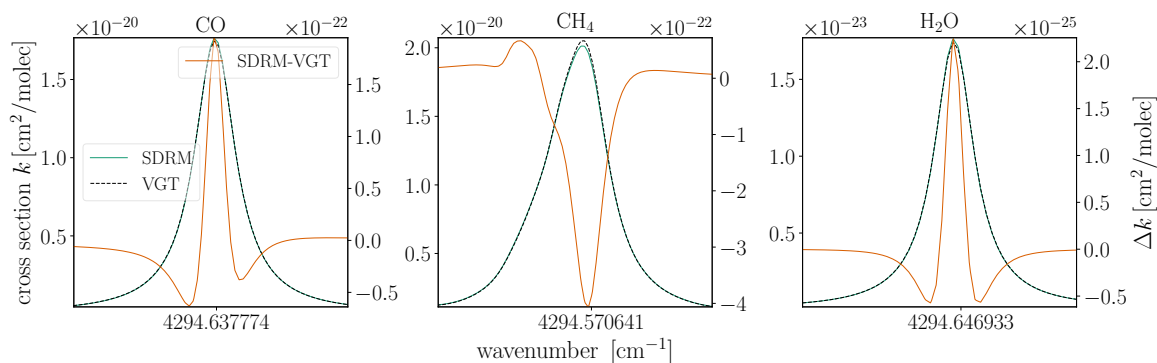

Figure 2. SDRM and Voigt (VGT) molecular cross sections for CO (left), CH$_4$ (center), and H$_2$O (right) using SEOM data. In addition, the difference between the two line profiles (SDRM-VGT) is depicted.

1.5. Previous Studies

An initial validation of SEOM tested the line list with atmospheric spectra from solar occultation measurements, and the new database was found to be a significant improvement over High Resolution Transmission (HITRAN) 2016 (M. Birk, personal communication).

An assessment of the operational TROPOMI CO product for various spectroscopic inputs, including SEOM, was recently published by Borsdorff et al. [58]; however, only the spectroscopic data was substituted, while the remaining retrieval settings were identical to the ones of the operational processing. The study quantified the quality of the spectral fits and biases in the CO column and found that "updating the CH$_4$ cross sections is the main reason for the improved CO product". They concluded that molecular spectroscopy data plays a key role for the quality of the retrieval.

In a recent study [59], we examined the impact of SEOM spectroscopic information on CO total columns from SCIAMACHY and found that the best retrieval results for SEOM line data are obtained when higher-order 'beyond Voigt' effects in molecular absorption are taken into account. The outcome indicates that, strictly speaking, the classic Voigt profile is not an adequate line model for the SEOM line list (confirmed by M. Birk, personal communication), although the largest impact on the improved

fitting residuals ($\approx$3% on average, up to 15% for individual observations) is attributed to the updated line parameters while the line profile has less contribution. Although the CO mole fractions increased by $4 - 11$% using SEOM spectroscopy the difference to H16 was found to be significant for only a limited set of spectra (see Reference [59, Figure 7 and A4]).

2. Methodology

In this study, we investigated the impact of the SEOM spectroscopy on the retrieved CO from TROPOMI SWIR observations by comparing the spectral fitting residuals, deduced columns, and corresponding retrieval errors with the most recent releases of HITRAN 2016 (H16, High Resolution Transmission; [60]) and GEISA 2015 (G15, Gestion et Etude des Informations Spectroscopiques Atmosphériques 2015 [61]). Besides the target gas CO, the impact on the interfering, and hence co-retrieved species CH_4 and H_2O (including their isotopologues), was examined. Note that the reason to stay with the current version of GEISA (2015 instead of 2019) was that not all molecules required for the retrieval of CO in the specified spectral range were updated at the time of submission (R. Armante, personal communication).

2.1. Retrieval Setup

The retrievals in this study were performed with the latest version of the scientific retrieval algorithm BIRRA (Beer Infrared Retrieval Algorithm [25,62]) which has been developed at the German Aerospace Center (DLR) since about 2005. In addition to enhancements in the GARLIC (Generic Atmospheric Radiation Line-by-line Infrared Code [63,64]) forward model described in [59], this most recent version of BIRRA incorporates TROPOMI calibration key data (CKD), such as tabulated instrument spectral response functions (ISRF).

2.2. Input Data

An updated framework is providing auxiliary data for the prototype retrieval of CO abundances from TROPOMI. Since information on the amount of CO is inferred by optimally varying forward model parameters during the inversion process, the quality of the input data affects the accuracy of the retrieval.

2.2.1. Calibrated Level 1b Spectra

The TROPOMI level 1b data (version 1.0) from band 7 contains spectrally and radiometrically calibrated Earth radiance and solar irradiance spectra in the 2305–2345 nm ($\approx$ 4338–4265 cm^{-1}) spectral range (see Figure 3). These quantities already include corrections from the CKD that account for several effects, such as offset, dark-current, pixel-quality, non-linear response, and noise, and were derived during the on-ground calibration campaign prior to launch [65,66]. In-flight, the CKD of, e.g., the pixel-quality, the ISRF, and stray-light correction, is monitored by TROPOMI's calibration unit and updated over the lifetime of the instrument if necessary. This is crucial as the operational level 0-1b processor marks data with quality assessment flags, e.g., in order to exclude bad and dead pixels that are deemed unusable for generation of the level 2 product. The actual number of available pixels in the selected retrieval window between 4277.20–4302.90 cm^{-1} is dependent on the bad and dead pixel-mask (BDPM) and ranges from 146–154 for observations considered in this study. Note that TROPOMI is commanded to perform a solar irradiance measurement near the day-night terminator at the northern side of the orbit only every 15 orbits, i.e., approximately once every calendar day [10].
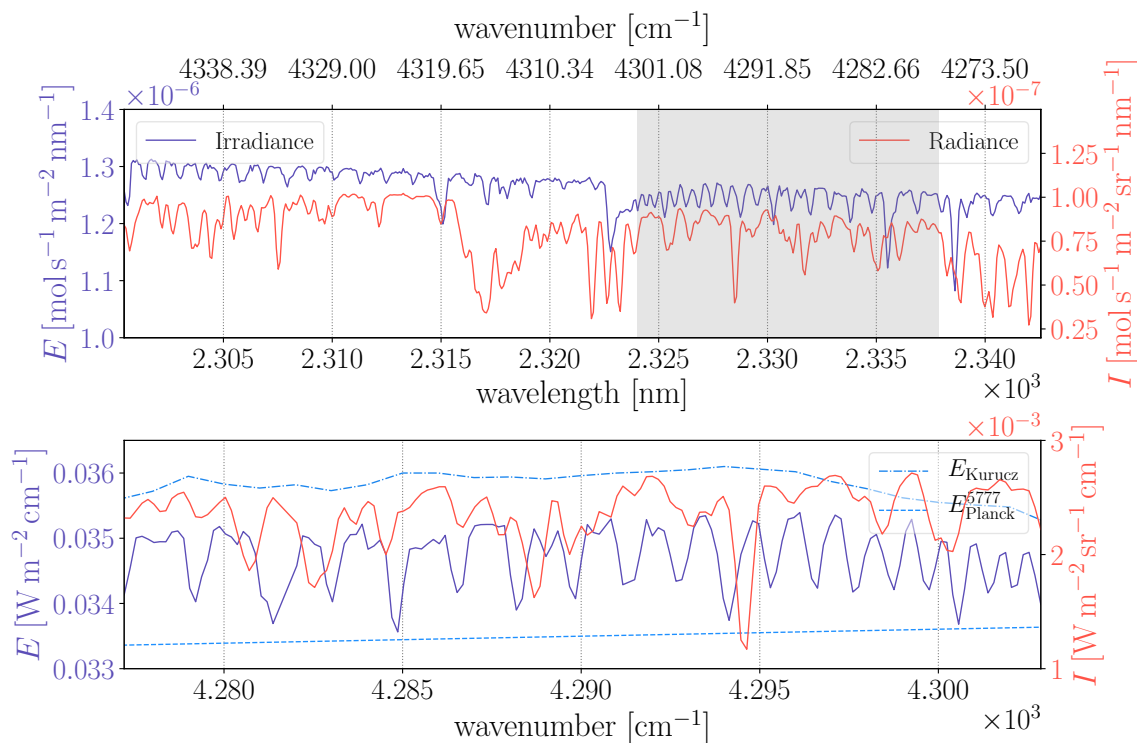
Figure 3. A calibrated level 1b photon radiance and irradiance spectrum from TROPOMI's band 7 with SI (Système international d'unités) units $\mathrm{mol\,s^{-1}\,m^{-2}\,nm^{-1}}$ and the retrieval window indicated in gray (top). The Earth radiance I was obtained in orbit 9093 at around $60°$ northern latitude, while the solar irradiance E was measured in orbit 9088. The lower panel depicts the retrieval window from 4277.20–4302.90 $\mathrm{cm^{-1}}$ with units converted to $\mathrm{W\,m^{-2}\,nm^{-1}}$. E_{Kurucz} and $E_{\mathrm{Planck}}^{5777}$ describe the incident solar irradiance at TOA (top of atmosphere) according to Kurucz [67] and an equivalent black body radiator at 5777 K, respectively. Both were added for illustrative purpose only.

2.2.2. The Instrument's Spectral Response

The forward model needs to include an accurate description of the ISRF $\mathcal{S}$ in order to model the physics of the measurement with adequate accuracy. In the SWIR, the TROPOMI ISRFs vary across the spectral and spatial dimension $\mathcal{S}_{ij}$ and are provided for each of the $1 \leq i \leq 256$ ground pixels, as well as for $1 \leq j \leq 24$ equally spaced central wavelengths of the spectral axis ranging from 2298 to 2344 nm [68]. Eight tabulated ISRFs remain within the range of sufficiently strong CO absorption lines defining our fitting window; hence, interpolated response values were used for most spectral pixels. The rather smooth variation in the spectral dimension is beneficial for the interpolation of responses to pixels where no tabulated values are available (see Figures 4 and 5). Nonetheless, accounting for those variations in the instrument's response is important, particularly when testing spectroscopic data and models.

During on-ground calibration, van Hees et al. [68] found that the accuracy of the ISRF CKD is well within the requirements for trace-gas retrievals. Moreover, van Kempen et al. [66] found that the differences between in-flight and on-ground CKD measurements are small, and no corrections need to be applied.
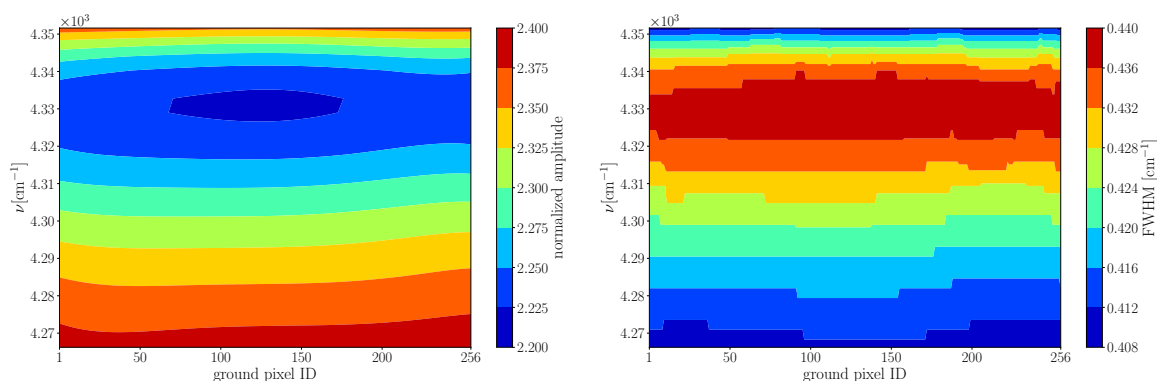
Figure 4. Distribution of the maximum normalized amplitudes (left) and corresponding full-width at half maxima (FWHM) (right) for the 24 × 256 tabulated responses [69].
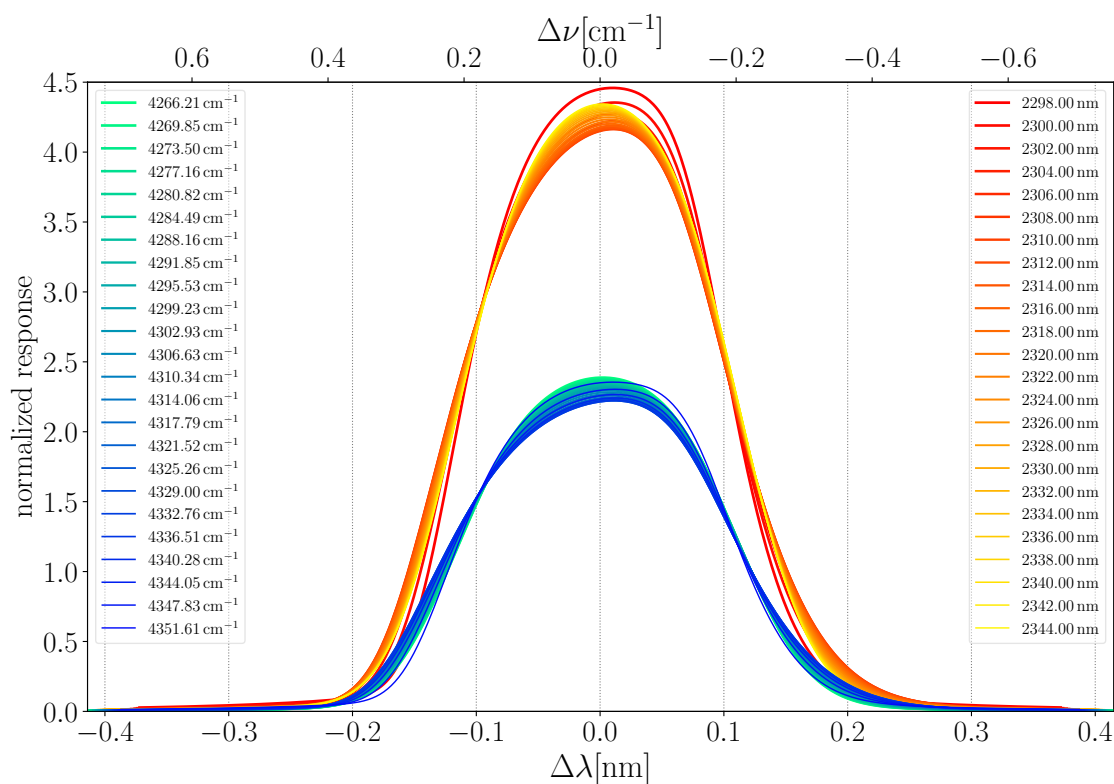

Figure 5. Tabulated responses for ground pixel 128 (out of 256) in TROPOMI's band 7. The instrument spectral response functions (ISRFs) are normalized $\int_{-\infty}^{+\infty} \mathcal{S}_{ij}(x)\,\mathrm{d}x = 1$ in the wavelength ($x = \lambda$) and wavenumber ($x = \nu$) domain, respectively.

2.2.3. Atmospheric Input Data

The physical description of a measurement by the forward model requires input for some atmospheric state-variables, such as pressure p, temperature T, and specific humidity q, since, e.g., the cross section of molecules k_m need to be calculated at different atmospheric levels in order to accurately model lbl absorption through the atmosphere.

Note that BIRRA [25] in Section 2.2.2 utilizes a separable least squares where the state-vector $\mathbf{x}$ is separated into two vectors $\boldsymbol{\eta} \subset \mathbf{x}$ and $\boldsymbol{\beta} \subset \mathbf{x}$ comprising the linear and nonlinear parameters [70],

and that initial guess values are only required for the nonlinear parameters. In the forward model, the 'true' optical depth of a molecule τ_m is described as

$$\tau_m(\nu) \;=\; \alpha_m \, \tau_m^{\mathrm{ref}}(\nu) \;=\; \alpha_m \int\limits_{\substack{\mathrm{double}\\\mathrm{path}}} n_m^{\mathrm{ref}}(s) \, k_m(\nu, s) \, \mathrm{d}s \, , \tag{1}$$

with $\alpha_m \in \boldsymbol{\eta}$ and n_m^{ref} the initial guess molecular number density. In Figure 6, the initial guess for CO and CH$_4$ mole fractions are shown. Both resemble AIRS (Atmospheric Infrared Sounder) version 6 initial guess profiles [71,72] with varying concentrations from the northern hemisphere to the southern hemisphere. The AIRS CO initial guess comes from MOZART (Model for OZone And Related chemical Tracers [73]) monthly mean hemispheric profiles, while CH$_4$ is described by a function of latitude and altitude. Pressure and temperature, as well as the specific humidity, were taken from the 4-times daily reanalysis product [74] maintained by the National Center for Environmental Prediction (NCEP).
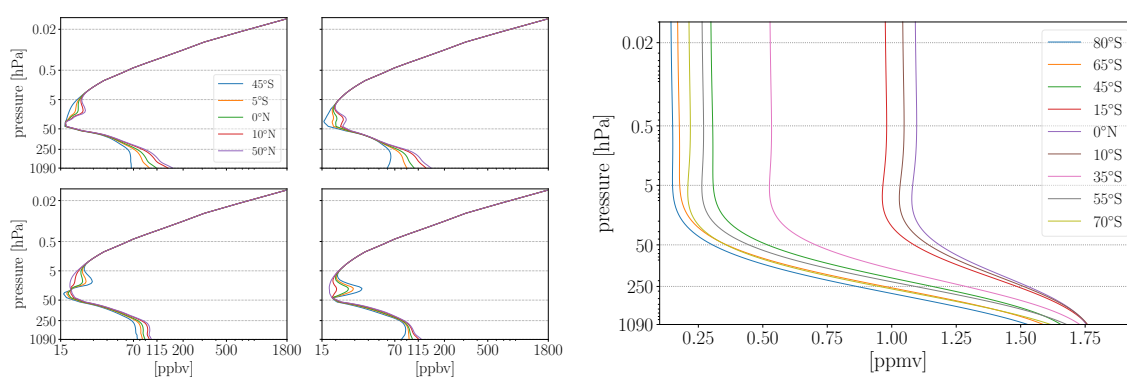

Figure 6. Initial guess mole fraction profiles for CO (four panels on the left) and CH$_4$ (right) resembling the Atmospheric Infrared Sounder (AIRS) version 6 first guess. The CO profiles are provided with monthly granularity for both hemispheres while only January, April, July and October (clockwise from top-left) are shown here.

2.2.4. Cloud Filtering and Topographic Information

Auxiliary input data on clouds was obtained from the S5P-NPPC product which is available for the TROPOMI bands 3, 6, and 7 (SWIR). The Visible Infrared Imaging Radiometer Suite (VIIRS [75]) aboard the Suomi-NPP (Suomi National Polar-Orbiting Partnership [76,77]) spacecraft leads ahead the S5P in loose formation orbit by 3.5 min in local time ascending node and reports cloud information with high spatial resolution on nominal and scaled TROPOMI field of views (FOV). The cloud mask data is grouped in four classes, namely confidently cloudy, probably cloudy, probably clear, and confidently clear. Prior to the CO retrieval the cloud fraction was calculated for the 1.5 scaled FOV of TROPOMI's band 7 (in the along and across-track dimension), and the ratio of the 'confident' and 'probable' classifications was formed. Besides the retrieved CH$_4$ absorption, those quantities serve to identify conditions that might lead to errors in the retrieved columns due to path modifications by clouds and aerosols (scattering) of the observed light, which is not yet considered in the forward model. In particular, an observation was rejected if the cloud fraction specified in the S5P-NPPC product exceeds 10% or if the number of VIIRS pixels that fall into the 'probable' classification (i.e., not the 'confident' classification) exceeds 20%. These rigorous filter criteria avoid observations with large retrieval inaccuracies caused by scattered photons [78] and minimize any bias that arises from changes in the retrieval's vertical sensitivity by modifications in the column averaging kernel (CAK [23], Section 5) in Figure 7 (also see Reference [28], Section 3, and Figure 4 [79] and Figure 3).

Furthermore, the calculation of the double-path transmission between the reflection point (e.g., Earth's surface) and observer and between Sun and reflection point (see Equation (1)) requires

topographic information on terrain elevation. Therefore, the ETOPO global relief model [80] with 2-min grid spacing and an adequate vertical and horizontal datum provides elevation data for each TROPOMI observation in the radiative transfer calculation.

2.3. Vertical Sensitivity and Relation to Priors

In the context of profile retrievals, the sensitivity of the inversion process to the true atmospheric state is given by the averaging kernel. Nonetheless, column density retrievals also have some altitude dependent sensitivity, i.e., the perturbation of elements in the state-vector **x** at different altitudes result in a non-uniform retrieval response [81] (Sections 2 and 3).

Figure 7 shows the altitude sensitivity for three elements of BIRRA's state-vector. The vertical sensitivity for the target gas and CH_4 reveals to be close to unity across the full range of TROPOMI observer zenith angles (OZA), while the CAK of H_2O tends towards zero at higher altitudes, where the retrieval is less sensitive to the true atmospheric state.
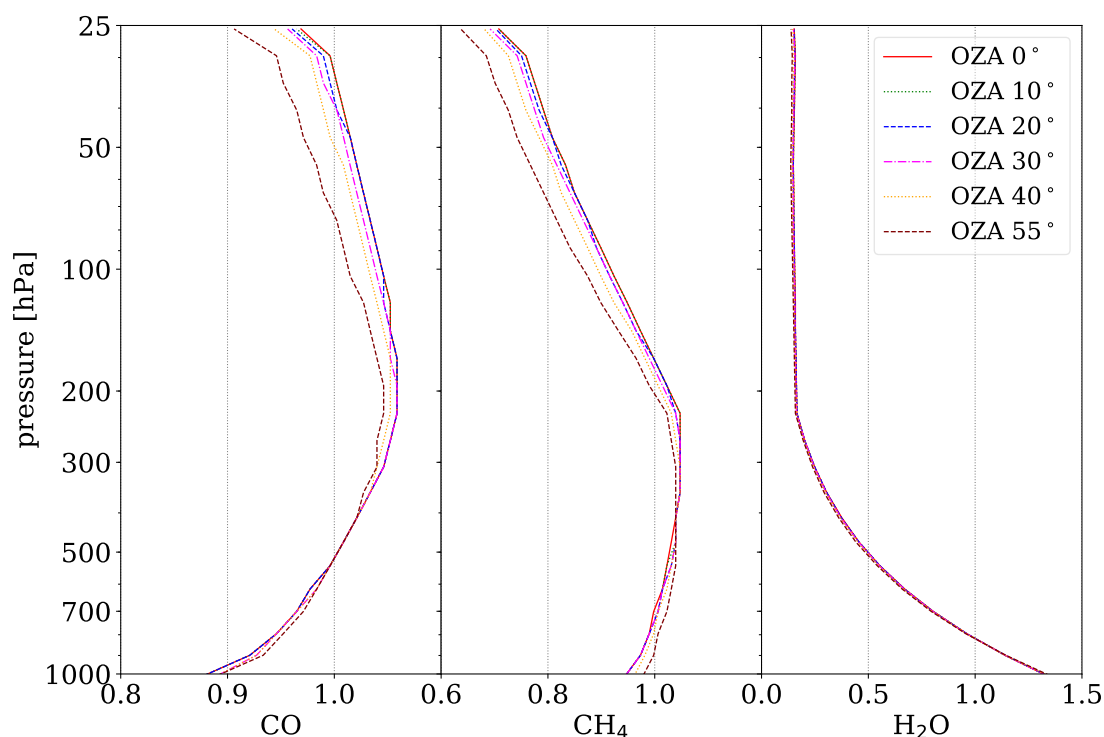

Figure 7. Altitude sensitivity of the CO retrieval in the 4277.20–4302.90 cm^{-1} spectral range. The column averaging kernels (CAKs) were calculated for different TROPOMI observer zenith angles (OZAs) with a solar zenith angle (SZA) of $30°$ and a surface albedo of $r_0 = 0.3$.

2.4. Assessing the Quality of the Fit

Standard diagnostics are used to assess deficiencies in the forward model of the retrieval. In particular, the scaled 2-norm of the spectral fitting residual

$$\rho = \mathbf{I}_{obs} - \mathbf{I}(\mathbf{x}), \tag{2}$$

$$\sigma^2 = ||\rho||^2 / (m - n) \tag{3}$$

is a suitable criterion as it becomes smaller the better the forward model $\mathbf{I}(\mathbf{x})$ can mimic the measurements $\mathbf{I}_{obs}$. Note that m is designating the number of available TROPOMI measurements in our fitting window and is dependent on the BDPM (see Section 3). n is the number of elements in the

state-vector **x** which is constant and in the setup for this study includes the molecular scaling factors $\boldsymbol{\alpha} \subseteq \boldsymbol{\eta}$ and a second-degree polynomial in wavenumber $\mathbf{r} \subseteq \boldsymbol{\beta}$ representing the surface reflectivity.

Furthermore, it is important to note that the absolute value of σ^2 is dependent on the SNR and since no normalization is applied, strictly speaking, only residuals from the same observation should be compared against each other. Nonetheless, it is deemed permissible to compare the scaled norm of the fitting residual within a region as the environment is quite homogeneous in terms of temperature, humidity, and, even at TROPOMI's spatial resolution, surface reflectance.

Errors in molecular spectroscopy can introduce systematic spectral residuals that in consequence result in larger retrieval errors of the corresponding quantities according to

$$\Xi = \sigma^2 \left(J^{\mathrm{T}} J\right)^{-1}. \tag{4}$$

The least squares covariance matrix Ξ contains the Jacobians $J \equiv \frac{\partial \mathbf{I}}{\partial \mathbf{x}}$ of the fitted parameters and diagonal elements represent the errors of the state-vector components [25,82]. Therefore, besides the analysis of the residual norms in Section 3.1, the fitted state-vector elements $\boldsymbol{\alpha}$, along with their error estimates, are examined in Section 3.2.

3. Results

The investigation was performed for a set of TROPOMI orbits in 2019 that cover various climatological regions, namely the Sahara, Central-Europe, Amazonia, and Siberia (for details, see Table 1). Aside from Central-Europe, the areas were selected according to different pairs of temperature and humidity values, and Europe contains strong anthropogenic sources (cities, large harbors, airports, etc.), as well as CO background levels (rural areas, many alpine regions, etc.). The individual orbits are given in Table 1 and were selected based on (low) cloud coverage. Nonetheless, post-processing steps include rigorous cloud filtering, the removal of non-converged retrievals, and the disposal of measurements with very small SNRs (e.g., observations above large bodies of water, such as lakes, rivers, etc.).

3.1. Spectral Fitting Residuals

The elements of the residual vectors according to Equation (2) are depicted in Figure 8. The histograms are separated by regions, starting with Sahara at the top-left and depicting Central-Europe, Siberia, and Amazonia in the clockwise direction. The residuals are following a normal distribution (except for Sahara) with an expected value around zero, indicating that the majority of the measurement errors are caused by random errors, such as instrument noise, etc. This is crucial in order to get the so-called best linear unbiased estimate for the state-vector **x** [83]. In the Sahara region, however, the distribution of residuals deviate from the Gaussian form in particular around the center of the curve and most significantly for G15. The non-uniform distribution over the Sahara is reduced for the SDRM retrievals. Note that the SEOM-Voigt case only considers spectral parameters that describe the mechanisms of pressure and Doppler broadening and was included in order to discriminate the impact of line data versus model.

Figure 9 shows one residual vector for a randomly picked measurement per region. It reveals that the modeled spectra for both H16 and G15 exhibit the largest disagreements in a spectral region close to 4295 cm^{-1}. This feature is significantly reduced when using SEOM line data and virtually eliminated when the SDRM line profile is applied (see Figure 2), as well causing a rather uniform distribution of the residuals across wavenumbers. In addition, G15 reveals some discrepancy round 4293 cm^{-1}. Note that spectral ranges with increased differences show the same positive or negative deviations across geographic regions, indicating that the radiance (transmission) is persistently over- or underestimated for those wavenumbers.

In Figure 10, the individual detector-pixel residuals are examined for all measurements across seasons. The average of the absolute differences is given by $E(|\rho|)$. It shows that the pixels close to

$4295\,\text{cm}^{-1}$ consistently exhibit major disagreements and that the retrievals over the Sahara (first panel) show the largest discrepancies on average since the absolute values of the elements in ρ are dependent on the SNR (see Section 2.4). Furthermore, in order to identify molecular transitions that possibly cause the discrepancies, optical depths of the absorbing molecules CO, CH_4, and H_2O are depicted in three separate but aligned panels below. It becomes obvious that the disagreements around $4295\,\text{cm}^{-1}$ coincide with three rather strong and overlapping absorption lines that were shown in Figure 2. SDRM and H16, particularly, do not agree in that part of the spectrum, and the overall spectral fit quality is improved by $\approx 4-8\%$ for SDRM retrievals.
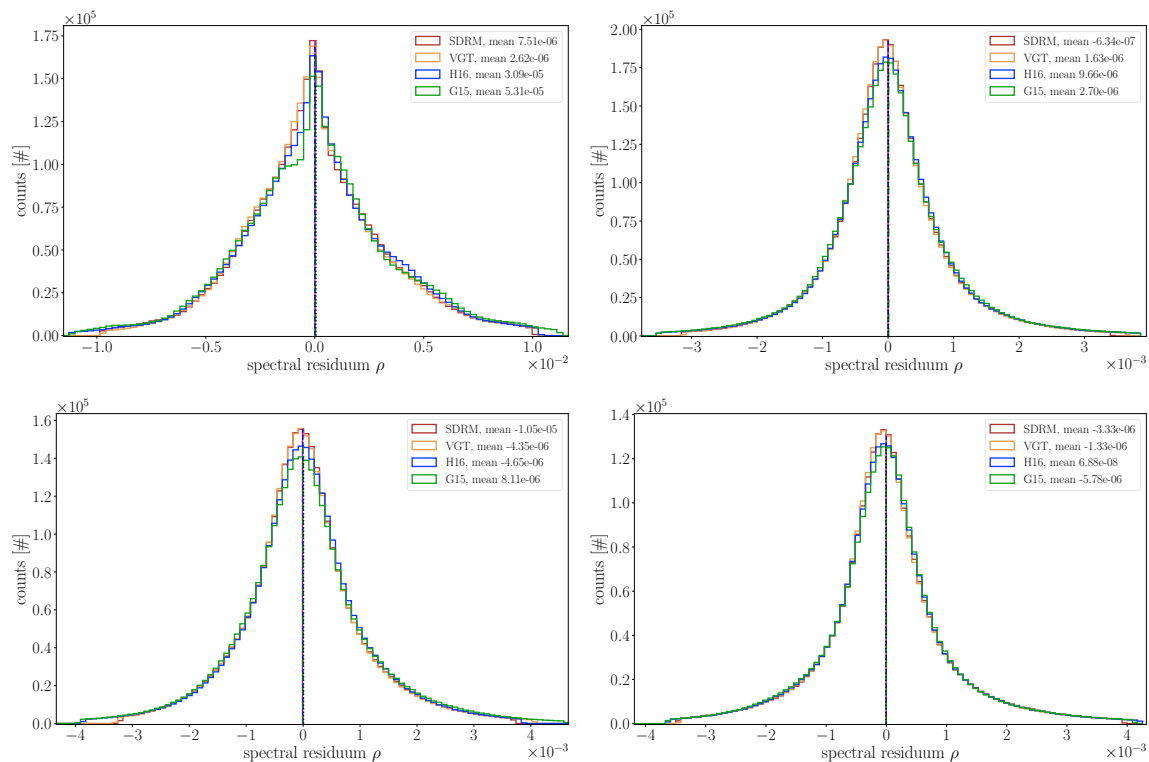
Figure 8. Distribution of the elements of the residual vectors for a single overpass per season in 2019 (Table 1). The histograms are shown per region, i.e. Sahara (top-left), Central-Europe (top-right), Amazonia (lower-left), and Siberia (lower-right). VGT represents cases for which the Voigt profile was calculated with SEOM line data input (SEOM-Voigt).

The substitution of H16 line data with SDRM for each molecule individually shows that the impact on the fit quality is small when CO is replaced ($<0.5\%$) but improves considerably when H_2O ($>7\%$) and CH_4 ($>5\%$) is updated. The given numbers are averages across all investigated regions. Since the residuals for G15 are similar with respect to H16 (except for some pixels between 4280–$4285\,\text{cm}^{-1}$ and around $4293\,\text{cm}^{-1}$ as indicated in Figure 9), they are not depicted in Figure 10.

The averages of the scaled norm σ^2 are itemized by region and season in Table 1. The results confirm above-mentioned findings, i.e., the retrievals using SEOM line data effectively cause smaller discrepancies to TROPOMI observations. More precisely, SDRM-based retrievals reduce σ^2 with respect to H16 and G15 by approximately 10–15% and 15–20%, respectively. The smaller disagreements between SDRM and H16 are likely attributed to the fact that both are not completely independent sets of line data and some of the updates from SEOM are already included in H16 [58].
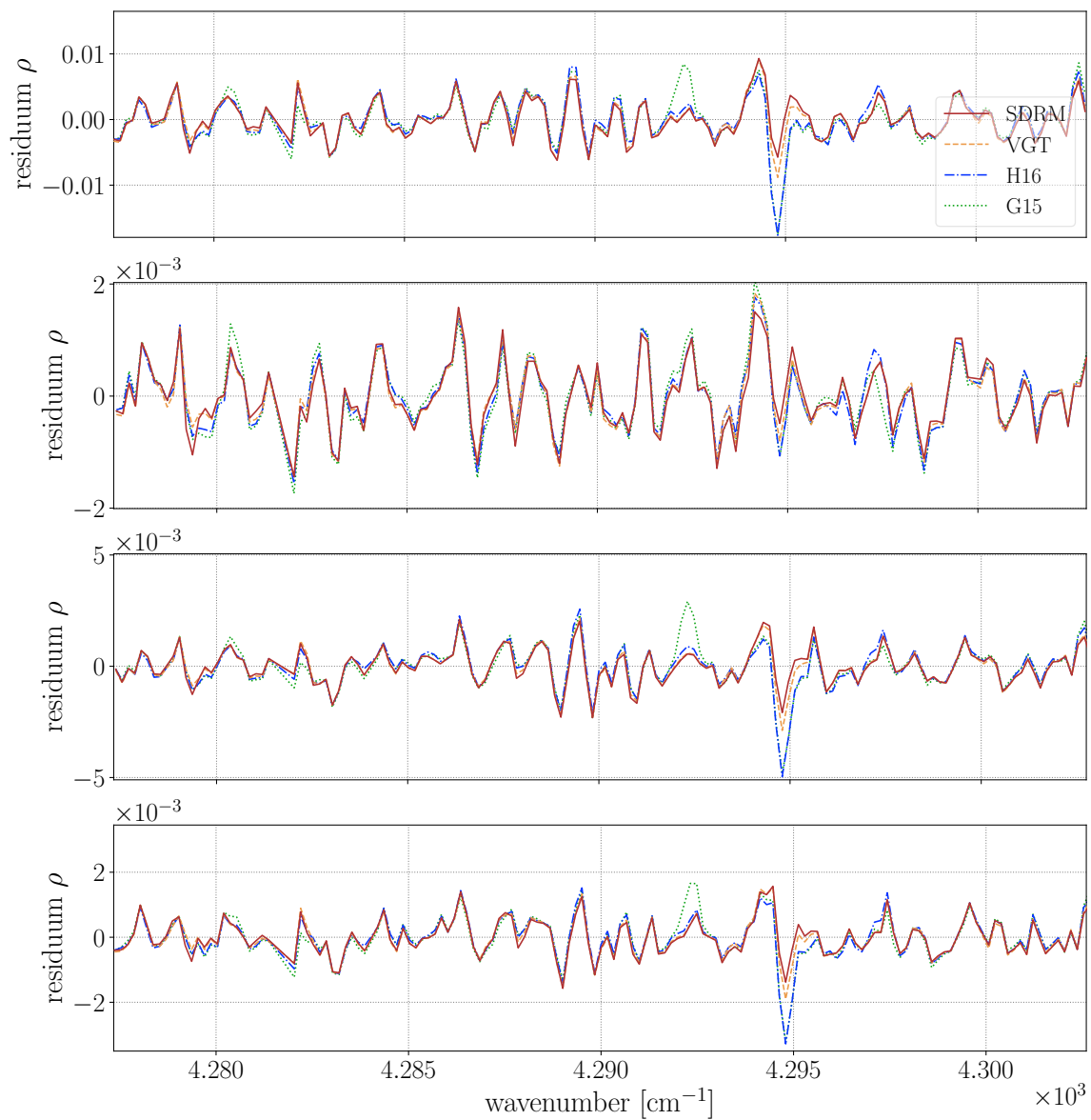
Figure 9. Spectral residuals for a measurement over the Sahara, Central-Europe, Amazonia, and Siberia (from top to bottom) in spring 2019, respectively. The observations were randomly picked; however, the largest disagreements appear to be at similar wavenumbers across the different regions, most prominent close to 4295 cm^{-1} and particularly pronounced in the fits with the two Voigt spectroscopies.
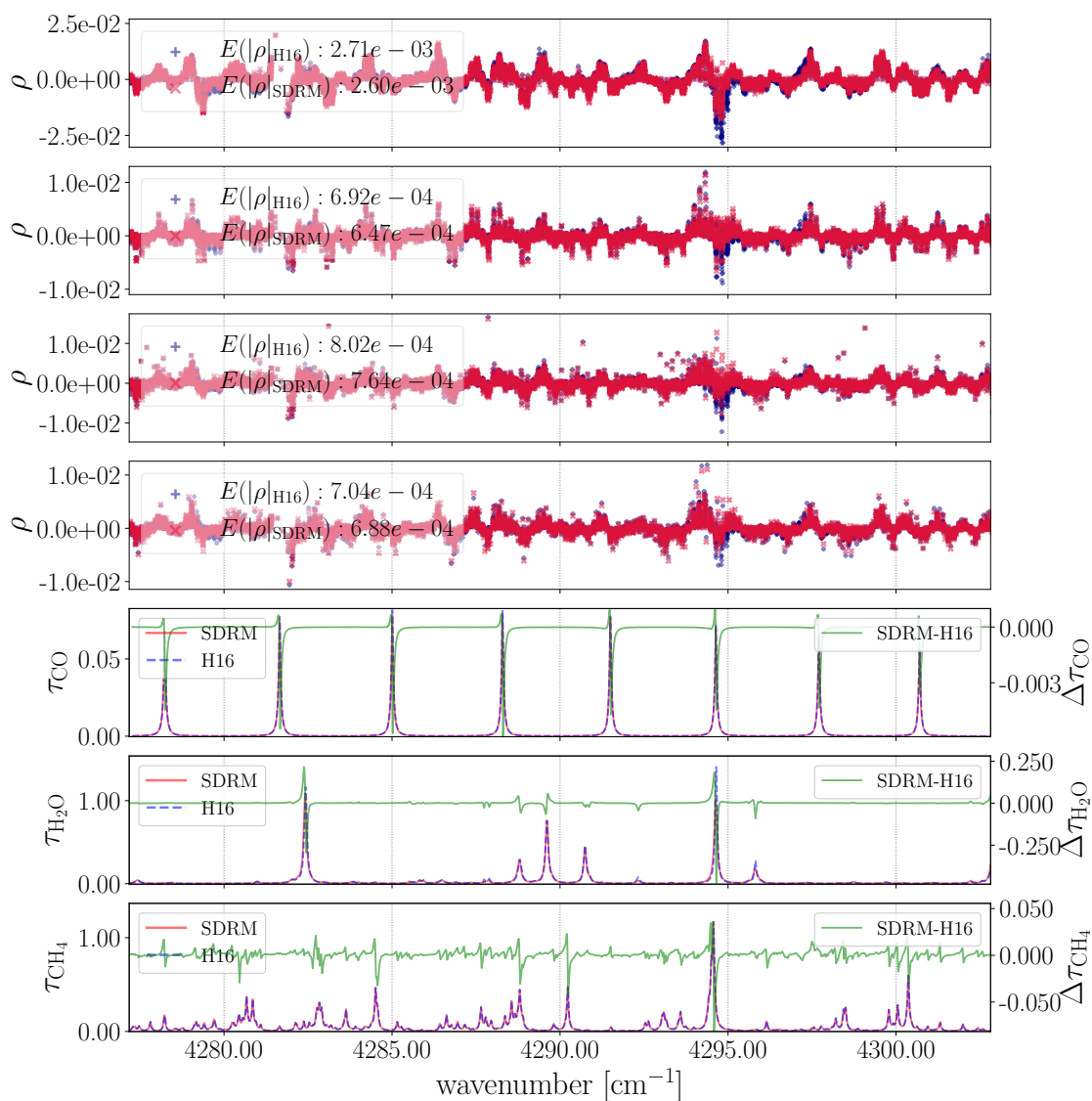
Figure 10. Spectral fitting residuals for TROPOMI observations in the four regions (same order top to bottom as in Figure 9) from the earlier mentioned subset in 2019. Large residuals in the upper four panels reveal to be co-located with strong (and overlapping) absorption lines depicted in the lower three panels. However, similar to Figure 9, the overlapping transitions close to 4295 cm^{-1} causes rather less significant fitting errors in the SDRM retrievals. Moreover, compared to H16, the average fitting residual is smaller for SDRM in all of the examined regions.

Table 1. Average of the scaled norm of the fitting residuals $E(\sigma^2)$ for individual TROPOMI overpasses in 2019.

Region	Input Data	Spring	Summer	Fall	Winter	Total
Sahara	Orbit	6967	7861	8812	10,542	
	G15	$1.93 \cdot 10^{-5}$	$2.69 \cdot 10^{-5}$	$2.40 \cdot 10^{-5}$	$8.65 \cdot 10^{-6}$	$1.97 \cdot 10^{-5}$
	H16	$1.79 \cdot 10^{-5}$	$2.52 \cdot 10^{-5}$	$2.20 \cdot 10^{-5}$	$7.63 \cdot 10^{-6}$	$1.82 \cdot 10^{-5}$
	SDRM	$1.65 \cdot 10^{-5}$	$2.18 \cdot 10^{-5}$	$1.74 \cdot 10^{-5}$	$6.24 \cdot 10^{-6}$	$1.55 \cdot 10^{-5}$
Central-Europe	Orbit	6967	7861	8811	10,542	
	G15	$5.93 \cdot 10^{-7}$	$8.28 \cdot 10^{-7}$	$7.56 \cdot 10^{-7}$	$6.89 \cdot 10^{-7}$	$7.17 \cdot 10^{-7}$
	H16	$5.95 \cdot 10^{-7}$	$7.70 \cdot 10^{-7}$	$5.72 \cdot 10^{-7}$	$5.92 \cdot 10^{-7}$	$6.32 \cdot 10^{-7}$
	SDRM	$4.92 \cdot 10^{-7}$	$7.11 \cdot 10^{-7}$	$5.70 \cdot 10^{-7}$	$5.65 \cdot 10^{-7}$	$5.84 \cdot 10^{-7}$
Amazonia	Orbit	7581	8517	9553	10,347	
	G15	$2.29 \cdot 10^{-6}$	$1.16 \cdot 10^{-6}$	$2.93 \cdot 10^{-6}$	$2.80 \cdot 10^{-6}$	$2.30 \cdot 10^{-6}$
	H16	$2.14 \cdot 10^{-6}$	$9.04 \cdot 10^{-7}$	$2.74 \cdot 10^{-6}$	$2.57 \cdot 10^{-6}$	$2.09 \cdot 10^{-6}$
	SDRM	$1.88 \cdot 10^{-6}$	$9.00 \cdot 10^{-7}$	$2.36 \cdot 10^{-6}$	$2.29 \cdot 10^{-6}$	$1.86 \cdot 10^{-6}$
Siberia	Orbit	7348	8231	9093	9958	
	G15	$1.45 \cdot 10^{-6}$	$2.11 \cdot 10^{-6}$	$2.94 \cdot 10^{-6}$	$1.25 \cdot 10^{-6}$	$1.94 \cdot 10^{-6}$
	H16	$1.33 \cdot 10^{-6}$	$1.93 \cdot 10^{-6}$	$2.81 \cdot 10^{-6}$	$1.15 \cdot 10^{-6}$	$1.81 \cdot 10^{-6}$
	SDRM	$1.29 \cdot 10^{-6}$	$1.14 \cdot 10^{-6}$	$2.62 \cdot 10^{-6}$	$1.14 \cdot 10^{-6}$	$1.55 \cdot 10^{-6}$

3.2. Impact on Retrieved Columns and Corresponding Errors

The effect of absorption line data on the retrieved CO total columns $N_m\,[\text{molec cm}^{-2}]$ and corresponding errors is shown in Figures 11–14. Each figure depicts the results for one region and contains the target and co-retrieved quantities according to

$$N_m \;=\; \alpha_m \int_{z_{\text{srf}}}^{z_{\text{TOA}}} n_m^{\text{ref}}(z) \;=\; \alpha_m\, N_m^{\text{ref}}. \tag{5}$$

The mole fractions for some molecular number densities n_m^{ref} are shown in Figure 6. It is important to note that H_2O and CH_4, although the latter has strong absorption lines across the CO spectral fitting window and can be used to identify light path modifications [78], are byproducts primarily considered due to their spectral interference with the target gas. The distribution of the errors for CH_4 and H_2O is shown in Appendix A.

For the Sahara region depicted in Figure 11, the majority of CO is distributed between $1.0 - 2.5 \cdot 10^{18}\,\text{molec cm}^{-2}$ and the histograms for SDRM and HITRAN are similar. The CO retrieval errors according to Equation (4) are illustrated in the top-right and are below $1.0 \cdot 10^{17}\,\text{molec cm}^{-2}$ across spectroscopies with the median around $1.7 \cdot 10^{16}\,\text{molec cm}^{-2}$ in case of SDRM and around $2.2 \cdot 10^{16}\,\text{molec cm}^{-2}$ for the other two line lists. Note that the majority of the CO errors for SDRM are even below $6.0 \cdot 10^{16}\,\text{molec cm}^{-2}$. Although there is almost no absolute difference in the medians of the CO columns for SDRM and H16 distributions, they were found to be significantly different according to the non-parametric Kolmogorov-Smirnoff test (p-value $< 1.0 \cdot 10^{-5}$) [84]. While this holds true for SDRM and G15 distributions (CO difference $\approx 1.0 \cdot 10^{17}\,\text{molec cm}^{-2}$), the magnitude of their retrieval errors is significantly different, as well.

A similar CO distribution is observed for Central-Europe in Figure 12, but values cover a greater range. Again, SDRM and H16 cause similar concentrations, while G15 is significantly different based on the distribution of errors. Moreover, the SDRM-based CO product is, again, the most precise, i.e., incorporates the smallest fitting residuals and thus retrieval errors.
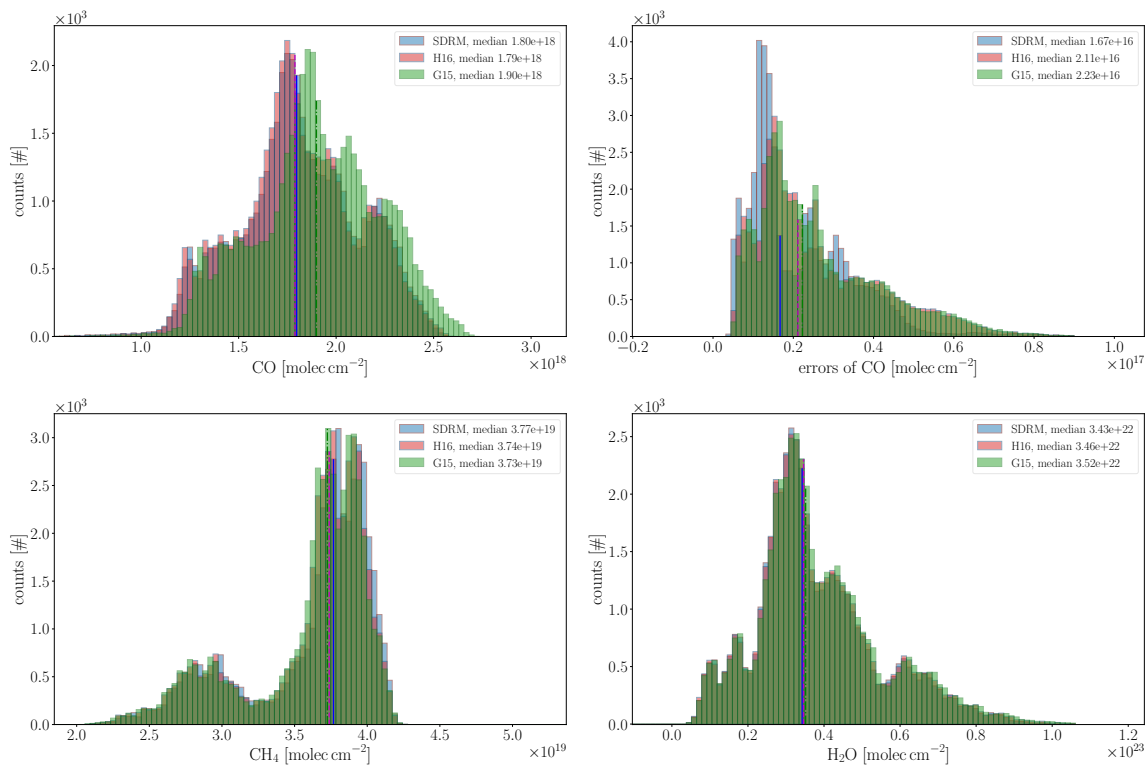
Figure 11. Distribution of retrieved quantities for Sahara. The CO columns and corresponding errors for different spectroscopic inputs and models are shown in the top-row. Histograms for the co-retrieved CH_4 and H_2O columns are depicted in the two panels below. The median of each distribution is shown in the corresponding color.
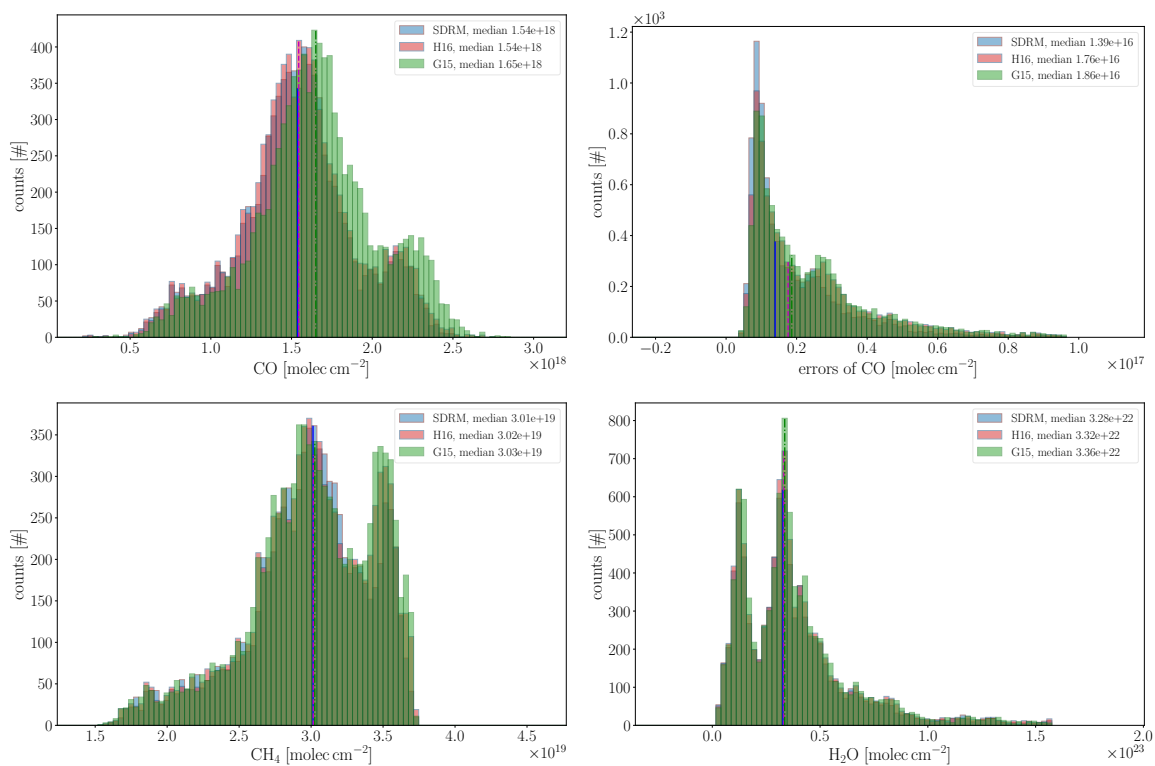

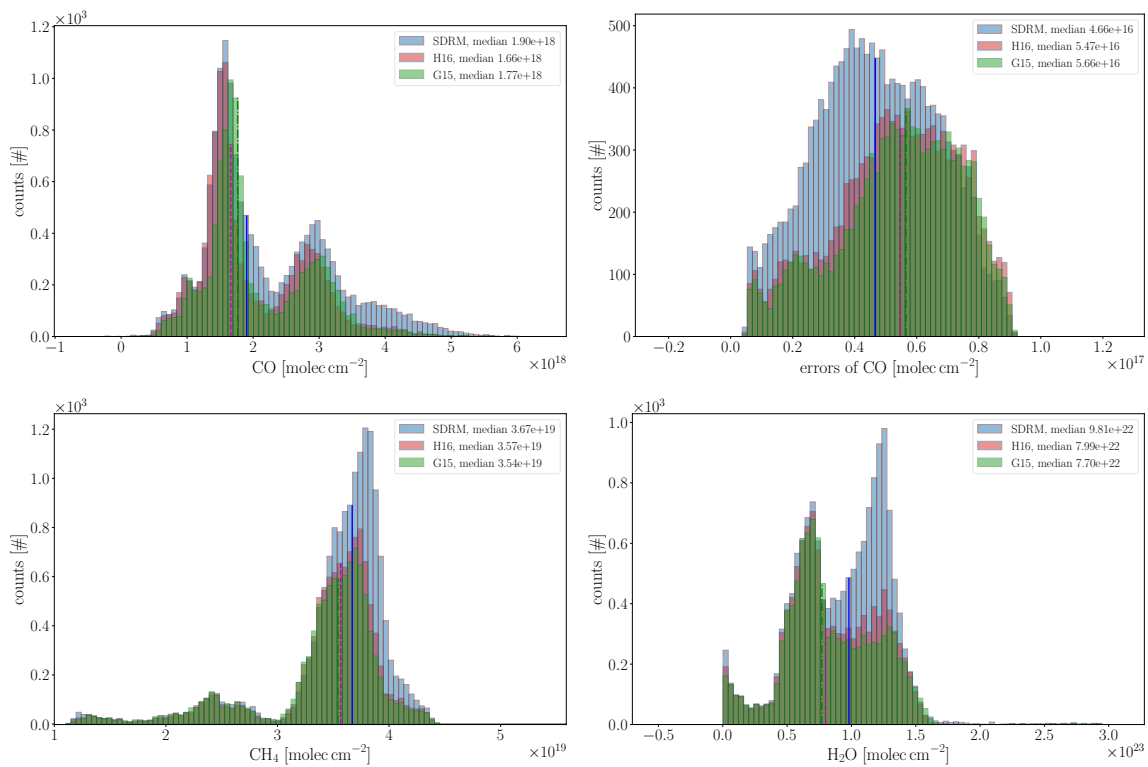**Figure 12.** Distribution of the retrieved values and errors for Central-Europe.

Figure 13. Distribution for the Amazonia region which include an overpass during the heavy forest fires in summer 2019 (orbit 9553 on 17 August 2019).
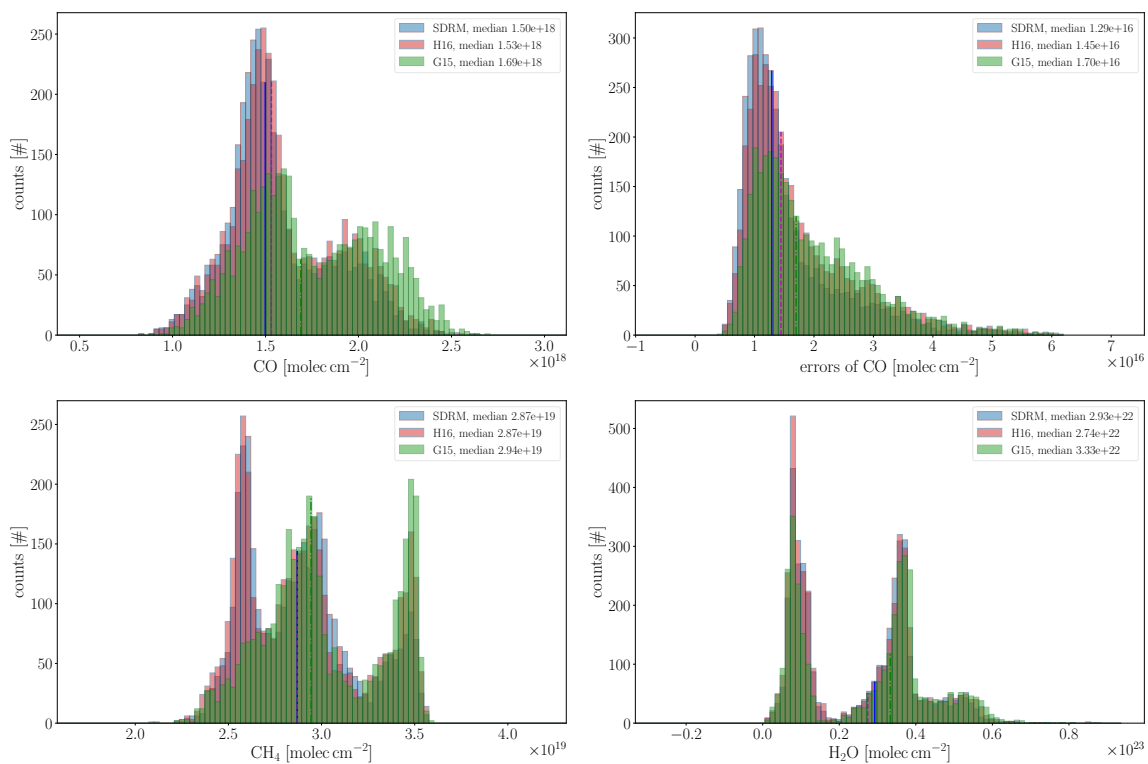
Figure 14. Distributions over Siberia.

In Figure 13, the CO columns over Amazonia show a rather different distribution where SDRM retrievals cause significantly higher concentrations. The difference is also significant for the two co-retrieved molecules. The errors in the top-left are larger compared to the two previous regions;

however, SDRM-based errors are, again, the smallest. The difference in CO for both SDRM-H16 and SDRM-G15 was found to be $2.4 \cdot 10^{17}$ and $1.3 \cdot 10^{17}$ molec cm^{-2}, respectively. The errors are consistently $<1.0 \cdot 10^{17}$ molec cm^{-2} which indicates that the differences in SDRM retrieved CO concentrations over Amazonia are significant with respect to both other line lists, suggesting that the 'true' value might be outside the specified error range of the product. Furthermore, it should be noted that the S5P-overpass in the second quarter

The distribution of CO over Siberia in Figure 14 is similar to that for Sahara and Central-Europe, i.e., the histograms for SDRM and HITRAN almost resemble each other, while GEISA is shifted towards higher columns on average. Again, this shift is larger than the product error and thus considered to be significant. In contrast to the other regions, the CH$_4$ columns for G15 are also shifted. For H$_2$O, the medians are rather equally spaced across cases as it is true for the errors.

3.3. CO *over Amazonia and Central-Europe*

The trend of differences between SDRM and H16-based CO mole fractions over Amazonia and Europe is depicted in Figure 15. The standard set of filter criteria described in Section 3 was applied on a mostly sunny, high pressure influenced day over Central-Europe on 21 September 2019 (orbit 10046) and a day with rather average cloud coverage over Amazonia on 17 August 2019 (orbit 9553). Note that only the cloud filter was applied for the calculation of the differences since filtering on errors, etc.,would create different sized datasets with even less measurements available for subtraction.
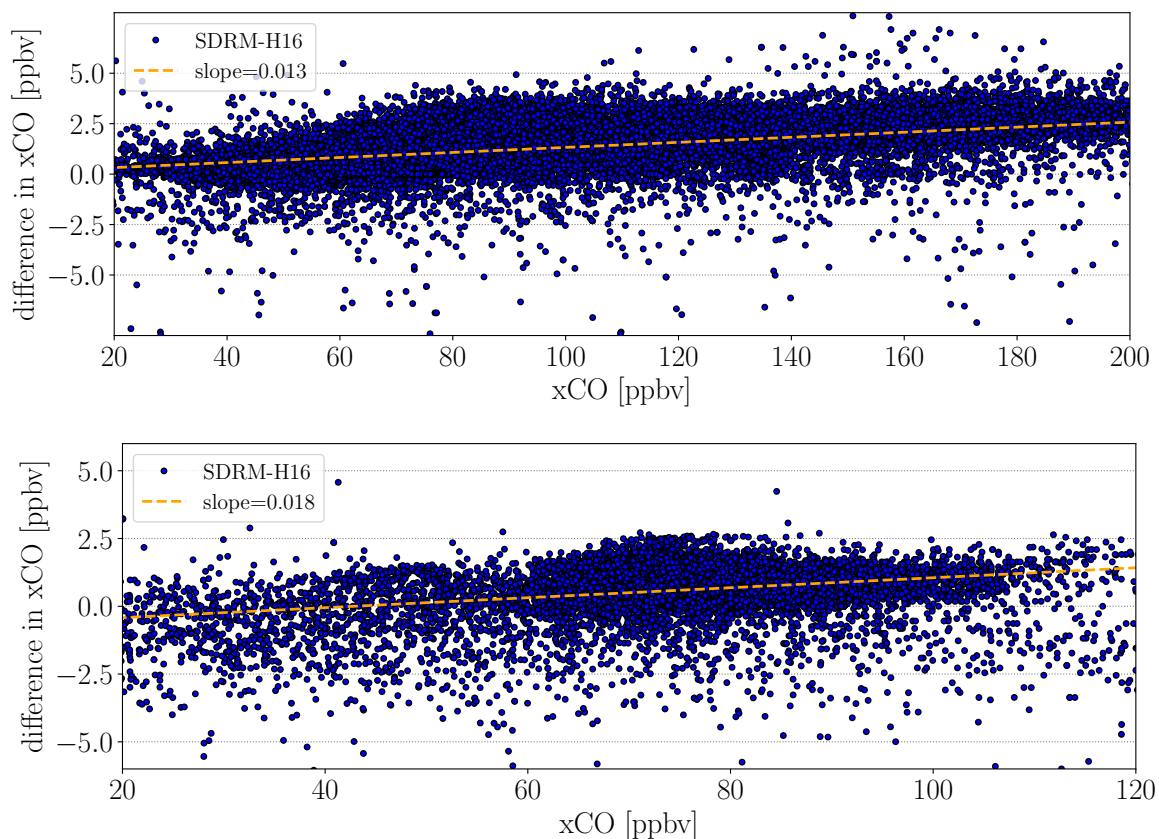

Figure 15. Differences in the retrieved CO mole fractions as a function of the CO concentrations over Amazonia (top) and Central-Europe (bottom).

Over both regions, the differences tend to be larger for higher CO mole fractions, although the trend is within the error of most observations (see second column of Figure 16). While the SDRM retrieval errors for background CO concentrations range between ≈0.5–1.5 ppbv, the errors over

elevated concentrations are somewhat larger, though the relative error is rather similar for either case. The spatial distribution of the CO differences is depicted in the third column of Figure 16, while SDRM-based mole fractions are depicted in the first column. Particularly, the difference plot over Europe reveals a striping pattern in the satellites' along track direction. It is a well-known but not yet understood feature of push-broom spectrometers that is changing from orbit to orbit [58].
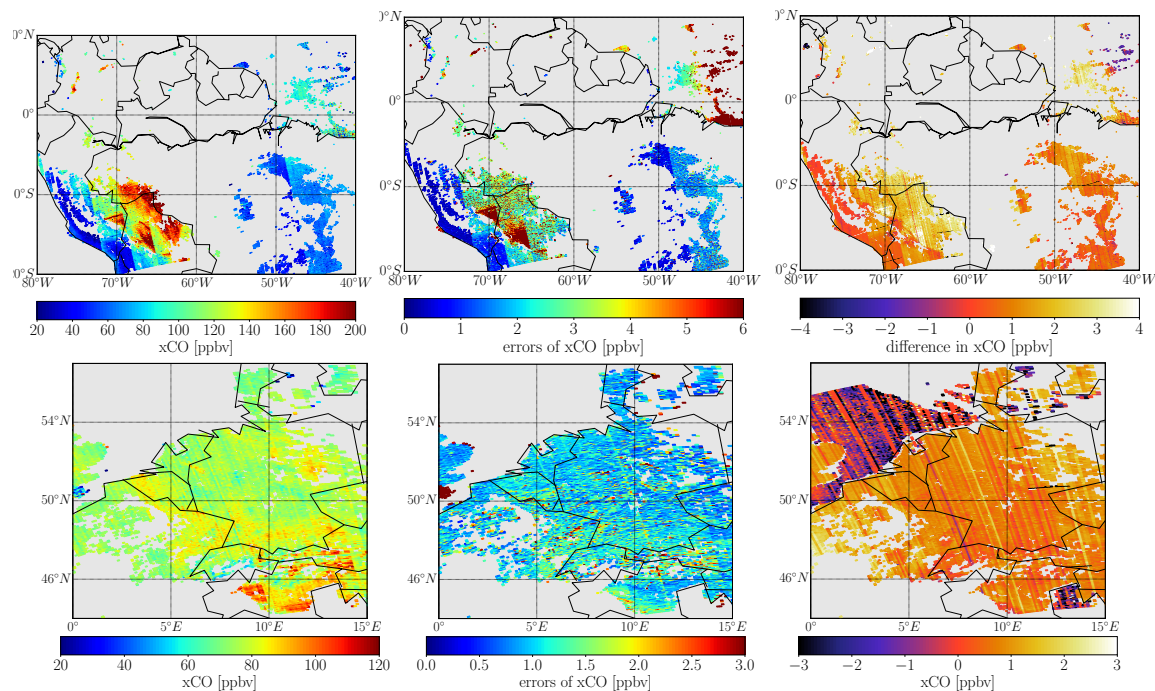

Figure 16. CO mole fractions (left column) and errors (center column) for SDRM-based retrievals, along with differences between SDRM and H16 (SDRM-H16) over Amazonia (top-row) and Central-Europe (bottom-row).

3.4. Comparison to Ground-Based Observations

In order to estimate the quality of the retrieval product, a comparison with observations from co-located TCCON (Total Column Carbon Observing Network [85,86]) and NDACC (Network for the Detection of Atmospheric Composition Change) ground-based (g-b) measurements was carried out.

Filtering of the TROPOMI retrievals is crucial in order to compare valid CO mole fractions to g-b references. However, it is important to note that since the errors in the retrievals are dependent on the spectroscopy, filter criteria based on errors might be appropriate for one retrieval (e.g., SDRM-based) but not the other (e.g., H16-based). The values given in Table 2 are mean and median values for the respective TROPOMI overpass on the specified day. Since SDRM inferred CO columns exhibit smaller errors (see Section 3.2), one set of strict filter criteria leads to different numbers of observations remaining for comparison after post-processing. The actual number of remaining measurements after filtering is primarily dependent on the weather conditions at the time of overpass and surface characteristics around a station. If no TROPOMI observations remained within a reasonable small radius around the g-b station after filtering (e.g., <50 km), the radius for co-location was increased in steps of 50 km up to 200 km. The mean value was calculated for measurements that remained after strict filtering (taking retrieval errors, etc., into account), while the median was computed for the non-filtered retrieval output (i.e., only filtered for clouds). By only rejecting cloudy pixels using the S5P-NPPC product (see Section 2.2.4), both retrievals deliver the same number of CO mole fractions after post-processing but include observations with large errors.

The comparison to g-b reference observations in Figure 17 shows that differences vary across sites. Although no consistent over- or underestimation of BIRRA retrieved CO mole fractions from

TROPOMI is obvious, most TCCON sites observe larger values. In accordance with results from Section 3.3, the validation shows that SDRM-based retrievals cause larger columns on average (also see Reference [59] Figure 9).

Table 2. Daily mean and median values for the SDRM- and H16-based CO mole fractions from TROPOMI measurements compared to Total Column Carbon Observing Network (TCCON) and Network for the Detection of Atmospheric Composition Change (NDACC) g-b observations. 'Non-filtered' specifies that only cloudy TROPOMI pixels were eliminated, while 'Filtered' additionally considers retrieval errors in the post-processing steps. 'Radius' designates the maximum distance for co-location, i.e., only TROPOMI observations from within that distance were compared to the g-b site. Values in brackets designate the number of observations after post-processing.

| | Ground-Based | | | TROPOMI | | | | | |
| | | | | Filtered | | | Non-Filtered | | |
Station	Date	Mean [ppbv]	Median [ppbv]	Mean [ppbv] SDRM	H16	Radius [km]	Median [ppbv] SDRM	H16	Radius [km]
Bremen (T)	08/05/18	86.32	86.35	92.82 (11)	90.81 (3)	50	91.93	91.05	30
Bremen (N)	11/10/18	79.77	79.82	79.26 (23)	77.29 (4)	50	78.57	77.58	20
Edwards (T)	01/07/19	85.11	87.20	88.62 (7)	86.99 (3)	50	94.40	95.07	50
Garmisch (T)	20/09/19	77.97	77.90	76.66 (6)	73.15 (4)	20	75.21	74.64	20
Karlsruhe (T)	20/09/19	77.83	78.00	84.68 (13)	84.34 (12)	15	86.19	85.40	15
Paris (T)	30/06/18	80.56	80.60	76.89 (21)	76.31 (8)	200	76.31	75.44	200
Park Falls (T)	13/06/19	77.42	77.10	81.77 (19)	84.18 (6)	150	80.40	79.44	150
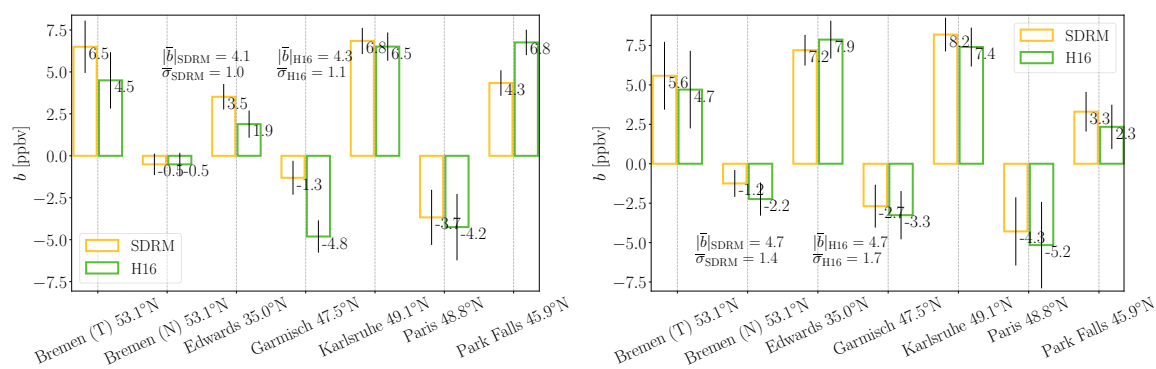
Figure 17. Differences of TROPOMI and TCCON/NDACC mean (left) and median (right) CO mole fractions from Table 2. Note that the mean was calculated upon the strictly filtered retrieval output, while the median includes all cloud-free retrievals.

4. Discussion

4.1. Spectral Residuals

The results in Section 3.1 indicate that the SEOM line data has positive impact on the retrieval by causing significantly smaller residuals on average. Furthermore, the impact on the CO fitting residuals underlines results from previous studies [58] and demonstrates that improvements in spectroscopy are important, particularly for retrievals from measurements with sufficient SNR and high spectral resolution [59].

4.2. CO Mole Fractions

As shown in Section 3.2, the number of measurements where differences in CO mole fractions become significant is considerably dependent on the specific sets of line data compared. The results in Section 3.3 indicate that the increase in retrieved CO is within the error bar of most observations, particularly with respect to H16. Nonetheless, Figures 15 and 16 also demonstrate that discrepancies can become significant over selected scenes, preferably for measurements over elevated CO

concentrations. A comparison to Reference [59] in Figure 7 shows that CO retrievals from SCIAMACHY using SEOM spectroscopy also exhibit a tendency towards larger values. Moreover, the outcome in Section 3.3 indicates that the ratio of change in CO to its errors is roughly proportional across most geographic areas between SCIAMACHY and TROPOMI.

4.3. Validation

Since g-b observed CO columns can vary considerably throughout a day (>15 ppbv), representation errors, such as spatial and temporal mismatch, should be taken into account when interpreting differences for a single TROPOMI overpass. Due to advection of CO over time (e.g., at $3\,\text{ms}^{-1}$), TROPOMI observations at some distance from the g-b station do also provide a valuable source of information in the comparison and should be considered, as well.

It is also important to note that TCCON data is calibrated to World Meteorological Organization (WMO) in situ trace gas measurement scales in order to tie its observations to in situ measurements [86,87]. The systematic difference between CO from TCCON and observations that are not tied to the WMO scale was examined in detail by Kiel et al. [88]. They found that this correction factor (1.0672 for the GGG2014 dataset) is the main source of the observed difference (also see Reference [62]) and that the choice of different spectroscopic line lists have only minor influence on the overall bias. This is in accordance with the results in Section 3.4, although different spectroscopies were compared.

5. Summary and Conclusions

The investigation on the impact of molecular spectroscopy on CO total columns from TROPOMI SWIR observations found that SEOM line data with the adequate model improves the spectral fit quality by reducing the residuals to TROPOMI measurements with respect to both H16 and G15.

The results demonstrate that molecular spectroscopy has a significant effect on the precision of the CO retrieval. The reduced spectral fitting residuals and smaller retrieval errors were found to be statistically significant across the examined regions, making the SDRM-based CO product more precise. The magnitude of the impact is dependent on the climatological region and spectroscopic reference but ranges from $\approx$ 10–20% (up to 30% for individual observations with respect to G15). Updates in the H_2O and CH_4 cross sections were identified to be the main reason for the improved fit quality. These findings underline the important role that accurate spectroscopic information plays in meeting the missions' requirements.

In contrast to the fitting residuals, the differences in CO columns between SDRM and H16 were found to be rather small across most regions ($\leq$ 3%), while some larger discrepancies were found for individual observations with elevated molecular concentrations, particularly over Amazonia. Similar to the spectral residuals, the average disagreements to SDRM are larger for G15, with the largest differences to SDRM-based retrievals found over Siberia. In the other two examined regions, the impact is less significant with respect to H16 but stays significant for the majority of G15-based retrievals.

The comparison to TCCON and NDACC g-b observations revealed that the smaller retrieval errors in the SDRM inferred columns are beneficial when comparing post-processed CO mole fractions to g-b references since stricter filter criteria can be applied on the TROPOMI observations within a given distance from the station.

Overall, many aspects of the findings underline recommendations from earlier investigations [41,42] and are in good agreement with similar conclusions from Hochstaffl and Schreier [59].

Author Contributions: P.H. developed the methods, tools, framework and strategy for this study and performed all retrievals. F.S. originally designed the forward model and retrieval algorithm. M.B. and G.W. compiled the SEOM–IAS database. The original draft of the manuscript was prepared by P.H. and then reviewed and commented by F.S., M.B., G.W. and D.G.F. The TCCON data listed in Table 2 was provided by J.N., R.S. and Y.T. (details see Table A1). All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are frequently used in this manuscript:

BIRRA	Beer Infrared Retrieval Algorithm
CRDS	Cavity Ring-Down Spectroscopy
FTS	Fourier Transform Spectrometer
FWHM	Full Width Half Maximum
GARLIC	Generic Atmospheric Radiation Line-by-line Infrared Code
G15	Gestion et Etude des Informations Spectroscopiques Atmosphériques 2015 (GEISA 2015)
H16	High Resolution Transmission 2016 (HITRAN 2016)
HT	Hartmann-Tran
HWHM	Half Width Half Maximum
lbl	line-by-line
NDACC	Network for the Detection of Atmospheric Composition Change
NIR	Near InfraRed
Py4CAtS	PYthon scripts for Computational ATmospheric Spectroscopy
S5P	Sentinel-5 Precursor
SCIAMACHY	Scanning Imaging Absorption SpectroMeter for Atmospheric CHartographY
SEOM	Scientific Exploitation of Operational Missions–Improved Atmospheric Spectroscopy
SDRM	SEOM with Speed-Dependent Rautian and line-Mixing
SNR	Signal-to-Noise Ratio
SWIR	ShortWave InfraRed
TCCON	Total Column Carbon Observing Network
TOA	Top Of Atmosphere
TROPOMI	TROPOspheric Monitoring Instrument
UVIS	Ultraviolet and VISible
VGT	SEOM with Voigt

Appendix A. Errors of the Retrieved Quantities

The subsequent Figures A1 and A2 show the errors of the corresponding CH_4 and H_2O columns from Section 3.2.
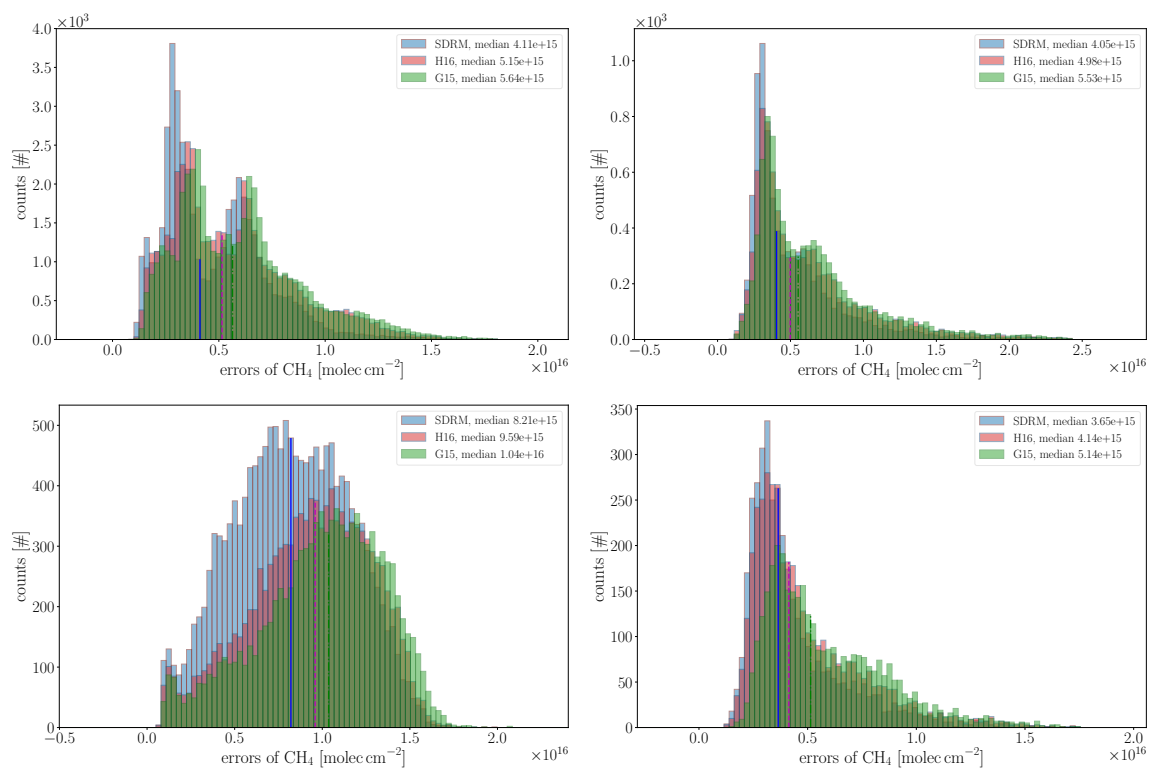
Figure A1. Errors of the CH$_4$ columns for Sahara, Central-Europe, Amazonia and Siberia (clockwise from top-left).
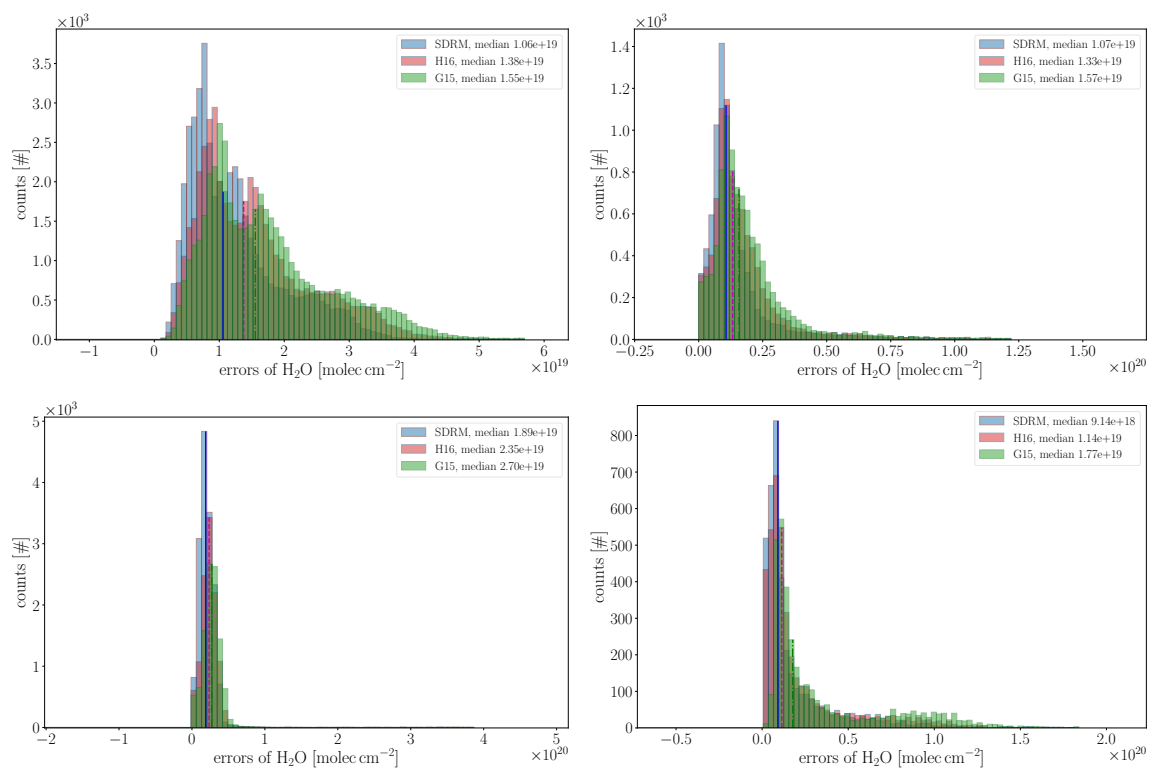

Figure A2. Errors of H$_2$O columns for Sahara, Central-Europe, Amazonia and Siberia (clockwise from top-left).

Appendix B. TCCON Data Providers

The TCCON data were obtained from sites listed in Table A1. The TCCON Data Archive is hosted by CaltechDATA, California Institute of Technology, CA (US), doi: https://doi.org/10.14291/tccon.archive/1348407.

Table A1. Total Carbon Column Observing Network (TCCON) sites with the references to the data used in this publication.

TCCON Site	Reference
Bremen	Notholt et al. [89]
Edwards	Iraci et al. [90]
Garmisch	Sussmann and Rettinger [91]
Karlsruhe	Hase et al. [92]
Paris	Té et al. [93]
Park Falls	Wennberg et al. [94]

Appendix C. NDACC Data Providers

The data used in this publication were obtained from from sites listed in Table A2 and are available through the NDACC website http://www.ndacc.org.

Table A2. NDACC site used in this publication, along with the station's principle investigator (PI) and cooperating institution.

NDACC Site	Station PI and Coorperating Institutions
Bremen	Prof. Dr. Justus Notholt
	Institute of Environmental Physics; University of Bremen, Germany

References

1. Struve, W.S. *Fundamentals of Molecular Spectroscopy*; Wiley-Interscience: New York, NY, USA, 1989; p. 379.
2. Clerbaux, C.; Hadji-Lazaro, J.; Turquety, S.; Mégie, G.; Coheur, P.F. Trace gas measurements from infrared satellite for chemistry and climate applications. *Atm. Chem. Phys.* **2003**, *3*, 1495–1508. [CrossRef]
3. Burrows, J.P.; Platt, U.; Borrell, P. *The Remote Sensing of Tropospheric Composition from Space*; Physics of Earth and Space Environment; Springer: Berlin/Heidelberg, Germany, 2011. [CrossRef]
4. Veefkind, J.; Aben, I.; McMullan, K.; Förster, H.; de Vries, J.; Otter, G.; Claas, J.; Eskes, H.; de Haan, J.; Kleipool, Q.; et al. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote. Sens. Environ.* **2012**, *120*, 70–83. [CrossRef]
5. Burrows, J.P.; Weber, M.; Buchwitz, M.; Rozanov, V.; Ladstätter-Weißenmayer, A.; Richter, A.; DeBeek, R.; Hoogen, R.; Bramstedt, K.; Eichmann, K.U.; et al. The Global Ozone Monitoring Experiment (GOME): Mission Concept and First Scientific Results. *J. Atmos. Sci.* **1999**, *56*, 151–175, doi:10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2. [CrossRef]
6. Munro, R.; Lang, R.; Klaes, D.; Poli, G.; Retscher, C.; Lindstrot, R.; Huckle, R.; Lacan, A.; Grzegorski, M.; Holdak, A.; et al. The GOME-2 instrument on the Metop series of satellites: instrument design, calibration, and level 1 data processing—An overview. *Atmos. Meas. Tech.* **2016**, *9*, 1279–1301. [CrossRef]
7. Bovensmann, H.; Burrows, J.; Buchwitz, M.; Frerick, J.; Noël, S.; Rozanov, V.; Chance, K.; Goede, A. SCIAMACHY: Mission Objectives and Measurement Mode. *J. Atmos. Sci.* **1999**, *56*, 127–150. [CrossRef]
8. Gottwald, M.; Bovensmann, H. (Eds.) *SCIAMACHY — Exploring the Changing Earth's Atmosphere*; Springer: Dordrecht, The Netherlands, 2011; doi:10.1007/978-90-481-9896-2. [CrossRef]
9. Levelt, P.; Oord, G.; Dobber, M.; Mälkki, A.; Visser, H.; Vries, J.; Stammes, P.; Lundell, J.; Saari, H. The Ozone Monitoring Instrument. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1093–1101. [CrossRef]
10. Vonk, F. *Input/Output Data Specification for the TROPOMI L01b Data Processor*; Technical Report; The Royal Netherlands Meteorological Institute KNMI: Utrecht, The Netherlands, 2017.

11. Kobayashi, H. Line–by–Line Calculation using Fourier–transformed Voigt Function. *J. Quant. Spectrosc. Radiat. Transfer* **1999**, *62*, 477–483. [CrossRef]

12. Deeter, M.N.; Emmons, L.K.; Francis, G.L.; Edwards, D.P.; Gille, J.C.; Warner, J.X.; Khattatov, B.; Ziskin, D.; Lamarque, J.F.; Ho, S.P.; et al. Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument. *J. Geophys. Res.* **2003**, *108*, 4399. [CrossRef]

13. Turquety, S.; Hadji-Lazaro, J.; Clerbaux, C.; Hauglustaine, D.; Clough, T.; Cassé, V.; Schlüssel, P.; Mégie, G. Operational trace gas retrieval algorithm for the Infrared Atmospheric Sounder Interferometer. *J. Geophys. Res.* **2004**, *109*, D21301 [CrossRef]

14. McMillan, W.; Barnet, C.; Strow, L.; Chahine, M.; McCourt, M.; Warner, J.; Novelli, P.; Korontzi, S.; Maddy, E.; Datta, S. Daily global maps of carbon monoxide from NASA's Atmospheric Infrared Sounder. *Geophys. Res. Lett.* **2005**, *32*, L11801. [CrossRef]

15. Rinsland, C.; Luo, M.; Logan, J.; Beer, R.; Worden, H.; Worden, J.; Bowman, K.; Kulawik, S.; Rider, D.; Osterman, G.; et al. Nadir Measurements of Carbon Monoxide Distributions by the Tropospheric Emission Spectrometer onboard the Aura Spacecraft: Overview of Analysis Approach and Examples of Initial Results. *Geophys. Res. Lett.* **2006**, *33*, L22806. [CrossRef]

16. Lu, Y.; Khalil, M. Methane and carbon monoxide in OH chemistry: The effects of feedbacks and reservoirs generated by the reactive products. *Chemosphere* **1993**, *26*, 641–655. [CrossRef]

17. Holloway, T.; Levy, H., II; Kasibhatla, P. Global distribution of carbon monoxide. *J. Geophys. Res.* **2000**, *105*, 12123–12147. [CrossRef]

18. Feilberg, K.L.; Sellevåg, S.R.; Nielsen, C.J.; Griffith, D.W.T.; Johnson, M.S. $CO + OH \rightarrow CO_2 + H$: The relative reaction rate of five CO isotopologues. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4687–4693. [CrossRef]

19. Daniel, J.S.; Solomon, S. On the climate forcing of carbon monoxide. *J. Geophys. Res.* **1998**, *103*, 13249–13260. [CrossRef]

20. Levy, H. Normal Atmosphere: Large Radical and Formaldehyde Concentrations Predicted. *Science* **1971**, *173*, 141–143. [CrossRef]

21. Crutzen, P.J.; Zimmermann, P.H. The changing photochemistry of the troposphere. *Tellus A* **1991**, *43*, 136–151. [CrossRef]

22. Buchwitz, M.; Rozanov, V.; Burrows, J. A near-infrared optimized DOAS method for the fast global retrieval of atmospheric CH_4, CO, CO_2, H_2O, and N_2O total column amounts from SCIAMACHY Envisat-1 nadir radiances. *J. Geophys. Res.* **2000**, *105*, 15231–15245. [CrossRef]

23. Buchwitz, M.; de Beek, R.; Bramstedt, K.; Noël, S.; Bovensmann, H.; Burrows, J.P. Global carbon monoxide as retrieved from SCIAMACHY by WFM-DOAS. *Atm. Chem. Phys.* **2004**, *4*, 1945–1960. [CrossRef]

24. Gloudemans, A.; Schrijver, H.; Hasekamp, O.; Aben, I. Error analysis for CO and CH_4 total column retrievals from SCIAMACHY 2.3 μm spectra. *Atm. Chem. Phys.* **2008**, *8*, 3999–4017. [CrossRef]

25. Gimeno García, S.; Schreier, F.; Lichtenberg, G.; Slijkhuis, S. Near infrared nadir retrieval of vertical column densities: methodology and application to SCIAMACHY. *Atmos. Meas. Tech.* **2011**, *4*, 2633–2657. [CrossRef]

26. Borsdorff, T.; Hasekamp, O.P.; Wassmann, A.; Landgraf, J. Insights into Tikhonov regularization: Application to trace gas column retrieval and the efficient calculation of total column averaging kernels. *Atmos. Meas. Tech.* **2014**, *7*, 523–535. [CrossRef]

27. Landgraf, J.; aan de Brugh, J.; Scheepmaker, R.A.; Borsdorff, T.; Houweling, S. Hasekamp, O. *Algorithm Theoretical Baseline Document for Sentinel-5 Precursor: Carbon Monoxide Total Column Retrieval*; Technical Report; SRON Netherlands Institute for Space Research: Utrecht, The Netherlands, 2018.

28. Landgraf, J.; aan de Brugh, J.; Scheepmaker, R.; Borsdorff, T.; Hu, H.; Houweling, S.; Butz, A.; Aben, I.; Hasekamp, O. Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements. *Atmos. Meas. Tech.* **2016**, *9*, 4955–4975. [CrossRef]

29. Schneising, O.; Buchwitz, M.; Reuter, M.; Bovensmann, H.; Burrows, J.P.; Borsdorff, T.; Deutscher, N.M.; Feist, D.G.; Griffith, D.W.T.; Hase, F.; et al. A scientific algorithm to simultaneously retrieve carbon monoxide and methane from TROPOMI onboard Sentinel-5 Precursor. *Atmos. Meas. Tech.* **2019**, *12*, 6771–6802. [CrossRef]

30. Schreier, F.; Gimeno García, S.; Hochstaffl, P.; Städt, S. Py4CAtS—PYthon for Computational ATmospheric Spectroscopy. *Atmosphere* **2019**, *10*, 262. [CrossRef]

31. Armstrong, B. Spectrum Line Profiles: The Voigt Function. *J. Quant. Spectrosc. Radiat. Transfer* **1967**, *7*, 61–88. [CrossRef]

32. Boone, C.; Walker, K.; Bernath, P. Speed–dependent Voigt profile for water vapor in infrared remote sensing applications. *J. Quant. Spectrosc. Radiat. Transfer* **2007**, *105*, 525–532. [CrossRef]

33. Schneider, M.; Hase, F. Improving spectroscopic line parameters by means of atmospheric spectra: Theory and example for water vapor and solar absorption spectra. *J. Quant. Spectrosc. Radiat. Transfer* **2009**, *110*, 1825–1839. [CrossRef]

34. Schneider, M.; Hase, F.; Blavier, J.F.; Toon, G.; Leblanc, T. An empirical study on the importance of a speed-dependent Voigt line shape model for tropospheric water vapor profile remote sensing. *J. Quant. Spectrosc. Radiat. Transfer* **2011**, *112*, 465–474. [CrossRef]

35. Tennyson, J.; Bernath, P.; Campargue, A.; Császár, A.; Daumont, L.; Gamache, R.; Hodges, J.; Lisak, D.; Naumenko, O.; Rothman, L.; et al. Recommended isolated-line profile for representing high-resolution spectroscopic transitions (IUPAC Technical Report). *Pure Appl. Chem.* **2014**, *86*, 1931–1943. [CrossRef]

36. Birk, M.; Wagner, G. Voigt profile introduces optical depth dependent systematic errors—Detected in high resolution laboratory spectra of water. *J. Quant. Spectrosc. Radiat. Transfer* **2016**, *170*, 159–168. [CrossRef]

37. Hartmann, J.M.; Tran, H.; Armante, R.; Boulet, C.; Campargue, A.; Forget, F.; Gianfrani, L.; Gordon, I.; Guerlet, S.; Hodges, J.T. Recent advances in collisional effects on spectra of molecular gases and their practical consequences. *J. Quant. Spectrosc. Radiat. Transfer* **2018**, *213*, 178–227. [CrossRef]

38. Frankenberg, C.; Warneke, T.; Butz, A.; Aben, I.; Hase, F.; Spietz, P.; Brown, L.R. Pressure broadening in the $2\nu_3$ band of methane and its implication on atmospheric retrievals. *Atm. Chem. Phys.* **2008**, *8*, 5061–5075. [CrossRef]

39. Frankenberg, C.; Bergamaschi, P.; Butz, A.; Houweling, S.; Meirink, J.F.; Notholt, J.; Petersen, A.K.; Schrijver, H.; Warneke, T.; Aben, I. Tropical methane emissions: A revised view from SCIAMACHY onboard ENVISAT. *Geophys. Res. Lett.* **2008**, *35*, L15811. [CrossRef]

40. Scheepmaker, R.A.; Frankenberg, C.; Galli, A.; Butz, A.; Schrijver, H.; Deutscher, N.M.; Wunch, D.; Warneke, T.; Fally, S.; Aben, I. Improved water vapour spectroscopy in the 4174–4300 cm^{-1} region and its impact on SCIAMACHY HDO/H$_2$O measurements. *Atmos. Meas. Tech.* **2013**, *6*, 879–894. [CrossRef]

41. Galli, A.; Butz, A.; Scheepmaker, R.A.; Hasekamp, O.; Landgraf, J.; Tol, P.; Wunch, D.; Deutscher, N.M.; Toon, G.C.; Wennberg, P.O.; et al. CH$_4$, CO, and H$_2$O spectroscopy for the Sentinel-5 Precursor mission: An assessment with the Total Carbon Column Observing Network measurements. *Atmos. Meas. Tech.* **2012**, *5*, 1387–1398. [CrossRef]

42. Checa-García, R.; Landgraf, J.; Galli, A.; Hase, F.; Velazco, V.; Tran, H.; Boudon, V.; Alkemade, F.; Butz, A. Mapping spectroscopic uncertainties into prospective methane retrieval errors from Sentinel-5 and its precursor. *Atmos. Meas. Tech.* **2015**, *8*, 3617–3629. [CrossRef]

43. Birk, M.; Wagner, G.; Loos, J.; Mondelain, D.; Campargue, A. ESA SEOM–IAS—Measurement database—2.3 µm region [Data set]. Zenodo. 2017. Available online: https://zenodo.org/record/1009122#.X5Iw-FARVhE (accessed on 27 August 2020)

44. Birk, M.; Wagner, G.; Loos, J.; Mondelain, D.; Campargue, A. ESA SEOM–IAS—Spectroscopic parameters database—2.3 µm region [Data set]. Zenodo. 2017. Available online: https://figshare.com/articles/ESA_SEOM-IAS_Measurement_database_2_3_m_region/6632627/1 (accessed on 27 August 2020)

45. Loos, J.; Birk, M.; Wagner, G. Measurement of positions, intensities and self-broadening line shape parameters of H$_2$O lines in the spectral ranges 1850–2280 cm^{-1} and 2390–4000 cm^{-1}. *J. Quant. Spectrosc. Radiat. Transfer* **2017**, *203*, 119–132. [CrossRef]

46. Ngo, N.; Lisak, D.; Tran, H.; Hartmann, J.M. An isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes. *J. Quant. Spectrosc. Radiat. Transfer* **2013**, *129*, 89–100. [CrossRef]

47. Ngo, N.; Lisak, D.; Tran, H.; Hartmann, J.M. Erratum to n isolated line-shape model to go beyond the Voigt profile in spectroscopic databases and radiative transfer codes [J. Quant. Spectrosc. Radiat. Transf. 129 (2013) 89 00]. *J. Quant. Spectrosc. Radiat. Transfer* **2014**, *134*, 105. [CrossRef]

48. Tran, H.; Ngo, N.; Hartmann, J.M. Efficient computation of some speed-dependent isolated line profiles. *J. Quant. Spectrosc. Radiat. Transfer* **2013**, *129*, 199–203. [CrossRef]

49. Tran, H.; Ngo, N.; Hartmann, J.M. Erratum to efficient computation of some speed-dependent isolated line profiles[J. Quant. Spectrosc. Radiat. Transfer 129 (2013) 199 03]. *J. Quant. Spectrosc. Radiat. Transf.* **2014**, *134*, 104. [CrossRef]

50. Rosenkranz, P. Shape of the 5 mm oxygen band in the atmosphere. *IEEE Trans. Antennas Propag.* **1975**, *23*, 498–506. [CrossRef]

51. Smith, E.W. Absorption and dispersion in the O_2 microwave spectrum at atmospheric pressures. *J. Chem. Phys.* **1981**, *74*, 6658–6673. [CrossRef]

52. Dicke, R.H. The Effect of Collisions upon the Doppler Width of Spectral Lines. *Phys. Rev.* **1953**, *89*, 472–473. [CrossRef]

53. Varghese, P.; Hanson, R. Collisional narrowing effects on spectral line shapes measured at high resolution. *Appl. Opt.* **1984**, *23*, 2376–2385. [CrossRef]

54. Boone, C.; Walker, K.; Bernath, P. An efficient analytical approach for calculating line mixing in atmospheric remote sensing applications. *J. Quant. Spectrosc. Radiat. Transfer* **2011**, *112*, 980–989. [CrossRef]

55. Kochanov, V.P. Speed-dependent spectral line profile including line narrowing and mixing. *J. Quant. Spectrosc. Radiat. Transfer* **2016**, *177*, 261–268. [CrossRef]

56. Schreier, F. Computational Aspects of Speed-Dependent Voigt Profiles. *J. Quant. Spectrosc. Radiat. Transfer* **2017**, *187*, 44–53. [CrossRef]

57. Schreier, F.; Hochstaffl, P. Computational Aspects of Speed-Dependent Voigt and Rautian Profiles. *J. Quant. Spectrosc. Radiat. Transfer* **2020**, In press. [CrossRef]

58. Borsdorff, T.; aan de Brugh, J.; Schneider, A.; Lorente, A.; Birk, M.; Wagner, G.; Kivi, R.; Hase, F.; Feist, D.G.; Sussmann, R.; et al. Improving the TROPOMI CO data product: Update of the spectroscopic database and destriping of single orbits. *Atmos. Meas. Tech.* **2019**, *12*, 5443–5455. [CrossRef]

59. Hochstaffl, P.; Schreier, F. Impact of Molecular Spectroscopy on Carbon Monoxide Abundances from SCIAMACHY. *Remote Sens.* **2020**, *12*, 1084. [CrossRef]

60. Gordon, I.; Rothman, L.; Hill, C.; Kochanov, R.; Tan, Y.; Bernath, P.; Birk, M.; Boudon, V.; Campargue, A.; Chance, K.; et al. The HITRAN2016 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transfer* **2017**, *203*, 3–69. [CrossRef]

61. Jacquinet-Husson, N.; Armante, R.; Scott, N.; Chédin, A.; Crépeau, L.; Boutammine, C.; Bouhdaoui, A.; Crevoisier, C.; Capelle, V.; Boonne, C.; et al. The 2015 edition of the GEISA spectroscopic database. *J. Mol. Spectrosc.* **2016**, *327*, 31–72. [CrossRef]

62. Hochstaffl, P.; Schreier, F.; Lichtenberg, G.; Gimeno García, S. Validation of Carbon Monoxide Total Column Retrievals from SCIAMACHY Observations with NDACC/TCCON Ground-Based Measurements. *Remote Sens.* **2018**, *10*, 223. [CrossRef]

63. Schreier, F.; Gimeno García, S.; Hedelt, P.; Hess, M.; Mendrok, J.; Vasquez, M.; Xu, J. GARLIC—A General Purpose Atmospheric Radiative Transfer Line-by-Line Infrared-Microwave Code: Implementation and Evaluation. *J. Quant. Spectrosc. Radiat. Transfer* **2014**, *137*, 29–50. [CrossRef]

64. Schreier, F. The Voigt and Complex Error Function: A Comparison of Computational Methods. *J. Quant. Spectrosc. Radiat. Transfer* **1992**, *48*, 743–762. [CrossRef]

65. Kleipool, Q.; Ludewig, A.; Babić, L.; Bartstra, R.; Braak, R.; Dierssen, W.; Dewitte, P.J.; Kenter, P.; Landzaat, R.; Leloux, J.; et al. Pre-launch calibration results of the TROPOMI payload on-board the Sentinel-5 Precursor satellite. *Atmos. Meas. Tech.* **2018**, *11*, 6439–6479. [CrossRef]

66. Van Kempen, T.A.; van Hees, R.M.; Tol, P.J.J.; Aben, I.; Hoogeveen, R.W.M. In-flight calibration and monitoring of the Tropospheric Monitoring Instrument (TROPOMI) short-wave infrared (SWIR) module. *Atmos. Meas. Tech.* **2019**, *12*, 6827–6844. [CrossRef]

67. Kurucz, R. *Model Atmosphere Codes: ATLAS12 andATLAS9 Intensity*; Harvard-Smithsonian Center for Astrophysics: Cambridge, MA, USA, 2014.

68. Van Hees, R.M.; Tol, P.J.J.; Cadot, S.; Krijger, M.; Persijn, S.T.; van Kempen, T.A.; Snel, R.; Aben, I.; Hoogeveen, W.M. Determination of the TROPOMI-SWIR instrument spectral response function. *Atmos. Meas. Tech.* **2018**, *11*, 3917–3933. [CrossRef]

69. Smeets, J.; Kleipool, Q.; van Hees, R.; Sneep, M. *README for TROPOMI Instrument Spectral Response Functions*; Technical Report; The Royal Netherlands Meteorological Institute KNMI: Utrecht, The Netherlands, 2018.

70. Golub, G.; Pereyra, V. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Problems* **2003**, *19*, R1–R26. [CrossRef]

71. Olsen, E.T.; Warner, J.; Zigang, W.; UMCP. *AIRS/AMSU/HSB Version 6* CO *Initial Guess Profiles (NH & SH)*; Technical Report; Jet Propulsion Laboratory (JPL): Pasadena, CA, USA, 2017.

72. Olsen, E.T.; Xiaozhen, X.; IMSG-NOAA/NESDIS/STAR. *AIRS/AMSU/HSB Version 6 CH₄ Initial Guess Profiles*; Technical Report; Jet Propulsion Laboratory (JPL): Pasadena, CA, USA, 2017.

73. Hauglustaine, D.A.; Brasseur, G.; Walters, S.; Rasch, P.; Müller, J.F.; Emmons, L.; Carroll, M. MOZART: A global chemical transport model for ozone and related chemical tracers. *J. Geophys. Res.* **1998**, *1032*, 28291–28336. [CrossRef]

74. Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; et al. The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Met. Soc.* **1996**, *77*, 437–472. [CrossRef]

75. Cao, C.; De Luccia, F.J.; Xiong, X.; Wolfe, R.; Weng, F. Early On-Orbit Performance of the Visible Infrared Imaging Radiometer Suite Onboard the Suomi National Polar-Orbiting Partnership (S-NPP) Satellite. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 1142–1156. [CrossRef]

76. Cao, C.; Xiong, J.; Blonski, S.; Liu, Q.; Uprety, S.; Shao, X.; Bai, Y.; Weng, F. Suomi NPP VIIRS sensor data record verification, validation, and long-term performance monitoring. *J. Geophys. Res.* **2013**, *118*, 11664–11678. [CrossRef]

77. Siddans, R.; Smith, A. *S5P-NPP Cloud Processor ATBD*; Technical Report; Rutherford Appleton Laboratory (RAL): Chilton, UK, 2018.

78. Scheepmaker, R.A.; aan de Brugh, J.; Hu, H.; Borsdorff, T.; Frankenberg, C.; Risi, C.; Hasekamp, O.; Aben, I.; Landgraf, J. HDO and H₂O total column retrievals from TROPOMI shortwave infrared measurements. *Atmos. Meas. Tech.* **2016**, *9*, 3921–3937. [CrossRef]

79. Vidot, J.; Landgraf, J.; Hasekamp, O.; Butz, A.; Galli, A.; Tol, P.; Aben, I. Carbon monoxide from shortwave infrared reflectance measurements: A new retrieval approach for clear sky and partially cloudy atmospheres. *Remote. Sens. Environ.* **2012**, *120*, 255–266. [CrossRef]

80. National Geophysical Data Center. *2-Minute Gridded Global Relief Data (ETOPO2) v2*; National Geophysical Data Center, NOAA: Boulder, CO, USA, 2006.

81. Rodgers, C.; Connor, B. Intercomparison of remote sounding instruments. *J. Geophys. Res.* **2003**, *108*, 4116. [CrossRef]

82. Gay, D. *Usage Summary for Selected Optimization Routines (PORT Mathematical Subroutine Library, Optimization Chapter)*; Computing Science Technical Report 153; AT&T Bell Laboratories: Murray Hill, NJ, USA, 1990. Available online: http://netlib.bell-labs.com/cm/cs/cstr/153.pdf (accessed on 17 October 2019).

83. Rust, B.W. Fitting nature's basic functions. II. Estimating uncertainties and testing hypotheses. *Comput. Sci. Eng.* **2001**, *3*, 60–64. [CrossRef]

84. Hodges, J.L. The significance probability of the Smirnov two-sample test. *Ark. Mat.* **1958**, *3*, 469–486. [CrossRef]

85. Wunch, D.; Toon, G.C.; Blavier, J.F.L.; Washenfelder, R.A.; Notholt, J.; Connor, B.J.; Griffith, D.W.T.; Sherlock, V.; Wennberg, P.O. The Total Carbon Column Observing Network (TCCON). *Philos. Trans. Roy. Soc. London Ser. A* **2011**, *369*, 2087–2112. [CrossRef]

86. Wunch, D.; Toon, G.C.; Sherlock, V.; Deutscher, N.M.; Liu, C.; Feist, D.G.; Wennberg, P.O. Documentation for the 2014 TCCON Data Release. *CaltechDATA* **2015**. [CrossRef]

87. Wunch, D.; Toon, G.C.; Wennberg, P.O.; Wofsy, S.C.; Stephens, B.B.; Fischer, M.L.; Uchino, O.; Abshire, J.B.; Bernath, P.; Biraud, S.C.; et al. Calibration of the Total Carbon Column Observing Network using aircraft profile data. *Atmos. Meas. Tech.* **2010**, *3*, 1351–1362. [CrossRef]

88. Kiel, M.; Hase, F.; Blumenstock, T.; Kirner, O. Comparison of XCO abundances from the Total Carbon Column Observing Network and the Network for the Detection of Atmospheric Composition Change measured in Karlsruhe. *Atmos. Meas. Tech.* **2016**, *9*, 2223–2239. [CrossRef]

89. Notholt, J.; Petri, C.; Warneke, T.; Deutscher, N.M.; Buschmann, M.; Weinzierl, C.; Macatangay, R.C.; Grupe, P. TCCON Data from Bremen (DE), Release GGG2014.R0. TCCON Data Archive, Hosted by CaltechDATA. 2014. Available online: https://data.caltech.edu/records/268 (accessed on 27 August 2020)

90. Iraci, L.T.; Podolske, J.; Hillyard, P.W.; Roehl, C.; Wennberg, P.O.; Blavier, J.F.; Allen, N.; Wunch, D.; Osterman, G.; Albertson, R. TCCON Data from Edwards (US), Release GGG2014.R1. TCCON Data Archive, Hosted by CaltechDATA. 2016. Available online: https://data.caltech.edu/records/268 (accessed on 27 August 2020)

91. Sussmann, R.; Rettinger, M. TCCON Data from Garmisch (DE), Release GGG2014.R2. TCCON Data Archive, Hosted by CaltechDATA. 2018. Available online: https://data.caltech.edu/records/268 (accessed on 27 August 2020)

92. Hase, F.; Blumenstock, T.; Dohe, S.; Gross, J.; Kiel, M. TCCON Data from Karlsruhe (DE), Release GGG2014.R1. TCCON Data Archive, Hosted by CaltechDATA. 2015. Available online: https://data.caltech.edu/records/268 (accessed on 27 August 2020)

93. Té, Y.; Jeseck, P.; Janssen, C. TCCON Data from Paris (FR), Release GGG2014.R0. TCCON Data Archive, Hosted by CaltechDATA. 2014. Available online: https://data.caltech.edu/records/268 (accessed on 27 August 2020)

94. Wennberg, P.O.; Roehl, C.M.; Wunch, D.; Toon, G.C.; Blavier, J.F.; Washenfelder, R.; Keppel-Aleks, G.; Allen, N.T.; Ayers, J. TCCON Data from Park Falls (US), Release GGG2014.R1. TCCON Data Archive, Hosted by CaltechDATA. 2017. Available online: https://data.caltech.edu/records/268 (accessed on 27 August 2020)